



HAL
open science

**Traitements sémantiques et émotionnels lors de la
perception rapide de scènes visuelles : étude du
fonctionnement normal chez l'adulte et application au
trouble déficitaire de l'attention avec ou sans
hyperactivité**

Anaïs Leroy

► **To cite this version:**

Anaïs Leroy. Traitements sémantiques et émotionnels lors de la perception rapide de scènes visuelles : étude du fonctionnement normal chez l'adulte et application au trouble déficitaire de l'attention avec ou sans hyperactivité. Psychologie. Université Côte d'Azur, 2020. Français. NNT : 2020COAZ2000 . tel-04348080

HAL Id: tel-04348080

<https://theses.hal.science/tel-04348080>

Submitted on 16 Dec 2023

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 DE DOCTORAT

Traitements sémantiques et émotionnels lors de la perception rapide de scènes visuelles :

Etude du fonctionnement normal chez l'adulte et l'enfant
et application au trouble déficitaire de l'attention
avec ou sans hyperactivité

Anaïs LEROY

Laboratoire d'Anthropologie et de Psychologie Cliniques, Cognitives et
Sociales

**Présentée en vue de l'obtention
du grade de docteur en Psychologie
de l'Université Côte d'Azur**
Dirigée par : Sylvane Faure
Co-dirigée par : Sara Spotorno
Soutenue le : 15 janvier 2020

Devant le jury, composé de :
Sara Bahadori, Psychiatre (Praticien Hospitalier)
Hôpital Robert Debré AP-HP, Institut de Psychiatrie et
Neuroscience de Paris
Sylvane Faure, Professeure, Université Côte d'Azur
Solène Kalenine, CNRS, HDR, Université de Lille
Olivier Koenig, Professeur, Université Lumière Lyon2
Sara Spotorno, Maître de Conférences, Keele University

**Traitements sémantiques et émotionnels lors de la
perception rapide de scènes visuelles :**

Etude du fonctionnement normal chez l'adulte et l'enfant
et application au trouble déficitaire de l'attention
avec ou sans hyperactivité

Jury

Rapporteur(e)s

Solène Kalenine, CNRS UMR 9193, HDR, Université de Lille

Olivier Koenig, Professeur, Université Lumière Lyon2

Examinatrice

Sara Bahadori, Pédiopsychiatre, hôpital Robert Debré, Institut de Psychiatrie et
de Neurosciences, Paris

RESUME

Lorsque nous regardons une scène visuelle, nous avons accès à un ensemble d'informations parmi lesquelles les contenus sémantiques et émotionnels ont un rôle majeur pour la compréhension et l'adaptation à l'environnement physique et social. Il a également été décrit que les associations entre des contenus de niveau « local » (les objets de premier plan) et de niveau « global » (le contexte, d'arrière-plan) permettent à l'observateur de catégoriser rapidement la scène. Or, nous avons mis en évidence, par la réalisation de trois revues critiques de littérature, une quasi absence de prise en compte de l'interaction des dimensions sémantique et émotionnelle lors de la catégorisation visuelle de l'objet et du contexte d'une scène visuelle. Nous avons analysé les traitements sémantiques et émotionnels, ainsi que leurs éventuelles interactions, grâce au paradigme de présentation visuelle rapide, dans une série de quatre études. Dans chacune, deux images (un objet et un contexte) apparaissaient rapidement hors de la vision fovéale. L'étude 1 a analysé l'effet de cohérence sémantique qui décrit des performances de catégorisation de l'objet et du contexte d'une scène visuelle supérieures lorsque les images sont présentées en situation de cohérence sémantique (i.e., selon une relation fortement cooccurrence comme un cochon et une ferme) plutôt qu'en situation d'incohérence (e.g., un cochon et une banquise). Cette étude a mis en évidence que les associations sémantiques entre objet et contexte se réalisent au stade de mise en correspondance entre un percept et une représentation en mémoire à long terme. L'étude 2 a analysé les relations entre traitements sémantiques et émotionnels en proposant une tâche de catégorisation du contenu sémantique et une tâche de catégorisation du contenu émotionnel d'une image neutre ou émotionnelle (positive ou négative). Cette image, de niveau local (objet) ou global (contexte), était présentée en condition de cohérence ou d'incohérence sémantique et émotionnelle (i.e., porteuse du même contenu émotionnel ou non) avec une image de l'autre niveau (objet ou contexte). Cette étude suggère l'existence de traitements sémantiques et émotionnels quelle que soit la tâche de l'observateur, et la présence d'interactions entre ces traitements, et entre l'objet et le contexte de la scène visuelle, variant en fonction des contraintes. Dans l'étude 3 nous avons analysé le développement de ces traitements auprès d'enfants scolarisés en primaire et répliqué les résultats obtenus chez l'adulte, à l'exception de l'effet de cohérence sémantique, qui semble ainsi dépendre de la quantité des expériences de vie. Enfin, l'étude 4 a analysé ces phénomènes auprès d'enfants qui présentent un Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H), condition à laquelle sont fréquemment associés des difficultés de reconnaissance et de gestion des émotions. Cette dernière étude a montré un traitement des émotions préservé lors

d'une présentation visuelle rapide, malgré de plus faibles performances de catégorisation sémantique du groupe TDA/H comparativement au groupe contrôle. Ces résultats suggèrent que le défaut de traitement des émotions relevé dans le TDA/H ne dépendrait pas de la prise initiale des informations, mais serait causé par un dysfonctionnement plus tardif dans la chaîne de traitements qui permet la prise de décision, probablement lié à un déficit attentionnel. D'autres recherches devront explorer si ce déficit provient d'un défaut des capacités d'attention soutenue ou de désengagement des informations émotionnelles. Au total, l'ensemble de ces résultats nous permet de décrire un modèle dynamique de la perception visuelle dans lequel les traitements sémantiques et émotionnels suivent des trajectoires développementales asynchrones et dont la priorisation d'une dimension de traitement sur une autre n'est pas préétablie, mais varie en fonction des contraintes liées à l'environnement et à l'observateur.

Mots clés : *Scènes visuelles, traitements sémantiques, traitements émotionnels, associations objet-contexte, catégorisation.*

ABSTRACT

When we observe a visual scene, we have access to a set of information, among which semantic and emotional contents have a key role in understanding and adapting to the physical and social environment. It has also been described that associations between the "local" (foreground objects) and "global" (context, background) content allow the observer to quickly categorize the scene. However, we have highlighted, in three critical reviews of the literature, that the interaction between semantic and emotional dimensions, during the categorization of the object and the context of a visual scene, have been largely neglected. We analyzed semantic and emotional processing, as well as their possible interactions, using the rapid visual presentation paradigm in a series of four studies. In each experiment, two images (an object and a context) quickly appeared outside foveal vision. Study 1 analyzed the semantic consistency effect, according to which categorization of the object and the context of a visual scene is superior when the images are presented in a semantically consistent situation (i.e., in a relationship that likely occurs in real life, such as a pig and a farm) rather than in an inconsistent situation (e.g., a pig and an iceberg). This study showed that the semantic associations between the object and the context act at the stage of matching between a percept and a long-term memory representation. Study 2 analyzed the relationships between semantic and emotional processing by proposing a semantic categorization task and an emotional categorization task of a neutral or an emotional (positive or negative) image. This image, either local (object) or global (context), was presented in a situation of semantic consistency or inconsistency and emotional consistency or inconsistency (i.e., carrying the same emotional content or not) with an image from the other level (object or context). This study suggested the existence of semantic and emotional processing regardless of the observer's task, and the presence of interactions between these processes that depend on the current constraints. In study 3, we analyzed the development of these processes in primary school-age children and replicated the results obtained in adults, with the exception of the effect of object-context semantic consistency, which thus seems to depend on the amount of life experience. Finally, study 4 analyzed semantic and emotional processing in children with Attention Deficit Disorder with or without Hyperactivity (ADHD), a condition to which deficits in recognition and management of emotions are frequently associated. This last study showed that emotional processing is preserved during rapid visual presentation, despite the lower semantic categorization performance in children with ADHD compared to the controls. These results suggested that the difficulties to process the emotions identified in ADHD do not depend on the initial gathering of information, but arise later,

probably due to the attention deficit characterizing this disorder. Further research will need to explore whether this deficit depends on a lack of sustained attention or on difficulties in disengagement from emotional information. Overall, all these results allow us to describe a dynamic model of visual perception, in which semantic and emotional processes follow asynchronous developmental trajectories and the prioritization of one dimension of processing over the other is not pre-established, but varies according to the constraints related to the environment and the observer.

Keywords: *Visual scenes, semantic processing, emotional processing, object-context associations, categorization.*

REMERCIEMENTS

Je tiens en premier lieu à remercier les membres du jury, Mme Kalénine, M. Koenig et Mme Bahadori, de m'avoir fait l'honneur d'accepter d'évaluer ce travail de thèse.

Je tiens ensuite à remercier vivement mes directrices, Sylvane Faure et Sara Spotorno. Merci pour votre soutien quelles que soient les circonstances (du problème informatique à la dernière minute d'envoi d'une soumission). Nous en aurons vécu des aventures depuis mon master 1 ; de nombreuses difficultés à surmonter mais également beaucoup de réussites, l'apprentissage du quotidien dans la recherche il me semble ! Merci de m'avoir accompagnée durant tout ce temps. Merci Mme Faure pour tous nos échanges, pour vos remarques constructives et votre confiance, qui m'ont permis de développer des compétences en recherche mais également d'enseignement. Merci Sara pour tout le temps passé sur Skype, et pour tout ce que vous m'avez appris, de l'utilisation d'outils de programmation à la rigueur scientifique. Merci également pour tous les bons moments passés ensemble notamment lors de colloques, j'espère que nous en aurons encore beaucoup d'autres !

Je souhaite ensuite remercier l'équipe du CERTA pour m'avoir accueillie pour les passations auprès des enfants avec TDA/H et plus particulièrement Catherine Fossoud qui a soutenu ce projet de recherche et a permis qu'il se réalise. Une des grandes réussites de ce doctorat est de m'avoir fait intégrer le CERTA en tant que psychomotricienne ! C'est un réel plaisir de refaire de la clinique et d'échanger avec l'ensemble de l'équipe, merci donc encore à Catherine Fossoud pour avoir soutenu mon embauche, pour ta bienveillance et pour toute l'énergie et le temps que tu donnes pour améliorer le quotidien des enfants avec trouble des apprentissages. Merci à Oriane Costini pour ton soutien et nos échanges, j'espère que maintenant nous pourrions monter pleins de projets ensemble ! Merci à Pascal Nellec pour ta disponibilité et ta gentillesse ! Merci bien évidemment au reste de l'équipe que je vois malheureusement peu mais les quelques moments avec vous sont très appréciables !

J'ai ensuite eu la chance d'intégrer le LINE pour mon poste d'ATER à l'INSPE, deuxième grande réussite de ce doctorat ! Merci à toute l'équipe et surtout merci à Margarida Roméro, ton enthousiasme dans la recherche (et dans les cubes en particulier !), ton dynamisme et surtout ton écoute et ta bienveillance m'ont vraiment aidée à tenir dans les moments les plus difficiles, dans les moments de doute. Merci de m'avoir écouté, merci de ta patience et de m'avoir permis d'organiser mon emploi du temps pour finaliser en premier lieu ma thèse, tu as permis en partie sa réalisation ! J'espère maintenant que nous allons enfin avoir plus de temps pour développer toutes nos idées de recherche, d'applications...

Bien entendu un très grand merci à l'ensemble des doctorants, actuels et anciens, du LAPCOS en particulier et de la MSHS en général, comment finir une thèse sans nos échanges et partages quotidiens ?

Merci en particulier à Hanane, ma sœur de thèse de psychologie cognitive ! On en aura partagé des galères et heureusement pleins de bons moments (de notre virée CogSci dans un hôtel Hilton à Londres au week-end Barcelonnais ! Je peux d'ailleurs ajouter notre petit Alex dans cette parenthèse qui, sans être frère de psycho cognitive, nous aura suivi à la découverte de l'ours paddigton). Romain merci pour tous les moments où tu as subi mon papotage, pour toutes nos discussions et j'espère que nous aurons l'occasion de faire à nouveau plus de pots maintenant ! Merci à tous ceux trop nombreux à citer, Victor et ses conseils culinaires, Camille qui est partie trop vite trop loin là bas à l'autre bout de la France, Candice pour ta gentillesse (il y a un vide au LAPCOS sans toi), Adrien, Paul, Lou, Selim, Vincent, Mustapha, Ismaël, Grand Alex et tous les autres ! Merci au LAPCOS et particulièrement à Dirk, Isabelle, Pierre, Xavier, Damien, Galina et bien entendu Diane pour m'avoir accueillie, accompagnée, soutenue et pour tous les bons moments passés !

Un grand merci également aux enfants et adultes qui ont participé à cette recherche, et aux écoles qui m'ont accueilli : La pinède et Jules Ferry à Cagnes sur Mer et les Magnolias à Nice.

Merci également à toutes les personnes qui ont participé d'une manière ou d'une autre à ce parcours de thèse, en commençant par Jean-Michel Albaret qui a encouragé dès le début mon attrait pour la recherche, merci aux « psychopotes » et notamment Mathilde et Laurène, merci aussi aux personnes rencontrées durant mon master 2 recherche en neuropsychologie et neurosciences cliniques, en particulier Morgane, je me souviens de notre hésitation à commencer ou non une thèse devant une galette à Bron ! Merci Laetitia, Lolo, Mumu et tous ceux qui ont subi cette période où j'étais peu disponible, j'espère me rattraper maintenant !

Et bien sûr merci à ma famille de leur soutien constant, de toujours prendre soin de moi, même à l'autre bout de la France, et merci Moumoune d'avoir lu tout ce pavé pour vérifier mes fautes d'orthographe ! Merci également à Anne-Marie et Laurent de votre écoute et de tous les moments passés ensemble qui m'ont permis de m'aérer.

Et enfin merci à Minouche de m'avoir forcé à me lever tous les matins et pour la ronron thérapie ! Et surtout merci à mon cher participant 6, Flo, qui a été obligé de subir l'ensemble d'une thèse, depuis l'obtention de la bourse jusqu'au rendu de ce manuscrit en passant par tous les moments de stress, de perte de motivation et tous les moments euphoriques de réussite ! Merci de m'avoir soutenu, d'avoir endurer tout ce quotidien de vie de thésard, merci de m'avoir forcée à m'aérer, à arrêter de travailler et à m'avoir emmenée au milieu des montagnes et des étoiles pour me permettre de relativiser et pour empêcher que mon cerveau n'explose ! Je ne te remercierai jamais assez pour tout ce que tu m'apportes, merci d'être toi !

TABLE DES MATIÈRES

TABLE DES MATIÈRES	8
<i>Index des figures et des tableaux</i>	11
Introduction générale	12
Le processus de catégorisation	15
Interaction entre les informations locales et globales au sein des scènes visuelles	16
La priorisation des traitements émotionnels	17
Présentation générale du travail de thèse	19
Structure de la thèse	21
Données théoriques	25
<i>Chapitre I. Traitements sémantiques et émotionnels des scènes visuelles complexes chez l'adulte</i>	26
I.1 Article I. Traitements sémantiques et émotionnels des scènes visuelles complexes : une synthèse critique de l'état actuel des connaissances	26
1. Traitement des informations sémantiques	30
2. Traitement des informations émotionnelles	39
3. Traitements sémantiques et émotionnels : indépendance ou interaction ?	48
Conclusion	51
Références bibliographiques	55
<i>Chapitre II. Développement des capacités de catégorisation des informations perceptives, sémantiques et émotionnelles de la naissance à l'adolescence</i>	66
II.1 Résumé français de l'article	66
II.2 Article II. Object categorization from infancy to adolescence: Interplay between perceptual, semantic and emotional information	66
Introduction	69
The development of object categorization: How perceptual and semantic information interact	70
Contextual effects on object categorization: How perception and cognition interact	75
Emotional categorization: Is emotional information special?	79
Conclusion and future directions	83
References	87
<i>Chapitre III. Traitement du contenu émotionnel des scènes visuelles par des enfants présentant un trouble déficitaire de l'attention avec ou sans hyperactivité</i>	98
III.1 Résumé français de l'article	98
III.2 Article III. Emotional Scene Processing in Children and Adolescents with Attention Deficit / Hyperactivity Disorder: A Systematic Review	99
Introduction	102
Methods	104
Results	105
Discussion	119
Conclusions	121
References	123

Partie expérimentale	129
<i>Chapitre IV. Synthèse théorique, problématique et méthodologie générale</i>	130
IV.1 Synthèse théorique	130
IV.2 Problématique	133
IV.3 Méthodologie générale	136
IV.3.1 Considérations générales pour la partie expérimentale	136
IV.3.2 Présentation des images en champ visuel divisé	137
IV.3.3 Tâches de catégorisation sémantique et de catégorisation émotionnelle	137
<i>Chapitre V : Etude du rôle des liens sémantiques entre l'objet et le contexte lors de la perception rapide de scènes visuelles complexes</i>	139
V.1 Rappels des apports théoriques principaux	139
V.2 Présentation de l'étude	140
V.3 Résumé de l'étude	142
V.4 Article IV. When two images make a scene: Reciprocal semantic predictions drive categorization of contexts and objects even when they are separate	143
Introduction	146
Results	149
Experiment 1	152
Experiment 2	153
Experiment 3	155
Experiment 4	156
Experiment 5	157
General Discussion	157
Conclusion	160
General Method	160
References	165
V.5 Discussion intermédiaire concernant les associations sémantiques objet/contexte	170
<i>Chapitre VI : Etude de l'interaction entre les traitements sémantiques et émotionnels lors de la perception rapide de scènes visuelles complexes chez l'adulte</i>	172
VI.1 Rappel des apports théoriques principaux	172
VI.2 Présentation de l'étude	173
VI.3 Résumé de l'étude	175
VI.4 Article V. Interplay between semantic and emotional information in visual scene processing.	176
Introduction	179
Experiment 1: Semantic categorization	191
Experiment 2: Emotional categorization	194
General Discussion	197
References	202
VI.5 Discussion intermédiaire	207
<i>Chapitre VII : Etude des traitements sémantiques et émotionnels lors de la perception rapide de scènes visuelles complexes chez l'enfant d'âge scolaire</i>	209
VII.1 Rappels des apports théoriques principaux	209
VII.2 Présentation de l'étude	211
VII.3 Résumé de l'étude	214

VII.4 Article VI. Semantic and emotional visual scene processing in school-aged children.	214
Introduction	218
General Method	223
Experiment 1: Semantic categorization	228
Experiment 2 : Emotional categorization	233
General discussion	236
Conclusion	241
Références	243
VII.5 Discussion intermédiaire	248
<i>Chapitre VIII: Application au Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H)</i>	249
VIII.1 Rappel des apports théoriques principaux	249
VIII.2 Présentation de l'étude	250
VIII.3 Analyses préliminaires	255
VIII.3.1 Tâche de catégorisation sémantique	255
VIII.3.2 Tâche de catégorisation émotionnelle	258
VIII.4 Discussion intermédiaire	261
Discussion générale	264
<i>1. Rappel de la démarche et synthèse des résultats</i>	<i>265</i>
<i>2. Implications théoriques</i>	<i>270</i>
2.1 Traitements sémantiques des scènes visuelles	270
2.2 Traitements émotionnels des scènes visuelles	273
2.3 Associations sémantiques et émotionnelles entre l'objet et le contexte au sein des scènes visuelles	278
2.4 Synthèse	279
<i>3. Application en clinique</i>	<i>284</i>
<i>Conclusion</i>	<i>285</i>
Références	287
Annexes	298

Index des figures et des tableaux

Figure 1: représentation schématique simplifiée des principaux modèles théoriques des associations objet/contexte.....	140
Figure 2 Représentation schématique d'un modèle prédictif de facilitation cognitive de la perception visuelle.	271
Figure 3 : Représentation schématique des traitements sémantiques et émotionnels lors de la présentation visuelle rapide de deux	282
Tableau 1 : Synthèse des principaux résultats des quatre études empiriques.....	267

INTRODUCTION GÉNÉRALE

“Perception is where cognition and reality meet” (Neisser, 1976)

Lorsque vous feuillotez les pages d'un magazine, que vous regardez la bande annonce d'un film ou bien encore lorsque les portes de l'ascenseur s'ouvrent sur une pièce inconnue, vous avez accès en une fraction de seconde à une multitude d'éléments auxquelles vous donnez sens. Il s'agit d'éléments locaux, décrits classiquement dans la recherche sur la perception visuelle sous le terme d'« objet », et d'éléments plus structurels, globaux, décrits sous le terme de « contexte ». L'ensemble de ces éléments locaux, qui peuvent être des objets animés (e.g., des animaux, des personnes, ...) ou inanimés (e.g., un parasol, une chaise, ...), et globaux (e.g., une plage, un bureau, ...) constitue votre environnement physique et social. La perception, notamment visuelle, de ces éléments vous permet d'obtenir certaines informations sur leur contenu sémantique, par exemple leur nature, ainsi que sur leur contenu émotionnel, c'est-à-dire sur les émotions qu'ils véhiculent et sur le ressenti émotionnel qu'ils vous procurent. Vous pouvez par exemple, pour des informations locales, décrire si l'animal que vous percevez est un chat ou un chien, s'il est agressif, et quelles émotions vous ressentez en le voyant (e.g., de la peur ou de l'attendrissement). Et de façon équivalente pour des informations contextuelles, globales, vous pouvez déterminer si vous êtes face à une plage paradisiaque un jour de beau temps ou durant une tempête. La totalité des informations sémantiques et émotionnelles, à la fois des constituants du niveau local et du niveau global, peut être initialement traitée en quelques dizaines de millisecondes, de manière plus ou moins automatique selon l'orientation attentionnelle de l'observateur. Ces traitements permettent de s'adapter à et d'agir efficacement sur le monde environnant. Par exemple votre réaction sera différente si vous rentrez chez vous, que votre porte est entrouverte et que vous découvrez une pièce saccagée suggérant que vous avez été cambriolé ou que vos amis vous attendent pour une surprise d'anniversaire. Pour étudier ces environnements, tels que le salon cambriolé ou une plage paradisiaque, les recherches sur la perception visuelle utilisent des scènes visuelles qui représentent des environnements réels, naturels. Les scènes visuelles sont constituées par un contexte de niveau global et un ou plusieurs objets de niveau local, elles sont présentées majoritairement via des photographies de scènes du quotidien, par des dessins ou plus rarement par des stimuli dynamiques tels que des vidéos ou par l'utilisation de la réalité virtuelle.

La grande rapidité de traitement des dimensions sémantiques et émotionnelles de ces scènes visuelles est encore plus surprenante lorsque nous considérons que nous sommes constamment soumis à un flux d'informations beaucoup trop dense pour être entièrement

appréhendé. Les stimulations reçues doivent ainsi être priorisées et regroupées afin d'éviter une surcharge cognitive et permettre à l'observateur de s'adapter et de réagir dès la survenue d'une modification de l'environnement, si cela est pertinent. L'observateur peut décider des informations qu'il va sélectionner, et orienter volontairement son attention. L'attention endogène est alors engagée et guidée par les buts et attentes de l'individu, le traitement est volontaire et descendant (processus *top-down*). Cependant la présence d'un élément saillant qui apparaît dans l'environnement visuel, bien que non pertinent, peut interférer avec l'activité en cours voire l'interrompre. Il s'agit alors de traitements non volontaires engagés lorsque les caractéristiques du stimulus attirent l'attention exogène (processus *bottom-up*). Le caractère saillant du stimulus peut dépendre notamment de caractéristiques perceptives (e.g., couleur, luminosité, contraste...), ou de caractéristiques sémantiques ou émotionnelles détectées car importantes pour l'individu (e.g., détection de la présence d'un serpent). Les traitements volontaires (dirigés par les buts) et non volontaires (dirigés par le stimulus) interagissent continuellement lors de la perception visuelle (Corbetta, Patel, & Shulman, 2008; Corbetta & Shulman, 2002). Bien que des enjeux scientifiques majeurs en découlent, nous n'aborderons pas la question de l'accès à la conscience de ces deux niveaux de traitements. Nous considérerons ici comme non volontaires, et probablement automatiques, les traitements des caractéristiques du stimulus hors du focus attentionnel endogène, par exemple si l'attention endogène de l'observateur est orientée vers le contenu sémantique d'une image et qu'un traitement non volontaire du contenu émotionnel a lieu, sans questionner si ce traitement est non-conscient ou conscient.

Parmi les traitements fondamentaux qui contribuent à la rapidité et à l'efficacité de la cognition visuelle, nous avons décidé de retenir, dans le cadre de ces travaux de thèse, certaines opérations cognitives particulières que nous regroupons sous trois grands axes : le processus de catégorisation, la mise en commun des informations locales et globales au sein des scènes visuelles et la priorisation du contenu émotionnel. Notre décision de nous intéresser particulièrement à ces domaines provient de constats lors de notre expérience clinique en tant que psychomotricienne. De nombreux enfants avec un trouble déficitaire de l'attention avec ou sans hyperactivité ne parviennent pas à synthétiser, à regrouper des informations et la mise en commun d'éléments de détails et d'éléments contextuels est difficile, ce qui impacte leur compréhension générale d'une histoire ou d'une image. De plus, ces enfants sont majoritairement en difficulté pour comprendre leurs émotions ainsi que les émotions des autres ce qui entraîne des conflits familiaux et avec leurs pairs. De ces constats nous nous sommes demandé si ces difficultés ne pourraient pas provenir d'une façon particulière de traiter et donc

de comprendre l'environnement, comme cela a pu notamment être décrit dans le cadre du trouble du spectre autistique. L'étude des traitements précoces visuels et de l'intégration des informations présentes dans le champ visuel en un tout cohérent nous a ainsi semblé nécessaire, à la fois dans le fonctionnement cognitif normal et à la fois pour améliorer la compréhension de fonctionnements cognitifs particuliers et ainsi adapter au mieux les protocoles thérapeutiques proposés. Nous introduisons succinctement ici ces trois axes qui seront repris plus en détail dans la partie théorique de ce manuscrit.

Le processus de catégorisation

Le processus de catégorisation est défini comme la capacité à regrouper sous une même entité différents spécimens qui partagent certains attributs. Ce processus évite une surcharge cognitive, facilite la communication ainsi que la mémorisation et permet d'inférer certaines caractéristiques d'un spécimen de la catégorie à un autre (Sloutsky & Fisher, 2011). Ainsi on peut regrouper un ensemble d'arbres sous le terme « forêt », des poissons, des oiseaux et des mammifères dans la catégorie « animal ». Pour certains auteurs, la catégorisation est ainsi considérée comme le socle sur lequel repose la cognition (Harnad, 2017).

La catégorisation peut se faire selon des caractéristiques sémantiques plus ou moins précises (le chat fait partie de la catégorie « vivant » ou « animal » ou « félin ») ou selon des caractéristiques perceptives (ce chat et cette boîte sont noirs), ou enfin selon une catégorie émotionnelle (ce chat exprime de la peur). La majorité des travaux sur la capacité de catégorisation visuelle rapide sémantique a utilisé des tâches de détection de la présence ou non dans une image d'un stimulus appartenant à une catégorie sémantique cible. Il a ainsi été décrit que l'observateur est capable de déterminer en 150 millisecondes (ms) si un animal (Thorpe, Fize, & Marlot, 1996) ou un moyen de transport (VanRullen & Thorpe, 2001) est présent dans une scène visuelle présentée pour seulement environ 20 ms. Bien que moins étudiées, des capacités similaires de catégorisation des contextes (i.e., niveau global) ont également été montrées (Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Rousselet, Joubert, & Fabre-Thorpe, 2005). Au niveau émotionnel, la catégorisation est principalement étudiée selon les théories dimensionnelles (e.g., Russell, 2003) qui avancent notamment une continuité sur un axe plaisant/positif versus déplaisant/négatif, ou selon la théorie des émotions de base (Ekman, 1992) qui différencie un nombre limité d'émotions distinctes telles que la joie, la colère, la tristesse, la peur, le dégoût et la surprise. L'analyse du traitement des informations émotionnelles s'est révélée néanmoins plus complexe que celle du traitement des informations

sémantiques de par la subjectivité des émotions. En effet, même si de nombreux travaux sont en faveur d'une identification universelle des expressions faciales dites de base (ou primaires, Ekman, 1992), un nombre grandissant d'études montre que certains facteurs peuvent moduler cette reconnaissance. Les facteurs culturels et sociaux (pour une revue voir: Gendron, Crivelli, & Barrett, 2018) ainsi que l'état émotionnel et les expériences antérieures de l'observateur peuvent contribuer à la subjectivité de traitement des informations émotionnelles. Par exemple, une araignée peut être catégorisée émotionnellement comme négative par une personne et positive par une autre alors qu'elle sera catégorisée sémantiquement sous le label « araignée » par ces deux individus. Un autre aspect important à souligner est que la majorité des études qui analysent la capacité de catégorisation des émotions utilisent comme matériel expérimental uniquement des visages expressifs, qui plus est, souvent présentés isolément et de face, donc d'une façon peu représentative du quotidien. En outre, le traitement spécifique des visages empêche une généralisation des résultats au traitement des informations émotionnelles véhiculées par exemple par un animal, une posture corporelle, un paysage... Malgré ces difficultés inhérentes au construit d'émotion et à l'approche expérimentale utilisée pour étudier son traitement, de nombreux travaux ont permis de mettre en évidence des capacités de catégorisation émotionnelle présentes dès la première année de vie dans le développement humain (Bayet, Pascalis, & Gentaz, 2014).

Interaction entre les informations locales et globales au sein des scènes visuelles

L'être humain peut donc catégoriser très rapidement aux niveaux sémantique et émotionnel les objets ou le contexte d'une scène visuelle. Certains travaux se sont intéressés à l'impact du lien entre ces niveaux - niveau local de l'objet et niveau global du contexte - sur les processus de catégorisation.

Palmer (1975) a montré que la présence d'un contexte cohérent sémantiquement avec l'objet (i.e., lorsque ces deux images se retrouvent fréquemment ensemble dans la vie quotidienne, comme un grille-pain et une cuisine) facilitait la catégorisation sémantique de celui-ci en diminuant le temps de réponse et en améliorant l'exactitude par rapport à une condition d'incohérence sémantique (i.e., lorsque les relations objet-contexte sont peu cooccurrentes comme un grille-pain et une forêt). Depuis, de nombreuses études ont répliqué ces résultats (bien que cela ne soit pas toujours le cas, voir Henderson & Hollingworth, 1999; Hollingworth & Henderson, 1998) et aussi précisé différents aspects interactionnels : le traitement du contexte influence la catégorisation sémantique de l'objet (Biederman,

Mezzanotte, & Rabinowitz, 1982; Joubert, Fize, Rousselet, & Fabre-Thorpe, 2008; Munneke, Brentari, & Peelen, 2013; Rémy, Vayssière, Pins, Boucart, & Fabre-Thorpe, 2014) et celui de l'objet influence la catégorisation sémantique du contexte (Davenport & Potter, 2004; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007). De manière analogue, il a également été montré une performance de catégorisation émotionnelle supérieure en situation de cohérence émotionnelle, c'est-à-dire lorsque l'objet et le contexte véhiculent la même émotion, par rapport à une situation d'incohérence émotionnelle où les deux niveaux sont porteurs d'émotions différentes (Kret & de Gelder, 2010; Nelson & Mondloch, 2017; Righart & de Gelder, 2008; Theurel et al., 2016). Cependant, ces effets facilitateurs, dits respectivement effet de cohérence sémantique et effet de cohérence émotionnelle, sont encore peu compris. Notamment, en ce qui concerne les traitements ascendants, la question du stade, dans un continuum d'un stade perceptif à un stade décisionnel, où se réalisent ces associations sémantiques et émotionnelles entre l'objet et le contexte est encore débattue. En outre, la perception visuelle ne se résume pas à des traitements uniquement ascendants. Pour approcher au mieux les processus de catégorisation, il conviendrait également de prendre en considération l'influence de traitements descendants sur le déroulement des traitements ascendants.

La priorisation des traitements émotionnels

La recherche sur les traitements émotionnels s'est largement développée depuis une trentaine d'années, notamment portée par la question des relations entre la cognition et les émotions et d'une éventuelle interdépendance entre ces dimensions (e.g., Lazarus, 1984; Zajonc, 1984). Actuellement, de nombreux chercheurs considèrent que le traitement des émotions fait partie des fonctions cognitives (e.g., Sander & Koenig, 2002) et décrivent les influences réciproques entre les émotions et les autres fonctions cognitives (e.g., Pessoa, 2008; 2009). Les théories de l'évaluation cognitive des émotions stipulent ainsi que la valeur émotionnelle accordée à un stimulus dépend de l'interprétation cognitive que l'on fera de ce stimulus. Cette interprétation dépend de la pertinence du stimulus (*appraisal*) pour l'individu, pertinence qui est constamment réévaluée en fonction des situations et besoins de l'individu (Brosch, Scherer, Grandjean, & Sander, 2013; Sander, Grandjean, & Scherer, 2005). Le développement de la neuroimagerie a permis de préciser les soubassements neuroanatomiques et fonctionnels de l'influence de la cognition sur les émotions mais également des émotions sur la perception, l'attention et la mémoire, effets déjà largement démontrés au niveau comportemental. Ainsi d'importantes interactions entre les centres cérébraux décrits

initialement comme dédiés aux traitements des émotions et ceux dédiés aux traitements cognitifs ont été mises en lumière (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012; Pessoa, 2008, 2010). L'amygdale, structure sous-corticale clé dans le traitement des informations émotionnelles, présente des connexions réciproques avec des régions liées à la mémoire (i.e., hippocampe), et de nombreuses régions corticales perceptives (i.e., cortex visuel primaire, cortex temporal inférieur) et exécutives (Pessoa, 2009). Selon certains modèles, cette connectivité permet à l'amygdale de renforcer le traitement perceptif initial et la représentation des stimuli émotionnels. De nombreuses études ont, en effet, démontré une priorisation du traitement d'informations comprenant un contenu émotionnel par rapport à des informations dites « neutres » car dépourvues de contenu émotionnel (Carretié, 2014; Pessoa, 2009). De plus, de par ses connexions multiples avec les aires corticales préfrontales, l'amygdale a une action sur les régions liées à la prise de décision (i.e., cortex latéral préfrontal et cortex orbito-frontal) et à la motivation (i.e., cortex cingulaire antérieur). Réciproquement, les réseaux préfrontaux peuvent réguler les émotions notamment en inhibant l'activité de l'amygdale.

De nombreux travaux sont ainsi en faveur d'une interaction entre les processus cognitifs et émotionnels : les traitements cognitifs et les traitements émotionnels s'influencent réciproquement. Cependant les liens plus spécifiques entre traitements émotionnels et un traitement cognitif particulier (ici traitement sémantique) sont encore peu étudiés. Les questions d'une indépendance et d'une séquentialité entre les traitements sémantiques et émotionnels divisent encore actuellement les chercheurs. Pour certains il existe une dépendance de traitement entre ces deux dimensions, et la « reconnaissance » sémantique d'un élément est nécessaire pour pouvoir le catégoriser émotionnellement (Lähteenmäki, Hyönä, Koivisto, & Nummenmaa, 2015; Nummenmaa, Hyönä, & Calvo, 2010; Storbeck, Robinson, & McCourt, 2006). Pour d'autres les traitements sémantiques et émotionnels se réalisent de manière indépendante, le jugement émotionnel pouvant même être accompli avant que l'individu ait reconnu l'objet.

Nous tenons à préciser que ce travail de thèse n'a pas pour ambition de documenter la question de la nature des traitements émotionnels ; notre objectif est d'étudier si, et dans quelles conditions, des interactions existent entre les traitements émotionnels et les traitements sémantiques. Pour ce faire, nous avons décidé d'inscrire nos travaux dans les approches théoriques qui intègrent le traitement des émotions dans les fonctions cognitives.

Présentation générale du travail de thèse

Pour résumer l'ensemble des points abordés ci-dessus, nous retenons que l'être humain a une remarquable capacité à se représenter rapidement son environnement visuel. Cette capacité est en partie due aux processus de catégorisation sémantique et émotionnelle, à la mise en commun d'informations provenant du niveau local et du niveau global des scènes visuelles et à la priorisation de traitement de certaines informations. Néanmoins, malgré l'ensemble des travaux réalisés depuis un demi-siècle sur la perception visuelle rapide de scènes visuelles et qui ont permis une avancée considérable dans la compréhension des processus en jeu, de nombreuses questions restent ouvertes. Notamment, l'état des lieux des travaux de recherche a mis en lumière l'existence de nombreuses études d'une part sur les traitements sémantiques et d'autre part sur les traitements émotionnels mais peu de prise en compte de l'interaction entre ces dimensions. L'étude de chaque dimension de façon séparée est certes une étape indispensable à la compréhension de leurs fonctionnements et de leurs spécificités, mais la compréhension scientifique de la perception de notre environnement ne peut se résumer à l'analyse d'une seule dimension. Une meilleure description et compréhension des influences réciproques, entre les traitements des niveaux local et global des scènes visuelles en fonction des traitements sémantique et émotionnel, et de leurs éventuelles interactions, est nécessaire à l'élaboration d'une explication scientifique de la cognition visuelle.

Nos travaux ont pour but de contribuer à la compréhension de la façon dont l'être humain perçoit son environnement visuel. Nous souhaitons ainsi dans cette thèse analyser si, et comment, les dimensions sémantique et émotionnelle, interagissent durant la perception visuelle rapide de scènes visuelles. Nous étudions les capacités de catégorisation sémantique et émotionnelle d'une image présentée rapidement, ce qui ne permet pas d'exploration visuelle donc de prise d'information de détail, afin d'analyser les traitements précoces lors de l'apparition d'un nouvel élément dans le champ visuel de l'observateur. Nous questionnons ainsi quelles dimensions (sémantique et/ou émotionnelle) d'une image vont être traitées, lorsque la prise d'information est limitée, en fonction de l'orientation attentionnelle de l'observateur et des contenus sémantiques et émotionnels d'une autre image présente dans le champ visuel. Nous souhaitons ainsi étudier si l'individu crée une représentation mentale de la scène visuelle même lorsque les deux composants de l'image sont séparés et si ces associations sémantiques et/ou émotionnelles entre objet et contexte influencent les capacités de catégorisation.

Pour approcher au mieux la vie quotidienne, nous utilisons des stimuli constitués d'images d'objets et de contextes habituels. Nous souhaitons ainsi pouvoir tester la

généralisation de résultats décrits dans la littérature avec des objets isolés, et, dans le cadre des études sur les traitements émotionnels, avec des visages isolés, c'est-à-dire sans information contextuelle, à l'utilisation de stimuli plus complexes et donc plus proches de ceux rencontrés couramment. Nous utilisons un paradigme de présentation visuelle rapide en champ visuel divisé dans toutes nos tâches expérimentales pour contrôler que chaque image soit présentée de manière identique, c'est-à-dire hors de la vision fovéale et avant qu'une exploration visuelle ne soit possible. La fovéa est la région centrale de la macula, qui permet la plus fine discrimination visuelle grâce à sa forte concentration en cônes alors que la région périphérique (constituée principalement de bâtonnets) présente une acuité médiocre et autorise une vision globale (Bringmann et al., 2018). Lorsqu'un nouvel élément apparaît dans notre champ visuel, hors de la vision fovéale, s'il nécessite une exploration plus fine une saccade oculaire sera déclenchée pour le placer dans la vision fovéale. Le paradigme de présentation visuelle rapide en champ visuel divisé permet de contrôler la position et le temps de présentation pour que le stimulus n'apparaisse pas et ne puisse pas être placé en vision fovéale (cf. chapitre IV.3.2 pour plus de précisions sur ce paradigme).

Les travaux réalisés dans le cadre de cette thèse se situent dans le fonctionnement cognitif normal chez l'adulte, puis chez l'enfant d'âge scolaire de huit à douze ans. Nous souhaitons caractériser les capacités de perception visuelle rapide de l'enfant à une période de développement très peu étudiée malgré sa place majeure dans le développement des apprentissages et des relations sociales. Ces travaux nous permettent également de renforcer la compréhension du fonctionnement adulte par la prise en compte du développement des capacités de catégorisation sémantique et émotionnelle. Enfin leurs résultats (i.e. effets des facteurs sémantiques et émotionnels et indications sur les influences réciproques entre ces dimensions dans la perception) auprès d'enfants au développement typique serviront de références pour l'étude des capacités de catégorisation sémantique et émotionnelle dans des populations qui présentent des troubles du traitement des émotions (identification ou contrôle). En effet, des particularités de traitement des informations émotionnelles sont décrites dans certains troubles psychoaffectifs, psychiatriques et neuro-développementaux. Une meilleure compréhension des traitements des informations sémantiques et émotionnelles dans le cadre du fonctionnement cognitif normal est indispensable pour permettre de mieux appréhender et spécifier les déficits particuliers présents dans certains troubles.

Parmi les troubles neuro-développementaux, le Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H) touche 3 à 7 % des enfants d'âge scolaire et persiste chez 60% d'entre eux à l'âge adulte (*American Psychiatric Association*, 2013). Il est caractérisé par une

triade symptomatologique - trouble attentionnel, impulsivité et hyperactivité - pouvant entraîner des difficultés dans le parcours académique et professionnel et des difficultés socio affectives. Ce trouble constitue un problème de santé publique, par sa forte prévalence et par sa répercussion sur l'ensemble de la sphère de la vie des patients TDA/H. Les personnes avec TDA/H présentent ainsi, comparativement à la population générale, plus de difficultés à trouver et à maintenir un emploi, plus de risques de conduites à risques et addictives, de risque d'exclusion sociale... Un rapport de la Haute Autorité de Santé (2014) a mis en avant l'importance du diagnostic et d'une prise en charge adaptée pour éviter l'aggravation des répercussions psychologiques, scolaires et sociales de ce trouble. Au-delà de ces symptômes primaires, le TDA/H est très fréquemment associé à des troubles du comportement, de la gestion des émotions et de la reconnaissance des émotions (pour des revues: Martel, 2009; Shaw, Stringaris, Nigg, & Leibenluft, 2014; van Stralen, 2016). Pour mieux appréhender ces déficits, qui se répercutent négativement sur le quotidien des enfants et des familles concernées, l'étude des particularités de traitement des informations émotionnelles constitue désormais une part importante des recherches sur le TDA/H. En analysant les capacités de catégorisation visuelle rapide à la fois des informations sémantiques et émotionnelles d'enfants présentant un TDA/H et en les comparant à celles d'enfants neurotypiques, notre recherche souhaite participer à l'amélioration de la compréhension des déficits présents dans ce trouble.

Structure de la thèse

La rédaction de ce travail de thèse s'articule autour d'articles scientifiques : dans la partie théorique nous présentons trois revues de la littérature puis dans la partie expérimentale trois articles empiriques, ainsi que les résultats préliminaires d'une étude réalisée auprès d'enfants qui présentent un TDA/H. Cette dernière étude est présentée comme un chapitre « classique » de thèse et non en format article.

Dans la partie théorique, le premier article (Leroy, Spotorno & Faure, accepté en février 2019) propose une revue critique des connaissances actuelles sur les interactions entre les traitements sémantiques et émotionnels lors de la perception rapide de scènes visuelles chez l'adulte. Le second article (Leroy, Spotorno & Faure, soumis) propose un état de l'art sur les interactions entre les informations perceptives, sémantiques et émotionnelles dans le processus de catégorisation des objets au cours du développement, du nouveau-né à l'adolescent. Au-delà des avancées majeures réalisées au cours des dernières années, notre analyse de la littérature a permis de mettre en relief des questions de recherche encore débattues ou négligées, dont

l'étude des mécanismes sous-jacents aux effets de cohérence sémantique et émotionnelle, l'utilisation encore dominante de stimuli isolés au détriment de stimuli plus représentatifs du quotidien tels que des scènes visuelles (paysages urbains, vues d'intérieurs...) et enfin le manque notable d'études sur la catégorisation visuelle chez les enfants d'âge scolaire. Le troisième et dernier article de cette partie théorique (Leroy, Spotorno & Faure, en révision) est une revue systématique de la littérature sur le traitement des informations émotionnelles portées par les scènes visuelles chez les enfants qui présentent un TDA/H. Cette synthèse met en avant la très faible présence d'études publiées ayant utilisé des scènes visuelles dans l'évaluation des capacités de catégorisation sémantique et émotionnelle dans cette population. Néanmoins, cette revue suggère que les particularités de traitement des informations faciales émotionnelles, décrites dans une littérature beaucoup plus fournie, sont également présentes lors du traitement de scènes visuelles complexes : déficit de catégorisation émotionnelle et/ou influence plus forte de la présence d'un contenu émotionnel non pertinent à la tâche.

La partie expérimentale est constituée de quatre études empiriques. Cette seconde partie commence par un chapitre décrivant la problématique générale de la thèse et la méthodologie générale utilisée pour les études empiriques.

Le quatrième article de cette thèse (Leroy, Faure & Spotorno, en révision) analyse les relations sémantiques objet-contexte qui déterminent l'effet de cohérence sémantique. En utilisant, pour la première fois à notre connaissance, une présentation simultanée et séparée de l'image d'un contexte et de l'image d'un objet qui sont en situation de cohérence sémantique ou non, cette étude montre que l'effet de cohérence sémantique est présent même lorsque les constituants de la scène ne sont pas présentés intégrés dans une seule image. Ceci suggère que l'effet de cohérence sémantique se réalise à un stade représentationnel du traitement de l'information et pas à un stade perceptif ni à un stade décisionnel. Ainsi, lorsque l'objet et le contexte sont cohérents sémantiquement, la quantité d'information perceptive nécessaire pour accéder à la représentation sémantique en mémoire à long terme de chacun de ces niveaux est réduite, ce qui permet des performances de catégorisation supérieures à celles d'une condition incohérente.

Le cinquième article (Leroy, Faure & Spotorno, soumis au jour de la soutenance) propose d'étudier les interactions entre les traitements sémantiques et émotionnels dans la perception visuelle rapide en fonction de l'orientation attentionnelle de l'observateur. Pour cela, une tâche de catégorisation sémantique ou de catégorisation émotionnelle d'une image cible (objet ou contexte) est proposée. Chaque image cible est présentée simultanément avec une

autre image. Les cohérences sémantique et émotionnelle entre les deux images (objet et contexte) sont manipulées de la même manière pour les deux tâches. Cette étude a permis de mettre en évidence plusieurs résultats importants. Premièrement nous répliquons l'effet de cohérence sémantique lors de la présentation séparée de l'objet et du contexte tel que décrit dans notre quatrième article. Deuxièmement, nous montrons que l'effet de cohérence émotionnelle, décrit dans la littérature pour des présentations en vision fovéale d'une scène visuelle composée d'un objet intégré dans un contexte, est également mis en évidence lors d'une présentation visuelle rapide et séparée de l'objet et du contexte et que cet effet n'interagit pas avec celui de cohérence sémantique. Troisièmement, notre étude suggère une interdépendance entre les traitements sémantiques et émotionnels de l'image cible. En effet, quelle que soit la tâche réalisée, les contenus sémantique et émotionnel de l'image cible apparaissent traités même s'ils sont non pertinents à la tâche.

Le sixième article (Leroy, Faure & Spotorno, en préparation) analyse les capacités de catégorisation sémantique et émotionnelle visuelles chez des enfants, au développement typique, âgés de 8 à 12 ans. Un objet et un contexte sont présentés rapidement et séparément à des enfants qui doivent réaliser soit une tâche de catégorisation sémantique, avec manipulation de la cohérence sémantique entre l'objet et le contexte, soit une tâche de catégorisation émotionnelle avec manipulation de la cohérence émotionnelle. Contrairement à l'étude chez l'adulte décrite précédemment (cinquième article) les cohérences sémantique et émotionnelle ne sont pas manipulées de façon orthogonale. Cette étude met en évidence qu'un traitement visuel rapide en parallèle d'un objet et d'un contexte pour accéder à des représentations de haut niveau est possible dès 8 ans. Les enfants présentent des capacités de catégorisation sémantique, en terme d'exactitude, qui sont sensiblement identiques à celles des adultes (mais avec des temps de réponse allongés). Toutefois, l'effet de cohérence sémantique entre les deux images présentées en même temps à chaque essai n'est pas répliqué, ce qui suggère que celui-ci se développe au fur et à mesure des expériences et des apprentissages, alors que l'effet de cohérence émotionnelle est mis en évidence chez les enfants dès 8 ans. De plus, la présence de contenu émotionnel interfère avec les processus de catégorisation sémantique (i.e., diminution de la performance de catégorisation des images à contenu négatif). L'ensemble de ces résultats suggère un traitement priorisé du contenu émotionnel et des trajectoires différentes pour le traitement des contenus sémantiques et émotionnels au cours du développement.

Nous exposerons ensuite les résultats préliminaires de la septième et dernière étude, la totalité des inclusions n'étant pas atteinte au jour de l'impression de ce document. Cette étude analyse, avec la même méthodologie que celle du sixième article, les capacités de catégorisation

sémantique et émotionnelle d'enfants qui présentent un TDA/H (âges de 7 à 12 ans). Les résultats préliminaires montrent des capacités de catégorisation sémantique globalement plus faibles chez les enfants avec TDA/H que dans le développement typique, mais avec un effet d'interférence du contenu émotionnel négatif semblable entre les deux groupes. Ceci évoque un traitement non volontaire, qui dépend de l'attention exogène, du contenu émotionnel d'images présentées rapidement préservé chez les enfants TDA/H. La tâche de catégorisation émotionnelle ne met en évidence aucune différence significative entre les deux groupes. Ce résultat suggère que le traitement rapide des informations visuelles émotionnelles est efficient chez les enfants TDA/H. Leurs difficultés de catégorisation des émotions, décrites dans la littérature sembleraient donc liées à un dysfonctionnement de traitements plus tardifs c'est-à-dire relatif à l'engagement de l'attention soutenue et des capacités de contrôle inhibiteur. Le déficit attentionnel inhérent à ce trouble pourrait être la cause de la perturbation de traitement des informations émotionnelles, soit par un défaut d'attention soutenue, qui empêcherait d'analyser plus finement les informations, soit par un défaut de désengagement de l'attention des informations émotionnelles.

Enfin, dans un dernier chapitre, nous discutons les résultats de l'ensemble des travaux réalisés dans le cadre de cette thèse afin de proposer un modèle dynamique de la perception visuelle dans lequel la priorisation d'une dimension de traitement sur une autre n'est pas préétablie mais semble varier en fonction des contraintes liées à l'environnement et à l'observateur.

DONNEES THEORIQUES

Chapitre I. Traitements sémantiques et émotionnels des scènes visuelles complexes chez l'adulte

Ce premier chapitre présente une revue des connaissances actuelles principales sur les traitements sémantiques et émotionnels, ainsi que leurs éventuelles interactions, dans la perception de scènes visuelles. Ce chapitre est constitué de la reproduction d'un article de revue de littérature accepté pour publication :

Leroy, A., Spotorno, S., & Faure, S. (accepté le 26/02/2019 pour publication). Traitements sémantiques et émotionnels des scènes visuelles complexes : une synthèse critique de l'état actuel des connaissances. *L'Année Psychologique/Topics in Cognitive Psychology*.

I.1 Article I. Traitements sémantiques et émotionnels des scènes visuelles complexes : une synthèse critique de l'état actuel des connaissances

Traitements sémantiques et émotionnels des scènes visuelles complexes :
une synthèse critique de l'état actuel des connaissances

Anaïs Leroy^a, Sara Spotorno^b, Sylvane Faure^a

^a Laboratoire d'Anthropologie et de Psychologie Cliniques, Cognitives et Sociales (LAPCOS),
Université Côte d'Azur, France.

^b School of Psychology, Keele University, United Kingdom.

Titre courant : Traitements sémantiques et émotionnels des scènes visuelles

Correspondance :

Anaïs Leroy
LAPCOS, MSHS Sud Est
Pôle Universitaire Saint Jean d'Angély
24 avenue des Diables Bleus
06357 Nice Cedex 4
Tel. : 04 89 88 14 60
Fax : aucun
Email : Anais.LEROY@univ-cotedazur.fr

Résumé

Nous avons la capacité de reconnaître très rapidement une scène visuelle même présentée un très court instant. Cette capacité concerne le sens de la scène mais également l'émotion qu'elle véhicule. Cette revue critique présentera d'abord une synthèse des principales recherches menées sur les traitements des contenus sémantiques et, puis, sur les traitements des contenus émotionnels lors de la perception des scènes. Il s'agira de mieux comprendre comment ces informations sont sélectionnées et représentées, et comment elles influencent l'exploration visuelle, contribuant ainsi à la rapidité du traitement de la scène. Nous questionnerons ensuite les éventuelles interactions entre les traitements sémantiques et émotionnels. Cette revue met en exergue l'importance des liens sémantiques et émotionnels entre les différents constituants d'une scène et le caractère dynamique, intégratif de la perception visuelle. Elle suggère également divers futurs axes de recherche.

Abstract:

Title: Processing of complex visual scenes: Between semantic and emotion understanding

We have a great ability to recognize complex, everyday visual scenes, even from very short presentations. This ability concerns the general meaning of the scene and the emotional content conveyed. The present review will first analyze previous research on how semantic information and emotional information within a scene are selected and represented, and how they influence the inspection of the scene. It will then examine the interplay between these two kinds of information during processing. This review highlights the importance of the semantic and emotional links between the different components of a scene and the dynamic, integrative nature of visual perception. It also suggests several areas for future research.

Nous avons la capacité de comprendre très rapidement une scène visuelle et de décrire ce que nous voyons en un seul coup d'œil, par exemple lorsque nous regardons une photographie, lorsque nous changeons les chaînes de télévision ou encore lorsque les portes de l'ascenseur s'ouvrent sur une pièce inconnue ... L'expérience subjective de la perception visuelle d'une scène nous fait apparaître cette scène telle un ensemble cohérent perçu instantanément. Cependant la fonction visuelle est très complexe. Elle nécessite la transduction neuronale des inputs lumineux reçus par la rétine, puis de nombreux relais entre les différents centres corticaux et sous corticaux impliqués afin d'accéder à une expérience consciente et cohérente de la scène visuelle. Dans un premier temps, nous devons extraire les informations les plus pertinentes parmi un flot de stimulations sensorielles, riche et continu et qui dépasse ainsi largement nos capacités de traitement. Cette priorisation, nécessaire à l'adaptation à notre environnement physique et social, est réalisée sur différents aspects élémentaires ou plus complexes, qui peuvent interagir : d'une part, les caractéristiques perceptives du stimulus, qui peuvent capturer l'attention exogène quand elles sont saillantes, et d'autre part, les représentations cognitives qui guident l'attention endogène (Corbetta & Shulman, 2002 ; Spotorno, Malcolm, & Tatler, 2014). Les différents stimuli provenant de notre environnement doivent être associés, interprétés et reconnus chacun en tant qu'entité afin de permettre une vision cohérente, globale et stable du monde visuel qui nous entoure. L'intégralité de ce processus de perception, de sélection et de traitement de l'information visuelle doit se faire en quelques millisecondes (ms) pour nous permettre d'appréhender notre environnement et de réagir au mieux. De nombreuses études ont analysé l'interaction entre les facteurs perceptifs (de bas niveau) et les facteurs cognitifs (de haut niveau) sur le déploiement de l'attention visuelle durant l'exploration de scènes visuelles (Borji, Sihite, & Itti, 2013 ; Hwang, Wang, & Pomplun, 2011 ; Koehler, Guo, Zhang, & Eckstein, 2014 ; Onat, Açık, Schumann, & König, 2014 ; Wu, Wang, & Pomplun, 2014). Ces études soulignent que, bien que la saillance perceptuelle (due à la luminosité, aux couleurs, aux orientations ...) puisse attirer l'attention visuelle, le contenu sémantique (par ex. pour une revue, Malcolm, Groen, & Baker, 2016) et le contenu émotionnel (Lang, Bradley, & Cuthbert, 1997 ; Nordström & Wiens, 2012 ; Simola, Le Fevre, Torniainen, & Baccino, 2015) ont un rôle prépondérant pour déterminer les informations prioritairement sélectionnées et analysées.

Cette revue propose une synthèse critique des recherches sur l'influence des informations sémantiques et émotionnelles lors de la perception visuelle. Il s'agit en particulier de comprendre comment les contenus sémantiques et émotionnels participent à la sélection des informations et influencent la façon dont les scènes visuelles sont explorées et interprétées,

contribuant ainsi à la rapidité du traitement des informations visuelles. Cette revue ne se veut pas exhaustive, la littérature sur le traitement des informations visuelles étant très vaste, mais a pour objectif de permettre au lecteur d'accéder à une vision intégrative des processus à l'œuvre lors de la perception visuelle rapide et de questionner certains aspects encore peu abordés dans cette littérature. En effet, alors que le traitement sémantique des informations visuelles d'un côté et celui des contenus émotionnels des informations visuelles de l'autre ont suscité de nombreuses études, la recherche sur les éventuelles relations entre ces deux types de traitements est bien moins riche. Nous ferons donc d'abord un état des lieux sur les travaux principaux réalisés sur les traitements des informations sémantiques visuelles, puis sur ceux réalisés sur les traitements des informations émotionnelles visuelles, afin de questionner ensuite leurs éventuelles interactions.

1. Traitement des informations sémantiques

1.1 Les scènes visuelles complexes

Toute scène visuelle de la vie quotidienne (autrement nommée « scène complexe » ou « scène naturelle ») est caractérisée par un contexte (niveau global, arrière-plan comme par exemple une forêt, ou une salle à manger) et un ou plusieurs « objets » correspondant à des entités vivantes ou non (niveau local, porteur de détails, comme par exemple un faon, ou des assiettes). En outre une scène visuelle présente des régularités et des propriétés spécifiques concernant sa structure globale. Celle-ci répond à des contraintes physiques (relatives à la structure, aux positions relatives des éléments présents) et sémantiques. Biederman, Mezzanotte et Rabinowitz (1982) ont réalisé une série d'expériences dans lesquelles ils ont manipulé la cohérence spatiale et sémantique de la scène en introduisant cinq types de violations : en faisant flotter des objets dans l'air (support), en laissant apparaître l'arrière-plan en transparence dans l'objet (interposition), en plaçant des objets dans une position inappropriée ou en modifiant leur taille (orientation et taille) et enfin en insérant des objets incohérents sémantiquement avec la scène (plausibilité). Les scènes étaient présentées rapidement (150 ms) à des participants qui devaient à chaque essai juger si l'objet inclus dans la scène correspondait à un objet qui venait de leur être nommé. La réponse de reconnaissance de l'objet perdait en exactitude et en vitesse lors de ces violations, ce qui met en avant l'importance des attentes du sujet (processus *top down*) lors de la perception visuelle. Ainsi, par vos expériences et leur répétition, lorsque vous devez chercher un sèche-cheveux, vous allez plutôt vous attendre à en trouver un dans une salle de bain que dans une cuisine. De telles attentes sur les relations entre les objets (niveau local)

et le contexte (niveau global) sont ainsi déterminées par vos représentations en mémoire à long terme (Biederman, Mezzanotte, & Rabinowitz, 1982 ; Friedman, 1979 ; Eckstein, Drescher, & Shimozaki, 2006 ; Spotorno, Malcolm, & Tatler, 2015). Ce traitement des stimuli de niveau global ou local a été initialement analysé au moyen de la tâche de Navon (1977) dans laquelle une grande lettre, le niveau global, est constituée de petites lettres réalisant le niveau local et où ces niveaux global/local peuvent être cohérents ou non : par exemple un grand « H » est constitué de petits « h » en situation de cohérence ou de petits « s » en situation d'incohérence. Des études ont adapté ce type de tâche – stimuli hiérarchiques – avec des stimuli naturels comme des animaux (Studer & Hübner, 2008) ou des scènes visuelles complexes (Peyrin, Mermillod, Chokron, & Marendaz, 2006) et ont mis en évidence un biais de traitement en faveur du niveau global, celui-ci étant perçu plus rapidement que le niveau local. Cette différence de vitesse de traitement entre stimulus de niveau global et local s'expliquerait par une composition spectrale différente. En effet un stimulus visuel est décomposé par le système visuel en différentes fréquences spatiales, qui correspondent à une variation plus ou moins rapide du contraste (Ginsburg, 1986). Les informations globales (*e.g.* le contexte pour les scènes visuelles) sont véhiculées par les basses fréquences spatiales (variations lentes du contraste) alors que le niveau local (*e.g.* les objets, les éléments de premier plan dans les scènes visuelles) correspond majoritairement aux hautes fréquences spatiales (variations rapides du contraste). Le système visuel est caractérisé par plusieurs dualités, la dualité des voies ventrale et dorsale (décrite ci-dessous, paragraphe 1.2) et la dualité entre les voies magnocellulaire et parvocellulaire. Le signal lumineux reçu par la rétine est traité par les photorécepteurs (cônes et bâtonnets) qui vont le traduire en signal électrique transmis aux neurones bipolaires puis aux cellules ganglionnaires. Ces cellules ganglionnaires sont divisées en deux groupes selon leur fonction : les cellules magnocellulaires et les cellules parvocellulaires. Les cellules magnocellulaires présentent de grands champs récepteurs, elles répondent à des stimulations brèves et elles sont sensibles au mouvement, alors que les cellules parvocellulaires présentent de plus petits champs récepteurs, elles répondent de façon tonique, continue, à la stimulation visuelle et elles sont sensibles aux variations de couleur (Bullier, 1998). Les cellules parvocellulaires répondent ainsi aux hautes fréquences spatiales et permettent un traitement de l'information plus détaillé (niveau local) que l'information extraite des basses fréquences spatiales (niveau global) par les cellules magnocellulaires. Cependant les voies magnocellulaires véhiculent l'information issue des cellules magnocellulaires plus rapidement que ne le sont les informations issues des hautes fréquences véhiculées par les voies parvocellulaires (Peyrin *et al.*, 2006 ; Schyns & Oliva, 1994).

1.2 Processus de catégorisation sémantique

Nous catégorisons, c'est-à-dire que nous classons constamment, de façon automatique, non volontaire (Greene & Fei-Fei, 2014), l'environnement visuel selon des attributs perceptifs (Rosch & Lloyd, 1978) ou sémantiques (Murphy & Medin, 1985). Par exemple, plutôt que de voir de façon isolée chaque grain de sable vous voyez une plage, et plutôt que de voir quatre pattes, de la fourrure et deux petites oreilles vous voyez un chat. Enfin, si vous voyez un dalmatien, un caniche et un doberman vous pouvez les regrouper sous la catégorie « chien ». Ce processus de catégorisation est essentiel pour éviter une surcharge des capacités cognitives (Hanrnat, 1978) et pour donner rapidement sens au monde qui nous entoure et nous y adapter en continu. Pour étudier le temps nécessaire à la catégorisation, Thorpe, Fize et Marlot (1996) ont utilisé les potentiels évoqués cognitifs (PE) dans une tâche de détection de la présence d'un animal dans une scène. Les images étaient présentées pour seulement 20 ms et la catégorisation visuelle était réalisée en 150 ms en moyenne. Ces résultats ont été reproduits en utilisant d'autres catégories de cibles, non biologiques, comme des moyens de transport (VanRullen & Thorpe, 2001a) et en présentant la scène latéralisée (Fize, Fabre-Thorpe, Richard, & Doyon, 2005). La catégorisation d'un objet se réalise ainsi en quelques centaines de millisecondes et automatiquement à un niveau optimal de performances (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001).

Au niveau cérébral, l'information visuelle est traitée par les différentes aires visuelles associatives en fonction de leurs caractéristiques, depuis V1 pour l'analyse de l'orientation des stimuli visuels, et se poursuit de façon hiérarchique : tout d'abord la catégorisation de base dans le cortex temporal antérieur, puis le traitement des relations contextuelles dans le cortex parahippocampique et enfin celui des relations sémantiques au niveau du cortex préfrontal (Bar, 2004). Selon l'hypothèse des deux voies de traitement visuel de Milner et Goodale (sans questionner ici leurs organisations, fonctions et interactions), la voie ventrale ou occipito-temporale, reliant les aires pré-striées au cortex inféro-temporal, permet la discrimination des objets visuels et leur attribution sémantique, dont l'accès à la fonction. La seconde voie, la voie dorsale ou occipito-pariétale, rejoint le cortex pariétal postérieur et permet le traitement de l'orientation spatiale, du mouvement, les comportements visuo-moteurs et les représentations pragmatiques (Milner & Goodale, 2006). Des études électrophysiologiques ont mis en évidence un traitement plus rapide de la voie dorsale (Bullier, Girard, & Salin, 1994) par rapport à la voie ventrale. Ces deux voies visuelles sont donc ascendantes depuis V1 mais elles présentent également des connexions réciproques descendantes (en *feed-back*) (Felleman & Van, 1991), ce qui suggère que la compréhension des scènes visuelles (contexte et objets) se réalise en

fonction des informations perceptives mais également des attentes du sujet et de ses représentations mnésiques. Cette hypothèse importante d'une influence des processus cognitifs sur la perception visuelle sera abordée plus en détail dans les prochaines sections.

1.3 Interaction entre les informations sémantiques contextuelles et locales

Le lien sémantique entre les différents constituants d'une scène visuelle (*i.e.* le contexte et les objets inclus dans ce contexte) influence nos capacités de traitement de cette scène. En effet, la catégorisation d'un objet sera plus rapide et plus exacte lorsque cet objet est inclus dans un contexte qui lui est sémantiquement cohérent (*i.e.* selon une association fréquente dans la vie quotidienne, par exemple un chameau et un désert) en comparaison à sa présentation dans un contexte sémantiquement incohérent (*e.g.* un chameau et une banquise). Cet effet dit de cohérence sémantique a été mis en évidence selon une double direction : la perception du contexte de la scène facilite la catégorisation des objets cohérents (Davenport, 2007 ; Davenport & Potter, 2004 ; Munneke, Brentari, & Peelen, 2013 ; Palmer, 1975 ; Rémy, Vayssière, Pins, Boucart, & Fabre-Thorpe, 2014 ; Sun, Simon-Dack, Robert, & Wolfgang, 2011), et la cohérence des objets inclus facilite la catégorisation du contexte (Davenport & Potter, 2004 ; Joubert *et al.*, 2007; Munneke *et al.*, 2013). L'hypothèse sur le plan des substrats cérébraux est que l'objet et le contexte activent chacun le réseau neuronal de leur catégorie. Lorsque l'objet et le contexte sont cohérents, leurs réseaux catégoriels sont proches et l'activation de ces populations de neurones adjacentes permet une sélection et une propagation de l'information plus rapide. Au contraire, en situation d'incohérence, il y a une interférence conflictuelle et la prise de décision est plus lente (Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007). L'ensemble de ces résultats comportementaux et en neuro-imagerie suggère que contexte et objet sont traités en parallèle dès les premiers instants, et que ces traitements interagissent très précocement (Bar, 2004 ; Wolfe, Võ, Evans, & Green, 2011). La présence d'une incohérence sémantique dans une scène visuelle perturbe donc le traitement automatique des informations visuelles : l'observateur détecte une « anomalie » en regard de ses attentes et ses représentations sur les scènes visuelles définies comme stables et devant correspondre aux réalités physiques et sémantiques du monde réel. Ainsi, dans des tâches en oculométrie, où l'on enregistre les mouvements oculaires de l'observateur lors de l'exploration visuelle d'une scène, il a été noté un temps de fixation des objets incohérents plus long que celui des objets cohérents avec le contexte (Võ & Hendersen, 2009). Pour certains auteurs les objets incohérents n'attireraient pas l'attention mais une fois fixés, ils seraient explorés plus longuement que les objets cohérents,

cela afin de lever l'ambiguïté sur leurs identités (Friedman, 1979 ; Hollingworth & Henderson, 1999, 2003 ; Mack, Clarke, Erol, & Bert, 2017). Des activations plus fortes en EléctroEncéphaloGraphie (EEG) et en Imagerie par Résonance Magnétique fonctionnelle (IRMf) sont également recueillies lors du traitement de scènes visuelles présentant une incohérence sémantique entre l'objet et le contexte en comparaison à la condition de cohérence. Ces activations sont néanmoins modulées par le type de tâche et de violation de la structure de la scène. Par exemple, on relève un effet d'incohérence plus fort lors d'une incohérence sémantique par comparaison à une variation de la plausibilité d'apparition d'un objet dans la scène (Coco, Araujo, & Petersson, 2017).

1.4 Modèles explicatifs de la reconnaissance des scènes visuelles

Plusieurs théories ont été proposées pour expliquer les mécanismes d'identification du contexte et de l'objet d'une scène visuelle et spécifier à quelle(s) étape(s) du processus de traitement des informations visuelles ces deux niveaux sont intégrés.

1.4.1 L'objet et le contexte sont traités de façon isolée

Le modèle d'isolation fonctionnelle (Hollingworth & Henderson, 1998 ; 1999) décrit une identification isolée de l'objet sans influence du contexte dans lequel il est inclus. Dans ce modèle le contexte et l'objet seraient traités en parallèle de façon indépendante, l'intégration des informations ne se faisant qu'à un stade post-perceptif, décisionnel. Hollingworth et Henderson, (1998) ont répliqué l'expérience de Biederman et collaborateurs (1982) (voir supra) en modifiant certains aspects de méthodologie. Biederman et collaborateurs donnaient, dans leur tâche originale, l'identité de l'objet cible aux participants avant l'apparition de la scène, et la facilitation contextuelle mise en évidence pouvait donc être interprétée comme provenant d'un biais de réponse et de l'influence de la scène lors de la prise de décision. Par exemple si l'observateur doit juger si une vache est présente ou non dans une scène et qu'on lui présente un paysage avec une ferme il aura plus tendance à répondre par l'affirmative que si on lui présente un bureau, même s'il n'a pas du tout vu l'objet en question (la vache). Hollingworth et Henderson ont donc répliqué l'étude de Biederman et collaborateurs mais en affichant après la présentation de la scène deux objets parmi lesquels les participants devaient choisir celui qu'ils avaient vu : aucun effet facilitateur du contexte sur la reconnaissance de l'objet n'a alors été mis en évidence. Cette absence d'effet facilitateur leur a permis de faire l'hypothèse d'une

identification de l'objet isolée et non en lien avec le contexte. Pour eux, l'effet de cohérence sémantique mis en évidence dans certaines études proviendrait donc d'un biais expérimental.

1.4.2 Le traitement de l'objet est précurseur à la catégorisation de la scène visuelle

Le modèle d'isolation fonctionnelle suppose ainsi un traitement isolé de l'objet et du contexte, d'autres modèles postulent que les informations issues de la présence d'un ensemble d'objets, de leurs relations spatiales et de la probabilité de leur cooccurrence, permettraient d'extraire le sens de la scène visuelle (Friedman, 1979 ; Biederman, 1987; De Graef, Christiaens, & d'Ydewalle, 1990). Selon ces hypothèses, les objets seraient systématiquement catégorisés avant les scènes, et la catégorisation de la scène serait ainsi secondaire à la catégorisation des objets (Riesenhuber & Poggio, 2000). Cependant, les approches centrées sur l'analyse de l'objet sont difficilement compatibles avec la grande rapidité de catégorisation du contexte (Joubert *et al.*, 2007). L'absence de détection d'un changement d'objet lorsque le contexte reste le même (Greene & Oliva, 2009) ou encore la possibilité de reconnaître une scène visuelle par la simple présence de contours morcelés, dans de faibles résolutions spatiales ou par une configuration particulière de couleurs (Oliva & Schyns, 2000) vont également à l'encontre de ces hypothèses théoriques. Par exemple, dans une série de trois expériences, Oliva et Schyns (2000) ont présenté rapidement (i.e. 120ms) des scènes visuelles complexes dans lesquelles les couleurs d'origine étaient ou non modifiées. La catégorisation des scènes était plus rapide lorsque les scènes étaient présentées dans leur couleur d'origine que dans des couleurs modifiées. Les auteurs ont conclu qu'une configuration de couleurs diagnostique d'une scène favorisait la reconnaissance de celle-ci, ce qui souligne le rôle des informations globales, diffuses, dans la catégorisation des scènes visuelles complexes.

1.4.3 Le traitement du contexte permet d'amorcer la reconnaissance de l'objet

D'autres hypothèses sont en faveur d'un traitement global, contextuel de la scène plutôt que local. L'approche « global vers local » (« *coarse to fine* » : Schyns & Oliva, 1994 ; Oliva & Schyns, 1997) décrit une analyse grossière préalable à une analyse secondaire plus fine des informations spatiales, analyses sous-tendues par la différence entre les fréquences spatiales basses et hautes respectivement. Oliva et Schyns ont mis en évidence le rôle des fréquences spatiales dans la perception visuelle en fonction du temps d'exposition par la présentation de stimuli hybrides : les auteurs ont présenté des stimuli visuels superposant les hautes fréquences spatiales d'une scène sur les basses fréquences d'une seconde scène pour des durées d'exposition variées afin d'évaluer le décours temporel du traitement des fréquences spatiales

(passage d'une analyse « *coarse to fine* » à une analyse « *fine to coarse* »). Le modèle de facilitation contextuelle de Bar (2004) rejoint cette conception de traitement « global vers local ». Selon ce modèle, en effet, les informations sur le contexte de la scène visuelle facilitent le traitement de l'objet à un stade précoce. Ainsi, lorsque nous regardons une scène, nous accédons à un premier traitement d'informations approximatives, peu détaillées, sur le contexte général de la scène (niveau global), provenant des basses fréquences spatiales et transmises majoritairement par l'activation des neurones de la voie magnocellulaire, la plus rapide. Cette reconnaissance rapide du sens général (*gist*) de la scène nous permet d'activer des représentations en mémoire à long terme et donc des attentes sur le type d'objet fréquemment associé à ce contexte. Des « hypothèses perceptives » sur l'objet à reconnaître sont ainsi créées dans le cortex orbito-frontal. Il s'en suit une compétition entre ces hypothèses pour déterminer l'objet réellement perçu grâce au cumul progressif d'informations plus fines sur l'objet (provenant des hautes fréquences spatiales) transmises par les neurones de la voie parvocellulaire. La représentation de l'objet peut alors être sélectionnée parmi les différents patrons représentationnels qui avaient été préactivés (Bar, 2004; Kauffmann, Ramanoël, & Peyrin, 2014; Kauffmann, Chauvin, Guyader, & Peyrin, 2015 ; Wolfe *et al.*, 2011). Des processus de haut niveau (*top down*), liés notamment aux attentes du sujet et à ses expériences antérieures, influenceraient ainsi la perception des objets (Ahissia, & Hochstein, 2004 ; Coco *et al.*, 2017 ; Greene, Botros, Beck, & Fei-Fei, 2015). Il peut même arriver que ces attentes « masquent » certaines incohérences d'objets présents dans une scène en leur attribuant une autre identité plus cohérente avec les représentations préexistantes (Mack *et al.*, 2017), créant ainsi ce qui est considéré comme un « déni perceptif » (Bruner & Postman, 1949). Par exemple, Mack et collaborateurs (2017) ont présenté des scènes visuelles contenant parfois des incohérences sémantiques, comme une femme léchant un tube de néon. La majorité des participants décrivait cette image comme une femme léchant une glace plutôt qu'un néon, et ce, que la scène soit présentée durant 100 ou 200 ms. Dans notre quotidien, le traitement plus précoce des informations contextuelles permettrait de réduire l'ambiguïté des objets en contraignant leur catégorisation. L'influence du contexte interviendrait ainsi aux étapes précoces de l'analyse perceptive de l'objet, en modulant directement l'influence de l'ambiguïté au début des processus de reconnaissance (Dyck & Brodeur, 2015). Les attentes du sujet influenceraient donc la perception visuelle dès les stades précoces, perceptifs, de la perception visuelle.

1.4.4 L'objet et le contexte sont traités de manière interactive

Davenport et Potter (2004) proposent l'hypothèse d'un traitement interactif (voir également Joubert, Fize, Rousselet, & Fabre-Thorpe, 2008 ; Joubert *et al.*, 2007) selon laquelle les traitements des objets et du contexte interagissent, chaque niveau pouvant faciliter la reconnaissance des autres niveaux. Enfin, selon d'autres auteurs (Bacon-Mace, Thorpe, & Fabre-Thorpe, 2005; Fabre-Thorpe *et al.*, 2001 ; VanRullen & Thorpe, 2001b), l'interaction entre l'objet et le contexte pourrait se faire dès le début du traitement des informations visuelles ce qui expliquerait l'effet de cette interaction sur les réponses comportementales très précoces, à un stade où seul un traitement uniquement ascendant, sans retour d'informations (de type « feed-forward ») est possible. C'est ce que montre l'étude de VanRullen et Koch (2003), avec un paradigme de catégorisation de scènes naturelles présentées très rapidement (26ms) et masquées : le traitement visuel peut effectivement se réaliser uniquement dans un mode *feed-forward*, les réponses des participants étant obtenues avant la possibilité temporelle d'un contrôle *feed-back*.

Ainsi, d'une conception initiale mettant l'accent essentiellement sur l'influence du niveau local (de l'objet), les théories explicatives de la perception des scènes visuelles ont évolué vers la prise en compte d'une influence plus grande du niveau global et/ou d'une interaction précoce entre ces deux niveaux global et local. La question du déroulement du traitement des scènes visuelles et plus précisément du traitement des associations contexte-objet au sein de cette scène reste néanmoins toujours débattue, notamment l'étape à laquelle les informations de l'objet et du contexte sont mises en commun : à un stade perceptif ou plus décisionnel.

1.5 Liens perception / cognition

De ce débat relatif aux relations contexte-objet découle une autre question fondamentale qui est celle du lien entre les traitements perceptifs et cognitifs. Les processus perceptifs du traitement des scènes visuelles se déroulent-ils selon l'hypothèse d'un modèle classique de traitement des informations sensorielles indépendant des connaissances sémantiques et des attentes (Pylyshyn, 1980), c'est-à-dire selon une séquentialité entre des traitements perceptifs suivis par des traitements cognitifs de plus haut niveau (i.e. la perception est encapsulée dans la cognition) ? Ou bien notre perception est-elle influencée par nos attentes et nos représentations (i.e. il existe une influence *top down* et la perception serait perméable à la cognition) ? Pour certains auteurs (pour une revue critique, Firestone & Scholl, 2016) les études

mettant en avant une influence *top down* ne sont pas fiables car elles présentent de nombreux biais méthodologiques, notamment parce qu'il est compliqué de différencier une modification réelle de l'expérience perceptive d'une modification du jugement perceptif (dans le sens de la réponse donnée par le participant à un niveau décisionnel et non plus perceptif). Schématiquement, on vous présente une image avec une personne qui mange une perle et vous décrivez cette image comme une personne qui mange une cerise et non une perle. Peut-on dire que votre système visuel perçoit vraiment une cerise (vos attentes ont influencé votre perception) ou bien, que vous avez perçu une perle mais qu'à un stade décisionnel, de traitement conscient de l'information, vous avez jugé ce stimulus comme une cerise et non comme une perle ? Il semble en effet questionnable de s'appuyer uniquement sur des résultats comportementaux (tels que l'exactitude de la catégorisation et les temps de réponse) pour différencier une origine perceptive versus décisionnelle de cette modification de l'interprétation. La neuroimagerie, en permettant un accès au fonctionnement cérébral lors de la perception de ces ambiguïtés, a permis d'enrichir ce débat (Truman & Mudrick, 2018). De nombreux travaux empiriques soutiennent ainsi la théorie de la pénétrabilité cognitive de l'expérience perceptive, théorie selon laquelle les facteurs cognitifs, notamment les connaissances sémantiques liées aux probabilités de rencontrer plus fréquemment certains objets dans des contextes spécifiques, influencent les traitements perceptifs (Brandman & Peelen, 2017 ; O'Callaghan, Kveraga, Shine, Adams, & Bar, 2017 ; Teufel & Nanay, 2017 ; Truman & Madrick, 2018). Une étude récente (Brandman & Peelen, 2017) a utilisé l'IRMf et la MagnétoEncéphaloGraphie (MEG) pour approcher la neuroanatomie fonctionnelle et les dynamiques neuronales qui sous-tendent l'effet de facilitation sémantique liée aux interactions entre contexte et objet lors du traitement de scènes visuelles complexes.

Différentes tâches sont proposées : deux tâches dites de « oddball » (les participants doivent répondre lorsqu'un chiffre est présenté à la place d'une image d'objet et/ou de scène visuelle) réalisées en IRMf et en MEG, et deux tâches comportementales, une tâche de reconnaissance d'objet et une tâche de détection de la présence d'un objet animé. Les mêmes stimuli sont utilisés dans les différentes tâches et sont présentés très rapidement (50ms). Il s'agit d'images présentant soit un objet dégradé seul (par une diminution du nombre de pixels), soit cet objet dégradé au sein d'une scène visuelle, ou enfin, d'images de la scène seule.

Cette étude met en avant une facilitation de la reconnaissance de l'objet grâce aux informations contextuelles et qui se réalise dans le cortex visuel à un stade perceptif. Cette facilitation serait réalisée grâce à des boucles de contrôle *feed-back* intervenant entre 320 et 340 ms après l'affichage de la scène. Notons que ces résultats vont, en outre, dans le sens du modèle

de Bar, 2004 et à l'encontre des théories d'un traitement agissant de façon *feed-forward* défendues notamment par VanRullen & Koch (2003).

En résumé, les études récentes sur le traitement des scènes visuelles remettent en question la séparation classique perception/cognition. L'interprétation des stimuli perceptifs et leur intégration en un tout cohérent seraient influencées par nos attentes et nos représentations. Des recherches sont encore néanmoins nécessaires afin de déterminer plus précisément la dynamique de traitement des associations sémantiques entre contexte et objet.

2. Traitement des informations émotionnelles

Lorsque nous percevons une scène visuelle nous n'avons pas uniquement accès à sa signification mais également aux ressentis émotionnels qui y sont liés (Barret & Bar, 2009 ; Malcolm *et al.*, 2016). Les informations émotionnelles ont un statut particulier car elles sont traitées de façon prioritaire par rapport aux informations « neutres » c'est-à-dire dépourvues de contenu émotionnel (Carretié, 2014 ; Pilarczyk & Kuniecki, 2014 ; Vuilleumier, 2015). Pour de nombreux auteurs, la présence d'un contenu émotionnel est en effet à l'origine d'un « biais affectif » dans la perception (pour des revues, Mulckhuyse, 2018 ; Pourtois *et al.*, 2012 ; Vuilleumier, 2015), qui influence l'exploration visuelle et renforce le traitement précoce du stimulus émotionnel (Phelps, Ling & Carrasco, 2006). Les aires corticales visuelles dès V1 sont ainsi plus fortement activées en présence d'un contenu émotionnel que pour un stimulus visuel qui en est dépourvu (Vuilleumier, 2005 ; 2015). La perception visuelle serait donc modulée par la présence d'un contenu émotionnel pour permettre un traitement plus rapide et plus précis du stimulus émotionnel afin d'adapter notre comportement le plus efficacement. Par exemple, si vous percevez un ensemble de traits fins noirs à proximité, vous allez automatiquement suspendre la tâche en cours pour fixer cet objet afin de l'identifier et pouvoir ajuster votre comportement suivant qu'il s'agisse d'une araignée ou de fils. Dans cette seconde partie nous nous intéressons aux particularités de traitement des informations émotionnelles présentes au sein des scènes visuelles.

2.1 Les scènes visuelles émotionnelles

Parmi les scènes visuelles complexes, les scènes à contenu émotionnel (ci-après « scènes émotionnelles ») forment une catégorie particulière, d'une grande importance dans les comportements et les interactions de la vie de tous les jours et qui est étroitement liée à la qualité de notre expérience subjective. Elles sont caractérisées par la présence d'un contexte

sémantique à valence émotionnelle (positive ou négative) et/ou d'éléments locaux (objets de la vie quotidienne, animaux, personnes, visages) porteurs d'une émotion (Sabatinelli, Fortune, Li, Siddiqui, Krafft, Oliver, Beck, & Jeffries, 2011). Historiquement, et toujours de façon majoritaire, les stimuli employés pour évaluer l'influence ou la reconnaissance des émotions sont des expressions faciales. Cependant nous ne sommes jamais confrontés dans notre quotidien à un visage qui survient isolé et nous devons extraire le contenu émotionnel parmi divers niveaux d'informations. Aussi, nous ne voyons pas toujours un visage de face et en vision centrale (fovéale), nous ne le voyons parfois que partiellement ou en vision périphérique. Qui plus est, y compris vues dans des conditions optimales, les expressions faciales sont ambiguës et une même expression peut exprimer différentes émotions : par exemple nous pouvons pleurer de joie ou de tristesse. Aviezer, Trope et Todorov (2012) ont présenté des photographies de visages de joueurs de tennis et ils ont demandé aux participants de dire si le joueur avait gagné ou perdu. Lorsque les participants ne voyaient que les visages des joueurs ils ne parvenaient pas à prédire correctement cette issue du match alors que l'observation des attitudes corporelles (sans voir le visage, celui-ci étant masqué) leur permettait de répondre correctement. Ces résultats ont été répliqués avec des photographies d'adultes et d'enfants exprimant une émotion intense (Wenzler, Levine, van Dick, Oertel-Knöchel, & Aviezer, 2016). Ces études mettent en question la conception d'une reconnaissance précise et innée des expressions émotionnelles faciales quels que soient les contextes dans lesquels elles surviennent (pour une revue de ces travaux et une discussion, Ekman, Friesen, & Ellsworth, 2013). Nous devons pouvoir inférer des émotions à partir d'attitudes corporelles, d'interactions sociales, et à propos d'éléments non humains tels que des animaux ou des objets. Enfin, nous devons pouvoir inférer le contenu émotionnel d'informations contextuelles. L'utilisation de scènes visuelles à contenu émotionnel (comme par exemple une photographie de l'effondrement d'un bâtiment, d'un tsunami ou à l'inverse du bal lors de la fête d'un village ou de personnes joyeuses devant un coucher de soleil) semble donc pouvoir permettre une approche plus écologique et valide du traitement des informations émotionnelles visuelles que celle de visages isolés. Cependant les scènes émotionnelles provoquent-elles des réactions émotionnelles différentes de celles en réponse à des visages émotionnels ? Au niveau comportemental (temps de réponse), une méta-analyse de 26 études (Carretié, 2014) conclut à une capture de l'attention exogène identique entre visages et scènes émotionnelles. En termes de réactions somatiques (ici la conductance cutanée) et des enregistrements en EEG, les scènes émotionnelles auraient un impact plus fort que les visages émotionnels (Wangelin, Bradley, Kastner, & Lang, 2012). Ainsi, les scènes neutres (*e.g.* qui ne comportent pas d'éléments émotionnels) auraient le même impact que des visages porteurs

d'expression émotionnelle, alors que les scènes « émotionnelles » avec fort niveau d'activation physiologique (« arousal ») susciteraient une activité cérébrale plus forte (Bradley, Sabatinelli, Lang, Fitzsimmons, King, & Desai, 2003) que les visages. Ces résultats sont cependant discutés, car pour d'autres auteurs (Hariri, Tessitore, Mattay, Fera, & Weinberger, 2002) toutes les émotions ne se valent pas. Les visages menaçants (expressions de peur et de colère) entraîneraient une activation plus forte de l'amygdale que les scènes émotionnelles. Les visages émotionnels attireraient plus précocement l'attention que les scènes émotionnelles, mais avec une activité en EEG d'intensité plus faible (Mavratzakis, Herbert, & Walla, 2016). Des différences qualitatives ont également été rapportées, avec des profils d'activations distincts sur le plan spatial et temporel en EEG et des résultats comportementaux (exactitude et temps de réponse) différents entre scènes et visages émotionnels. La littérature a ainsi mis en évidence ce qui est appelé un « biais négatif » pour les scènes émotionnelles (réactions plus fortes et performances meilleures pour les scènes émotionnelles négatives que pour les positives) et un biais vers le contenu émotionnel positif pour les visages (Carretié, Kessel, Carboni, López-Martín, Albert, Tapia, Mercado, Capilla, & Hinojosa, 2013). Cependant, cette notion de biais positif pour les visages est à considérer avec précaution. La joie apparaît certes identifiée plus rapidement et/ou avec plus d'exactitude que d'autres émotions sur un visage, mais ceci serait lié à une caractéristique faciale unique et saillante : le sourire. En effet, comme il n'y a pas d'ambiguïté possible avec d'autres expressions émotionnelles, le sourire faciliterait la reconnaissance de l'émotion de joie (Calvo & Nummenmaa, 2011 ; Calvo & Beltran, 2013). A noter néanmoins que les visages « menaçants » (*i.e.* expressions faciales de peur et de colère qui suscitent un ressenti de menace) seraient traités plus rapidement que les visages « joyeux », de façon automatique, non consciente, mais que la catégorisation de l'expression émotionnelle réalisée dans un second temps serait plus lente pour les émotions de peur et de colère car il y a plus de compétition entre les émotions négatives. Ceci serait dû à de plus fortes similarités perceptives entre les émotions négatives. Par exemple, la bouche fermée et les sourcils froncés peuvent se retrouver lors de l'expression de la colère, du dégoût, voire de la tristesse (Calvo & Beltran, 2013).

Il faut retenir que la diversité des résultats des études analysant les variations de réponses émotionnelles entre visages et scènes émotionnelles peut s'expliquer par des différences de méthodologies et de matériel. En effet, il est parfois difficile de synthétiser un ensemble de résultats qui paraissent discordants. Il est alors important de considérer que les particularités de méthodologie de chaque étude en termes de matériel ou de présentation des stimuli, selon la tâche demandée au participant (jugement explicite d'une émotion, visualisation passive, tâche

de catégorisation sémantique, ...) mais également selon le type de données analysées (réponses comportementales, réactions somatiques, activité cérébrale...) conduisent à évaluer différents niveaux de traitement de l'information sous tendus par des processus cognitifs distincts. Par exemple si le jugement explicite d'un contenu émotionnel est demandé, différents niveaux de traitement sont sollicités selon qu'il s'agit de juger du contenu affectif (i.e. plaisant ou déplaisant) ou émotionnel (i.e. en termes d'émotion de base : joie, peur, colère, dégoût ou tristesse) (Aguado *et al.*, 2018). La confrontation de ces résultats est donc riche d'informations et nécessaire à la compréhension générale du déroulement des traitements en jeu dans une activité cognitive. Cependant il est important de savoir précisément ce que l'on souhaite étudier et de rester prudent dans l'interprétation des résultats. Il est ainsi actuellement difficile de conclure précisément sur les similarités et particularités des visages et scènes émotionnels. Néanmoins, quel que soit le niveau du stimulus émotionnel, celui-ci présente une attractivité attentionnelle plus forte que celle d'un stimulus neutre.

2.2 Attractivité des scènes émotionnelles

Les stimuli émotionnels, présents dans les scènes visuelles, attirent l'attention exogène de façon supérieure aux stimuli neutres quelle que soit la tâche expérimentale (Bradley, Codispoti, Cuthbert, & Lang, 2001; pour une revue, Carretié, 2014 ; Pilarczyk & Kuniecki, 2014). Les scènes visuelles émotionnelles sont donc traitées plus rapidement que les scènes neutres (pour une revue, Vuilleumier, 2015) et ce de façon irrésistible. Par exemple, dans une tâche où l'on présente simultanément une scène émotionnelle et une scène neutre avec comme consigne d'ignorer la scène émotionnelle, la première saccade oculaire se fera malgré tout vers la scène émotionnelle (Nummenmaa, Hyönä & Calvo, 2006). De manière analogue, lors de l'exploration libre d'une scène visuelle, la première saccade et les fixations sont principalement dirigées vers le contenu émotionnel, quelle que soit sa valence émotionnelle, positive ou négative (Calvo & Lang 2005 ; Carniglia, Caputi, Manfredi, Zambarbieri, & Pessa, 2012 ; Kissler & Keil, 2008 ; pour une revue, Mulckhuyse, 2018 ; Nummenmaa *et al.*, 2006). Les saccades oculaires sont des mouvements des yeux brefs réalisés entre deux fixations (c'est-à-dire positions stables). Nous réalisons une moyenne de 3 à 4 saccades par seconde. Ces saccades permettent de placer l'objet cible en vision fovéale, beaucoup plus précise que la vision périphérique qui présente une acuité médiocre et autorise une vision globale sans détails. Ainsi, lors de l'exploration d'une scène nouvelle, les saccades permettent de placer en vision fovéale plusieurs points nécessitant une analyse plus fine pour la compréhension de la scène. Lorsque

les saccades s'orientent préférentiellement dans une zone de l'image pour permettre une fixation sur un objet, elles sont le signe d'un intérêt pour cet élément. Les saccades peuvent être déclenchées de manière exogène en fonction des propriétés perceptives d'un stimulus (Tatler, Hayhoe, Land, & Ballard, 2011) ou de manière endogène (contrôle *top down*) selon certaines propriétés sémantiques (effet de mémoire ou d'attente particulière liée à l'objet recherché) et émotionnelles (Pilarczyk & Kuniecki, 2014) des stimuli. La saillance affective serait plus prédictive de l'exploration visuelle que la saillance perceptive (Niu, Todd, & Anderson, 2012 ; Schomaker, Walper, Wittmann, & Einhäuser, 2017). Pour d'autres auteurs cependant, il n'y aurait pas de « *pop-up effect* », c'est-à-dire que le contenu émotionnel se serait pas saillant au point d'attirer la première saccade ; mais, une fois l'élément émotionnel fixé, le temps de fixation est plus long que pour les éléments non émotionnels (Acunzo & Hendersen, 2011), un résultat qui soulignerait un attrait particulier pour cet élément et la difficulté à désengager son attention d'un stimulus émotionnel. Le traitement de l'information émotionnelle nécessiterait alors d'allouer des ressources attentionnelles et ne serait pas automatique (Pessoa, Kastner, & Ungerleider, 2002). Néanmoins, au niveau électrophysiologique, il a été inféré une analyse automatique des stimuli émotionnels (Müller, Andersen, & Keil, 2007), très rapide, les premiers potentiels évoqués visuels (P1 et C1) apparaissant dès 100 ms post-stimulus (pour une revue, Olofsson, Nordin, Sequeira, & Polich, 2008). De nombreux auteurs défendent ainsi la conception d'un traitement automatique, non conscient des stimuli émotionnels et lié à leur importance biologique et à leur valeur adaptative (Öhman, Flykt, & Esteves, 2001 ; Vuilleumier, Armony, Driver, & Dolan, 2001). Pour d'autres, la capture attentionnelle liée aux stimuli émotionnels est discutée en termes de modulation par la pertinence avec la tâche en cours ou les besoins du sujet (Mazzietti, Sellem, & Koenig, 2014), par la nature du stimulus et/ou par la valence et le niveau d'éveil de celui-ci (par ex. pour une revue, Okon-Singer, Lichtenstein-Vidne, & Cohen, 2013). Enfin, les contenus émotionnels sont également mieux mémorisés que les contenus non émotionnels (Dolcos, LaBar, & Cabeza, 2004 ; Kensinger & Corkin, 2003), ce qui présente une valeur adaptative certaine puisque cela permet une réaction plus rapide lors d'un nouveau contact avec ce même élément émotionnel.

Nous avons donc vu que la présence d'un contenu émotionnel a une influence sur l'attention exogène et le désengagement de l'attention. Cependant l'attractivité des stimuli émotionnels est modulée par des caractéristiques propres à l'individu. De nombreuses études ont notamment mis en évidence des différences dans le traitement des informations émotionnelles en fonction du sexe (pour une revue voir Kret & de Gelder, 2012). Les stimuli négatifs (notamment les images d'armes ou les scènes d'agressions) suscitent une orientation

attentionnelle plus forte chez les hommes que chez les femmes (Fernández-Martín, Gutiérrez-García, Capafons, & Calvo, 2017). Celles-ci cherchent plus à éviter les stimuli négatifs et jugent plus négativement que les hommes les stimuli déplaisants sur une échelle de valence émotionnelle (Bradley, Codispoti, Sabatinelli, & Lang, 2001). Pour les stimuli à valence positive, selon les bases de données, les scènes érotiques suscitent une attraction plus forte chez les hommes alors que les images de nourrissons ou de familles présentent un attrait plus fort pour les femmes. L'histoire personnelle de l'individu ou la présence de psychopathologies (telles que dépression, phobie ou autre) peuvent également renforcer ou diminuer l'importance de certains stimuli émotionnels (Bishop, 2007; Brosch, Pourtois, Sander, & Vuilleumier, 2011 ; Burkhouse, Owens, Feurer, Sosoo, Kudinova, & Gibb, 2017 ; Duque & Vázquez, 2015; Mueller, Hardin, Mogg, Benson, Bradley, Reinholdt-Dunne, & Ernst, 2012 ; Öhman et al., 2001 ; Pourtois, Schettino, & Vuilleumier, 2012). Enfin il est également à prendre en considération que le même stimulus peut ne pas exercer le même attrait en fonction des besoins de l'individu à un moment donné (Mazzietti et al., 2014). Par exemple un gâteau paraîtra très appétissant et sera très attrayant à l'heure du repas mais ce même gâteau pourrait provoquer une sensation de dégoût après un repas trop important. Cette influence des attentes de l'individu (attention endogène) sur le niveau d'attractivité exogène des stimuli a suscité l'émergence de la conception d'une « attention émotionnelle » relative à l'influence attentionnelle des stimuli émotionnels. L'attention émotionnelle dépendrait de circuits cérébraux spécialisés, en lien avec l'amygdale (Vuilleumier, 2005) et dissociés des réseaux fronto-pariétaux sous tendant l'allocation de l'attention visuelle endogène et exogène (Corbetta & Schulman, 2002). L'observation de dissociations électrophysiologiques lors de la mise en jeu de ces trois types d'attention - endogène (liée aux motivations et attentes du sujet), exogène (liée aux caractéristiques perceptives du stimulus) et émotionnelle (liée au contenu émotionnel) - permet en effet d'envisager l'existence de mécanismes neuronaux différents pour chacun (Brosch *et al.*, 2011).

2.3 Rapidité de traitement des scènes émotionnelles

Il est primordial de pouvoir extraire très rapidement puis analyser précisément une information émotionnelle afin d'adapter au mieux nos comportements, notamment d'approche ou d'évitement (Barrett & Bliss-Moreau, 2009). Cette importance est soulignée par le fait que la présence d'une scène émotionnelle (non pertinente à la tâche) perturbe la réalisation de la tâche en cours même lorsque la scène émotionnelle est hors du focus attentionnel (Calvo,

Nummenmaa, & Hyona, 2007 ; Calvo, Gutiérrez-García, & del Líbano, 2015) ou en périphérie du champ visuel (Calvo, Nummenmaa, & Hyona, 2008 ; Calvo *et al.*, 2015 ; D'Hondt, Lassonde, Collignon, Lepore, Honoré, & Sequeira, 2013 ; Rigoulot, D'Hondt, Honoré, & Sequeira, 2012). Le contenu émotionnel des scènes visuelles est ainsi traité de façon automatique, il guide notre regard. Les scènes émotionnelles sont aussi bien perçues en vision périphérique qu'en vision fovéale car l'information émotionnelle serait véhiculée par les basses fréquences spatiales (information approximative mais véhiculée ainsi de façon rapide) dans la voie magnocellulaire (Calvo *et al.*, 2008 ; Phelps, Ling, & Carrasco, 2006). Les informations approximatives ainsi obtenues orienteraient notre regard vers le contenu émotionnel pour permettre une exploration plus fine et précise du stimulus émotionnel (Nummenmaa, Hyönä, & Calvo, 2009). Les émotions renforceraient le traitement visuel rapide grossier (« *coarse* ») aux dépens d'un traitement plus précis (« *fine* ») mais moins rapide (Bocanegra & Zeelenberg, 2009 ; 2011). La détection de stimuli dangereux serait ainsi permise même sur des images floues, grossières, grâce à ce traitement priorisé des basses fréquences spatiales (Mermillod, Droit-Volet, Devaux, Schaefer, & Vermeulen, 2010) : ceci permet à l'observateur d'agir rapidement. D'autres auteurs ont précisé un rôle différencié des hautes et basses fréquences spatiales en fonction de la tâche demandée. Ainsi les basses fréquences spatiales seraient priorisées dans les comportements d'approche ou d'évitement alors que les hautes fréquences spatiales, qui permettent un traitement plus précis et détaillé de l'information, seraient priorisées dans les tâches de jugement du contenu émotionnel de scènes visuelles (Fradcourt, Peyrin, Baciú, & Campagne, 2013). Il est cependant à noter à nouveau que la majorité des études des liens entre traitement de l'information émotionnelle et fréquences spatiales utilisent des visages comme stimuli (pour une revue voir De Cesarei & Codispoti, 2013). Nous présenterons ainsi une étude qui utilise des visages comme matériel expérimental, et non des scènes visuelles, mais qui permet néanmoins d'étayer les rôles différenciés des fréquences spatiales hautes et basses lors du traitement du contenu émotionnel. Pour tester l'effet de prédiction générée par le traitement des informations globales, Beffara et collaborateurs (2015) ont utilisé une tâche d'amorçage de type « stroop émotionnel » dans laquelle les participants devaient reconnaître une émotion sur un visage présenté accompagné d'un mot cohérent ou incohérent (par exemple un visage de joie accompagné du mot « joie » ou du mot « colère » respectivement). L'étude consistait en la présentation rapide d'une amorce, c'est-à-dire d'un visage émotionnel exprimant la même émotion que celle devant être jugé sur le visage cible, présentée en hautes ou basses fréquences spatiales. L'amorce était suivie par la présentation du visage cible en situation de cohérence ou d'incohérence. L'effet d'interférence-Stroop est réduit dans la

condition d'amorçage en basses fréquences comparée à l'amorçage en hautes fréquences. Ainsi, la présence d'un mot incohérent émotionnellement avec le visage qu'il accompagne entraîne un allongement du temps de reconnaissance de l'expression du visage mais de façon moindre si ce visage est précédé par une amorce en basses fréquences. Ceci conduit les auteurs à proposer que les attentes, qui dépendent de l'extraction précoce d'information à partir des basses fréquences, facilitent l'inhibition de l'information non pertinente lors du traitement de l'expression émotionnelle. Ces résultats confortent le modèle de Bar (2004, voir supra) et l'hypothèse de la prédiction affective de Barrett et Bar (« *the affective prediction hypothesis* », 2009) qui étend le modèle de reconnaissance des objets de Bar (2004) aux objets émotionnels. La rapidité de traitement du contenu émotionnel d'un stimulus serait ainsi permise par le système magnocellulaire qui projette dans le cortex orbito-frontal, lequel est connecté à l'amygdale qui a un rôle clé dans la détection et l'encodage d'informations saillantes au niveau émotionnel (Pessoa & Adolphs, 2010). En effet, l'amygdale présente des connexions réciproques avec de nombreuses régions corticales : ceci lui permet de recevoir de riches informations sensorielles et de renforcer la représentation neuronale des stimuli émotionnels par des *feed-back* grâce aux voies sensorielles, aux régions liées à l'attention (cortex pré-frontal et lobe pariétal) et aux régions liées à la mémoire (hippocampe). De plus les connexions multiples avec les aires préfrontales permettent à l'amygdale un relai dans les régions liées à la prise de décision et à la motivation (Brosch, Scherer, Grandjean, & Sander, 2013 ; Pessoa, 2008).

2.4 Interaction entre les informations émotionnelles contextuelles et locales

Nous avons la capacité de reconnaître et de classifier rapidement et efficacement les informations émotionnelles (Hietanen & Astikainen, 2013 ; pour une revue, Brosch, Pourtois, & Sander, 2010), ce qui nous permet de rapidement évaluer et comprendre l'état émotionnel d'autrui et d'obtenir des informations importantes sur notre environnement. L'étude de la façon dont nous traitons les stimuli émotionnels est principalement fondée sur la théorie des émotions de base (Ekman, 1992) différenciant un nombre limité d'émotions basiques ou discrètes (la joie, la colère, la peur, la tristesse, le dégoût et parfois la surprise) ou selon les approches dimensionnelles (Russell, 2003) qui prennent en compte la valence émotionnelle (selon l'axe positif/image plaisante versus négatif/image déplaisante) et le niveau d'éveil (selon le continuum relaxé/calme versus excité/activé). La cohérence émotionnelle entre le contexte et les objets (*i.e.* lorsque ces deux niveaux sont porteurs de la même catégorie émotionnelle)

facilite la reconnaissance de l'émotion par rapport à une situation d'incohérence émotionnelle (de Borst & de Gelder, 2016 ; de Gelder, Meeren, Righart, Van den Stock, van de Riet, & Tamietto, 2006 ; Kret & de Gelder, 2010 ; Kret, Roelofs, Stekelenburg, & de Gelder, 2013 ; Righart & de Gelder, 2008). Les informations émotionnelles contextuelles (du niveau global) sont ainsi traitées automatiquement et intégrées avec les informations émotionnelles du niveau local et ce, même lorsque l'on demande expressément aux participants d'ignorer le contexte (Righart & de Gelder, 2008). Le contexte émotionnel influence la reconnaissance d'expressions faciales (de Gelder *et al.*, 2006 ; Righart & de Gelder 2008) et d'attitudes corporelles (Kret & de Gelder, 2010) lors d'une présentation simultanée de ces deux niveaux mais également lors d'une présentation séquentielle en amorçage, c'est-à-dire lorsque le contexte est présenté avant l'objet (Hietanen, & Astikainen, 2012). De manière similaire, l'attitude corporelle influence la reconnaissance de l'émotion faciale, et ce que l'on utilise des stimuli statiques (Meeren, van Heijnsbergen, & de Gelder, 2005) ou dynamiques (Nelson & Mondloch, 2017). De plus, les effets contextuels peuvent se combiner. L'expression émotionnelle d'un visage sera en effet mieux reconnue sur un corps présentant une attitude corporelle cohérente émotionnellement (donc de même catégorie émotionnelle) et cette reconnaissance sera encore facilitée si le corps est lui-même intégré dans un contexte émotionnellement cohérent (Reschke, Knothe, Lopez, & Walle, 2018). De façon analogue, les expressions faciales influencent le jugement d'attitudes corporelles (Willis, Palermo, & Burke, 2011). Toutefois, cette influence de l'association émotionnelle entre contexte et objet reste peu comprise. Pour certains auteurs, le rôle du contexte serait modulé par l'âge et par l'attention de l'observateur (processus *top down*). Par exemple si ce dernier est informé que le contexte n'est vraiment pas pertinent pour l'évaluation de l'émotion de l'objet, l'effet contextuel sera présent mais plus faible (Ngo & Isaacowitz, 2015). Pour d'autres chercheurs, la reconnaissance des expressions faciales ne serait pas seulement modulée par le contexte mais dépendante des informations contextuelles (Aviezer *et al.*, 2012 ; Barrett, Mesquita, & Gendron, 2011 ; Wenzler *et al.*, 2016 ; pour des revues : Aviezer, Ensenberg, & Hassin, 2017 ; Wieser & Brosch, 2012). De la même manière que pour les liens sémantiques entre contexte et objet, l'étape de traitement à laquelle se joue l'effet facilitateur de la cohérence émotionnelle sur la reconnaissance de l'émotion reste incertaine. Les informations émotionnelles du contexte et de l'objet peuvent être mises en commun à un niveau décisionnel (Klauer & Musch, 2003) ou à un niveau précoce, perceptif (Spruyt, De Houwer, Hermans, & Eelen, 2007 ; Van den Stock & de Gelder, 2012 ; 2014). Cette dernière hypothèse est soutenue par des études en EEG où l'effet de la valence émotionnelle du contexte sur le jugement de l'émotion d'un visage émotionnel est enregistré précocement (Hietanen &

Astikainen, 2013 ; Righart & de Gelder, 2008). Les recherches futures devront spécifier le rôle et la dynamique des associations entre les informations émotionnelles élaborées à partir du contexte et à partir de l'objet.

Au final, l'effet facilitateur de la cohérence émotionnelle sur le jugement d'une émotion indique donc que le traitement en parallèle et très rapide des niveaux global (contexte) et local (objets) au sein d'une scène est mené non seulement sur les plans perceptif (informations structurelles et spatiales) et sémantique, mais aussi sur celui des informations émotionnelles. Ces dernières, de par leur importance biologique, sont sélectionnées, explorées et analysées prioritairement.

3. Traitements sémantiques et émotionnels : indépendance ou interaction ?

Un débat constant dans la littérature concerne la question de l'indépendance ou de l'interaction entre les traitements des informations sémantiques et émotionnelles. Ainsi, par exemple, sur un même stimulus-visage expressif, le traitement nécessaire à la catégorisation du genre sera-t-il distinct et indépendant de celui de la reconnaissance de l'expression faciale ? Ou bien existe-il une interaction entre ces traitements (*e.g.* une interférence des émotions traitées de manière non volontaire lors de la tâche de catégorisation du genre et une catégorisation non volontaire du genre lors de la tâche de jugement émotionnel) ? La question du décours temporel de ces traitements se pose également. Si ces traitements sont indépendants, l'un se réalise-t-il plus précocement que l'autre, et, s'ils sont interdépendants, sont-ils traités de façon sérielle ou en parallèle ? Cet exemple d'un éventuel lien entre le traitement de l'expression faciale et de la catégorisation de genre est un des exemples le plus fréquemment décrit dans la littérature mais le même questionnement survient avec toutes sortes de stimuli : doit-on reconnaître une araignée selon ses attributs sémantiques avant de pouvoir l'évaluer comme déplaisante ? Peut-on juger une scène visuelle comme plaisante sans avoir accédé à son sens ou à son identité ?

3.1 Le traitement du contenu émotionnel est indépendant de celui des informations sémantiques

En s'appuyant sur diverses études de neuropsychologie, qui ont mis en évidence une double dissociation entre les traitements de l'identité faciale et de l'expression émotionnelle, Bruce et Young (1986) ont proposé un modèle théorique de reconnaissance des visages dans lequel le traitement de l'identité faciale et celui de l'expression émotionnelle du visage seraient réalisés de façon indépendante en parallèle. De façon plus générale, cette question d'indépendance des traitements émotionnels et sémantiques rejoint le fameux débat entre

Zajonc et Lazarus sur la question de l'indépendance entre le traitement des émotions et les traitements cognitifs. Pour Zajonc et son hypothèse de la primauté affective, une réponse émotionnelle peut se produire directement sur la base d'informations sensorielles non transformées donc sans nécessiter une évaluation cognitive délibérée et consciente, de façon rapide et automatique (Zajonc, 1980 ; LeDoux, 1996). Cette hypothèse s'appuie notamment sur le phénomène de simple exposition qui décrit une réaction émotionnelle sans conscience ni catégorisation sémantique de l'objet émotionnel. Ainsi, si vous vous retrouvez face à un danger potentiel vous allez d'abord ressentir la peur puis dans un second temps pouvoir définir sémantiquement quel est l'objet dangereux. Cette hypothèse veut donc que l'évaluation émotionnelle, soit est indépendante de l'évaluation sémantique d'un objet ou évènement, soit se déroule avant le traitement sémantique. L'indépendance des traitements émotionnels et sémantiques serait permise par une indépendance des circuits neuronaux spécialisés pour chaque type de traitement (LeDoux, 1995 ; Vuilleumier, 2005). L'influence de l'amorçage émotionnel (*i.e.* la reconnaissance d'une émotion est facilitée si le stimulus est précédé par la présentation rapide d'un autre véhiculant la même émotion en comparaison à la présentation d'émotions différentes) et celle de la présentation subliminale de stimuli émotionnels, donc non perçus consciemment, vont également dans le sens de l'hypothèse de primauté affective (pour une revue, Tamietto & de Gelder, 2010).

3.2 Le traitement du contenu émotionnel est secondaire au traitement sémantique

Pour Lazarus (1984), à l'inverse de l'hypothèse d'indépendance, l'évaluation cognitive d'une situation ou d'un objet est une précondition nécessaire à l'évaluation de l'émotion. L'évaluation émotionnelle se ferait ainsi après l'évaluation sémantique du stimulus. Autrement dit, vous ressentez de la peur face à un ours parce que vous l'avez reconnu. Nummenmaa, Hyönä et Calvo (2010) ont réalisé une série de sept expériences de catégorisation sémantique et de jugement du contenu émotionnel, dans lesquelles ils ont fait varier les positions et les temps de présentation de scènes visuelles. Dans toutes les conditions expérimentales, l'étude met en évidence des temps de réponse plus courts pour la tâche de catégorisation sémantique que pour la tâche de jugement émotionnel. L'étude est ainsi en faveur d'un modèle sériel du traitement des informations sémantiques et émotionnelles. Pour ces auteurs, lorsque nous nous retrouvons devant un ours, nous accédons dans un premier temps à un traitement global des informations sémantiques (au niveau superordonné de la catégorie sémantique, donc ici « un animal ») puis ensuite à un traitement plus fin (au niveau de base de la catégorie sémantique,

ici « un ours ») et enfin dans un dernier temps, nous accédons à l'analyse du contenu émotionnel. Ainsi, pour de nombreux auteurs, la catégorisation émotionnelle serait effectivement précédée par une catégorisation sémantique (Nummenmaa, Hyönä, & Calvo, 2010 ; Pilarczyk & Kuniecki, 2014 ; Storbeck, Robinson, & McCourt, 2006) et elle nécessiterait une appréhension consciente de la perception visuelle (Lähteenmäki, Hyönä, Koivisto, & Nummenmaa, 2015). L'effet d'amorçage émotionnel serait en fait un effet d'amorçage sémantique (Storbeck *et al.*, 2006).

3.3 Les traitements émotionnels et sémantiques sont interdépendants

Enfin, de nombreuses recherches actuelles sont en faveur d'un traitement interactif entre les informations sémantiques et émotionnelles, qu'il s'agisse d'évaluer le genre et l'expression émotionnelle faciale (Curby, Johnson, & Tyson, 2012 ; Ganel & Goshen-Gottstein, 2004 ; Kaufmann & Schweinberger, 2004 ; Yankouskaya, Rotshtein, & Humphreys, 2014) ou le genre et l'expression émotionnelle corporelle (Van den Stock & de Gelder, 2014). Au niveau neuronal il a été mis en avant l'activation de mêmes ensembles de neurones dans le sulcus temporal supérieur et dans le cortex inféro-temporal antérieur en réponse à l'identification du genre et de l'expression émotionnelle d'un visage (Perrett, Smith, Potter, Mistlin, Head, Milner, & Jeeves, 1984 ; Sugase, Yamane, Ueno, & Kawano, 1999). L'interdépendance des processus émotionnels et cognitifs est défendue par de nombreux auteurs (Brosch *et al.*, 2010 ; 2013 ; Okon-Singer, Stout, Stockbridge, Gamer, Fox, & Shackman, 2017 ; Pessoa, 2008, 2018 ; Storbeck & Clore, 2007). Cette conception d'interdépendance est soutenue par plusieurs types d'observations : les émotions peuvent moduler la cognition et la perception visuelle, et inversement la cognition peut réguler les émotions (Pessoa, 2008 ; Storbeck & Clore, 2007). En outre, au niveau cérébral, il semble que les territoires du « cerveau cognitif » jouent un rôle dans la régulation des émotions et des comportements motivationnels et que les territoires de ce qu'il est convenu d'appeler le « cerveau émotionnel » (depuis notamment LeDoux, 1998) régulent la cognition via l'influence sur certains neurotransmetteurs (Brosch *et al.*, 2013 ; Okon-Singer *et al.*, 2017 ; Pessoa, 2018).

Finalement, pour d'autres auteurs les traitements sémantiques et émotionnels peuvent se réaliser de façon indépendante ou en interaction, ceci selon le type de tâche (Lai, Hagoort & Casasanto, 2012 ; Mueller, Fritsch, Hofmann, & Kuchinke, 2017). Ainsi, en utilisant des stimuli hybrides où sont superposés deux visages de genre différent avec des expressions émotionnelles différentes, Schyns et Oliva (1999) ont montré que les mêmes stimuli entraînaient des réponses

différentes suivant la tâche – de catégorisation de genre ou émotionnelle –, ce qui souligne également une perméabilité entre la cognition et la perception.

Pour résumer, de nombreux arguments sont actuellement en faveur d'une interdépendance entre les traitements émotionnels et cognitifs mais les modalités de cette interaction restent à préciser. La question du déroulement temporel des traitements sémantiques et émotionnels des informations visuelles n'est ainsi pas résolue et doit encore être explorée. Il serait également important que les recherches questionnent ces aspects à l'aide de stimuli plus écologiques et plus complexes que des visages isolés.

Conclusion

Cette revue a présenté les principaux travaux sur l'influence des informations sémantiques et émotionnelles lors de la perception visuelle avec l'objectif principal de contribuer à une approche intégrative du traitement des scènes visuelles complexes.

Premièrement, nous pouvons retenir qu'au travers des travaux récents, un consensus important semble émerger, celui d'un processus de perception visuelle dynamique et interactif (Brosch *et al.*, 2010) : la perception visuelle influence, mais est également influencée par, les connaissances sémantiques et émotionnelles (Brandman & Peelen, 2017 ; O'Callaghan *et al.*, 2017 ; Teufel & Nanay, 2017 ; Trapp & Barr, 2015 ; Truman & Mudrick, 2018). La perception visuelle nécessite une intégration constante entre des informations de bas niveaux provenant du stimulus (*bottom-up*) et des informations de haut niveau (*top-down*) en lien avec les attentes de l'observateur. Les interactions continues entre ces niveaux d'information permettent une compréhension rapide et peu coûteuse de notre environnement (Bar, 2004 ; Barret & Bar, 2009 ; Trapp & Bar, 2015). Ainsi, les informations sémantiques et émotionnelles contenues dans les scènes visuelles, notamment parce qu'elles influencent les processus perceptifs, nous permettent de reconnaître rapidement le monde visuel qui nous entoure, d'y guider nos explorations et actions et d'adapter nos comportements en fonction de la pertinence des stimuli par rapport à nos attentes, pertinence qui varie constamment. La proposition de l'existence d'une « attention émotionnelle » (Vuilleumier, 2005) va ainsi dans ce sens d'une attractivité des stimuli émotionnels modulée par le niveau de pertinence de ces informations pour le sujet. Cette subjectivité dans le traitement d'une même information émotionnelle explique également en partie certains résultats expérimentaux contradictoires.

Les recherches ont aussi permis d'établir que les associations entre les informations locales et globales au sein des scènes visuelles contribuent à la rapidité de notre perception

visuelle notamment en autorisant des prédictions, en limitant l'ambiguïté de la perception (Dyck & Brodeur, 2015 ; Trapp & Bar, 2015) et en guidant l'exploration visuelle (Borji *et al.*, 2013 ; Hwang *et al.*, 2011 ; Koehler *et al.*, 2014 ; Onat *et al.*, 2014 ; Wu *et al.*, 2014). En effet, la répétition des rencontres et des expériences au cours de notre vie nous permet d'inférer des cooccurrences entre objets et contextes et ainsi de catégoriser rapidement et efficacement notre environnement visuel quotidien (Bayet, Pascalis, & Gentaz, 2014 ; Widen, 2013). Nous donnons ainsi sens à ce que nous voyons et à nos ressentis. Ces capacités sont fortement liées aux connaissances stockées en mémoire à long terme sur les régularités des configurations sémantiques, spatiales et émotionnelles qui nous permettent de prédire les objets liés à un contexte, leurs objets associés et d'anticiper les actions possibles (Biederman *et al.*, 1982 ; Eckstein *et al.*, 2006 ; Friedman, 1979 ; Mareschal & Quinn, 2001 ; Spotorno *et al.*, 2015). Les réseaux sémantiques et émotionnels, formés entre les représentations des objets et de leurs contextes, ont donc une importance primordiale pour notre adaptation.

Cette revue a également souligné le recours de plus en plus fréquent à l'approche par neuroimagerie dans les recherches actuelles. Ces travaux ont permis une avancée notable dans la connaissance des soubassements cérébraux à l'œuvre dans les traitements perceptifs, sémantiques et émotionnels. Le recours à ces technologies permet en effet d'éprouver les hypothèses théoriques et d'argumenter leur pertinence. Les données sur le fonctionnement cérébral enrichissent également les réflexions en permettant un accès simultané à l'ensemble des processus à l'œuvre lors de la réalisation d'une activité tout en les confrontant aux données issues des réponses comportementales.

Néanmoins, au-delà de ces avancées, cette revue a également souligné certaines questions de recherche qui restent négligées ou qui font encore débat. Nous retenons notamment la question des étapes de mise en commun des informations locales et globales lors des traitements sémantiques et des traitements émotionnels (à des stades perceptifs ou décisionnels), ainsi que celle de l'interaction entre ces traitements sémantiques et émotionnels. Cette dernière question, comme nous l'avons vu précédemment, a souvent été abordée via l'étude d'un éventuel lien entre la catégorisation du genre et le jugement de l'expression émotionnelle d'un même visage ou dans le contexte de la question plus générale des liens entre émotion et cognition. Peu de recherches se sont intéressées aux éventuelles relations entre catégorisation sémantique et catégorisation émotionnelle notamment de stimuli autre que des visages (avec de rares exceptions, par exemple Nummenmaa, Hyönä, & Calvo, 2010). Ceci est d'autant plus dommageable que, au-delà des implications théoriques fondamentales, cet axe de recherche présente un enjeu dans la compréhension de plusieurs conditions pathologiques. De

nombreuses pathologies sont en effet caractérisées par des particularités dans la prise et le traitement d'informations visuelles (notamment émotionnelles), qu'il s'agisse de psychopathologies comme la phobie (Öhman *et al.*, 2001), la dépression (Burkhouse *et al.*, 2017 ; Duque & Vázquez, 2015) et les troubles anxieux (Mueller *et al.*, 2012) ou de troubles du neuro-développement, en particulier les troubles du spectre autistique et le trouble déficitaire de l'attention avec hyperactivité (Tye, Mercure, Ashwood, Azadi, Asherson, Johnson, & McLoughlin, 2013). Une connaissance plus approfondie des particularités de traitement des informations visuelles pourrait ainsi permettre de progresser dans la compréhension de déficits présents dans ces pathologies et aider à leur diagnostic notamment différentiel. Mais pour cela il est important de comprendre quels traitements cognitifs sont sollicités lorsque l'on demande d'évaluer, en la classifiant, l'émotion véhiculée par un stimulus. En particulier, est-ce que cette catégorisation émotionnelle se réalise indépendamment de, ou secondairement à, une catégorisation sémantique ? L'intégration des émotions dans le champ de la psychologie et des neurosciences cognitives de la perception visuelle, l'utilisation de scènes visuelles complexes plutôt que de stimuli isolés avec la prise en compte des interactions contexte-objet constituent autant de pas vers la prise en compte de ce qu'est notre cognition visuelle au quotidien.

Recommandations pour les études futures

Cette revue a permis de dégager les questions vives de ce domaine de recherche sur les dimensions sémantiques et émotionnelles du traitement des scènes visuelles ; chacune pointe des axes de recherche à initier ou intensifier :

- Utiliser des scènes visuelles en complément de l'utilisation de visages dans les recherches sur le traitement des informations émotionnelles, parce que le traitement des émotions ne peut pas être réduit à celui des expressions faciales.
- Proposer des situations expérimentales plus écologiques, notamment avec différents niveaux de stimuli présentés dans l'ensemble du champ visuel.
- S'appuyer sur les travaux déjà réalisés sur les associations sémantiques contexte-objet afin d'améliorer la compréhension des processus à l'œuvre lors de la mise en commun des différents niveaux d'information (global/local) au sein des scènes visuelles émotionnelles.
- Développer l'analyse des dynamiques temporelles de la mise en relation des dimensions contextuelles et locales des informations émotionnelles notamment à l'aide des potentiels évoqués et/ou de la magnétoencéphalographie.

- Utiliser le même matériel expérimental dans des études évaluant à la fois la catégorisation sémantique et le jugement du contenu émotionnel afin d'améliorer la compréhension des liens entre traitements sémantiques et traitements émotionnels. Nous avons vu que ce paradigme a été appliqué aux visages (tâches de détection de genre et d'expression faciale émotionnelle), il serait important de l'utiliser avec des scènes visuelles complexes.
- Développer l'étude des effets du contenu émotionnel sur la catégorisation sémantique mais également de la catégorisation sémantique sur le contenu émotionnel (par exemple par le biais de la manipulation de la cohérence sémantique) afin d'enrichir les débats sur les liens entre traitements sémantiques et émotionnels.

Références bibliographiques

- Acunzo, D. J., & Henderson, J. M. (2011). No emotional “pop-out” effect in natural scene viewing. *Emotion, 11*(5), 1134-1143.
- Aguado, L., Martínez-García, N., Solís-Olce, A., Dieguez-Risco, T., & Hinojosa, J. A. (2018). Effects of affective and emotional congruency on facial expression processing under different task demands. *Acta Psychologica, 187*, 66-76.
- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*(10), 457-464.
- Aviezer, H., Ensenberg, N., & Hassin, R. R. (2017). The inherently contextualized nature of facial emotion perception. *Current Opinion in Psychology, 17*, 47-54.
- Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*, 1225-1229.
- Bacon-Macé, N., Macé, M. J. M., Fabre-Thorpe, M., & Thorpe, S. J. (2005). The time course of visual processing: Backward masking and natural scene categorisation. *Vision Research, 45*(11), 1459-1469.
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617-629.
- Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*(5), 286-290.
- Barrett, L. F., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1521), 1325-1334.
- Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology, 41*, 167-218.
- Bayet, L., Pascalis, O., & Gentaz, É. (2014). Le développement de la discrimination des expressions faciales émotionnelles chez les nourrissons dans la première année. *L'Année Psychologique/ Topics in Cognitive Psychology, 114*(3), 469-500.
- Beffara, B., Wicker, B., Vermeulen, N., Ouellet, M., Bret, A., Funes Molina, M. J., & Mermillod, M. (2015). Reduction of interference effect by low spatial frequency information priming in an emotional Stroop task. *Journal of Vision, 15*(6): 16, 1–9.
- Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review, 94*(2), 115-147.
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*(2), 143-177.
- Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: an integrative account. *Trends in Cognitive Sciences, 11*(7), 307-316.
- Bocanegra, B. R., & Zeelenberg, R. (2009). Emotion improves and impairs early vision. *Psychological Science, 20*(6), 707-713.
- Bocanegra, B. R., & Zeelenberg, R. (2011). Emotional cues enhance the attentional effects on spatial and temporal resolution. *Psychonomic Bulletin & Review, 18*(6), 1071-1076.
- Borji, A., Sihite, D. N., & Itti, L. (2013). Quantitative analysis of human-model agreement in visual saliency modeling: A comparative study. *IEEE Transactions on Image Processing, 22*(1), 55-69.

- Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral Neuroscience*, *117*(2), 369-380.
- Bradley, M.M., Codispoti, M., Cuthbert, B.N., & Lang, P.J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276-298.
- Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: Sex differences in picture processing. *Emotion*, *1*, 300-319.
- Brandman, T., & Peelen, M. V. (2017). Interaction between scene and object processing revealed by human fMRI and MEG decoding. *Journal of Neuroscience*, *37*(32), 7700-7710.
- Brosch, T., Pourtois, G., Sander, D., & Vuilleumier, P. (2011). Additive effects of emotional, endogenous, and exogenous attention: behavioral and electrophysiological evidence. *Neuropsychologia*, *49*(7), 1779-1787.
- Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition and Emotion*, *24*(3), 377-400.
- Brosch, T., Scherer, K. R., Grandjean, D. M., & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*, *143*, w13786.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305-327.
- Bruner, J. S., & Postman, L. (1949). On the perception of incongruity: A paradigm. *Journal of Personality*, *18*(2), 206-223.
- Bullier, J. (1998). Architecture fonctionnelle du système visuel. In M. Boucart, M. A. Hénaff & C. Belin (Eds.), *Vision : aspects perceptifs et cognitifs* (pp. 11-42). Cahors: SOLAL.
- Bullier, J., Girard, P., & Salin, P. A. (1994). The role of area 17 in the transfer of information to extrastriate visual cortex. In A. Peters & K.S. Rockland (Eds.), *Cerebral Cortex, Vol. 10* (pp. 301-330). New York: Plenum Press.
- Burkhouse, K. L., Owens, M., Feurer, C., Sosoo, E., Kudinova, A., & Gibb, B. E. (2017). Increased neural and pupillary reactivity to emotional faces in adolescents with current and remitted major depressive disorder. *Social Cognitive and Affective Neuroscience*, *12*(5), 783-792.
- Calvo, M. G., Gutiérrez-García, A., & del Líbano, M. (2015). Sensitivity to emotional scene content outside the focus of attention. *Acta Psychologica*, *161*, 36-44.
- Calvo, M. G., & Beltrán, D. (2013). Recognition advantage of happy faces: tracing the neurocognitive processes. *Neuropsychologia*, *51*(11), 2051-2061.
- Calvo, M. G., & Nummenmaa, L. (2011). Time course of discrimination between emotional facial expressions: The role of visual saliency. *Vision Research*, *51*(15), 1751-1759.
- Calvo, M. G., Nummenmaa, L., & Hyönä, J. (2008). Emotional scenes in peripheral vision: Selective orienting and gist processing, but not content identification. *Emotion*, *8*(1), 68-80.
- Calvo, M. G., Nummenmaa, L., & Hyönä, J. (2007). Emotional and neutral scenes in competition: Orienting, efficiency, and identification. *The Quarterly Journal of Experimental Psychology*, *60*(12), 1585-1593.

- Calvo, M. G., & Lang, P. J. (2005). Parafoveal semantic processing of emotional visual scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 31(3), 502-519.
- Carniglia, E., Caputi, M., Manfredi, V., Zambarbieri, D. & Pessa E. (2012). The influence of emotional picture thematic content on exploratory eye movements. *Journal of Eye Movement Research* 5(4), 1-9.
- Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, Affective, & Behavioral Neuroscience*, 14(4), 1228-1258.
- Carretié, L., Kessel, D., Carboni, A., López-Martín, S., Albert, J., Tapia, M., Mercado, F., Capilla, A., & Hinojosa, J. A. (2012). Exogenous attention to facial vs non-facial emotional visual stimuli. *Social Cognitive and Affective Neuroscience*, 8(7), 764-773.
- Coco, M. I., Araujo, S., & Petersson, K. M. (2017). Disentangling stimulus plausibility and contextual congruency: Electro-physiological evidence for differential cognitive dynamics. *Neuropsychologia*, 96, 150-163.
- Corbetta, M., & Shulman, G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Neuroscience*, 3, 201-215.
- Crouzet, S. M., & Cauchoix, M. (2011). When does the visual system need to look back? *Journal of Neuroscience*, 31(24), 8706-8707.
- Curby, K. M., Johnson, K. J., & Tyson, A. (2012). Face to face with emotion: holistic face processing is modulated by emotional state. *Cognition and Emotion*, 26(1), 93-102.
- Davenport, J. (2007). Consistency effects between objects in scenes. *Memory & Cognition*, 35(3), 393-401.
- Davenport, J., & Potter, M. (2004). Scene consistency in object and background perception. *Psychological Science*, 15, 559-564.
- de Borst, A. W., & de Gelder, B. (2016). Clear signals or mixed messages: inter-individual emotion congruency modulates brain activity underlying affective body perception. *Social Cognitive and Affective Neuroscience*, 11(8), 1299-1309.
- De Cesarei, A., & Codispoti, M. (2013). Spatial frequencies and emotional perception. *Reviews in the Neurosciences*, 24(1), 89-104.
- de Gelder, B., Meeren, H. K., Righart, R., Van den Stock, J., van de Riet, W. A., & Tamietto, M. (2006). Beyond the face: Exploring rapid influences of context on face processing. *Progress in Brain Research*, 155, 37-48.
- De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, 52(4), 317-329.
- D'Hondt, F., Lassonde, M., Collignon, O., Lepore, F., Honoré, J., & Sequeira, H. (2013). "Emotions guide us": behavioral and MEG correlates. *Cortex*, 49 (9), 2473-2483.
- Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron*, 42(5), 855-863.
- Duque, A., & Vázquez, C. (2015). Double attention bias for positive and negative emotional faces in clinical depression: Evidence from an eye-tracking study. *Journal of Behavior Therapy and Experimental Psychiatry*, 46, 107-114.

- Dyck, M., & Brodeur, M. B. (2015). ERP evidence for the influence of scene context on the recognition of ambiguous and unambiguous objects. *Neuropsychologia*, *72*, 43-51.
- Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science*, *17*(11), 973-980.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (2013). *Emotion in the human face: Guidelines for research and an integration of findings*. Oxford, England: Pergamon Press.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3-4), 169-200.
- Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience*, *13*(2), 171-180.
- Felleman, D. J., & Van, D. E. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1-47.
- Fernández-Martín, A., Gutiérrez-García, A., Capafons, J., & Calvo, M. G. (2017). Adaptive attunement of selective covert attention to evolutionary-relevant emotional visual scenes. *Consciousness and Cognition*, *51*, 223-235.
- Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*, e229.
- Fize, D., Fabre-Thorpe, M., Richard, G., Doyon, B., & Thorpe, S. J. (2005). Rapid categorization of foveal and extrafoveal natural images: Associated ERPs and effects of lateralization. *Brain and Cognition*, *59*(2), 145-158.
- Fradcourt, B., Peyrin, C., Baciú, M., & Campagne, A. (2013). Behavioral assessment of emotional and motivational appraisal during visual processing of emotional scenes depending on spatial frequencies. *Brain and Cognition*, *83*(1), 104-113.
- Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*(3), 316-355.
- Ganel, T., & Goshen-Gottstein, Y. (2004). Effects of familiarity on the perceptual integrality of the identity and expression of faces: the parallel-route hypothesis revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(3), 583-597.
- Ginsburg, A. P. (1986). Spatial filtering and visual form perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance*, Vol. 2. *Cognitive processes and performance* (pp. 1-41). Oxford, England: John Wiley & Sons.
- 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.
- Greene, M. R., & Fei-Fei, L. (2014). Visual categorization is automatic and obligatory: Evidence from Stroop-like paradigm. *Journal of Vision*, *14*(1), 14-14.
- Greene, M. R., & Oliva, A. (2009). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology*, *58*(2), 137-176.
- Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F., & Weinberger, D. R. (2002). The amygdala response to emotional stimuli: a comparison of faces and scenes. *Neuroimage*, *17*(1), 317-323.
- Harnad, S. (1987). *Categorical perception: The groundwork of cognition*. Cambridge: Cambridge University Press.

- Hietanen, J. K., & Astikainen, P. (2013). N170 response to facial expressions is modulated by the affective congruency between the emotional expression and preceding affective picture. *Biological Psychology*, *92*(2), 114-124.
- Hollingworth, A., & Henderson, J. M. (2003). Testing a conceptual locus. *Memory & Cognition*, *31*(6), 930-940.
- Hollingworth, A., & Henderson, J. M. (1999). Object identification is isolated from scene semantic constraint: Evidence from object type and token discrimination. *Acta Psychologica*, *102*(2-3), 319-343.
- Hwang, A. D., Wang, H. C., & Pomplun, M. (2011). Semantic guidance of eye movements in real-world scenes. *Vision Research*, *51*(10), 1192-1205.
- Joubert, O. R., Fize, D., Rousselet, G. A., & Fabre-Thorpe, M. (2008). Early interference of context congruence on object processing in rapid visual categorization of natural scenes. *Journal of Vision*, *8*(13), 1-18.
- Joubert, O., Rousselet, G., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, *47*(26), 3286-3297.
- Kauffmann, L., Chauvin, A., Guyader, N., & Peyrin, C. (2015). Rapid scene categorization: Role of spatial frequency order, accumulation mode and luminance contrast. *Vision Research*, *107*, 49-57.
- Kauffmann, L., Ramanoël, S., & Peyrin, C. (2014). The neural bases of spatial frequency processing during scene perception. *Frontiers in Integrative Neuroscience*, *8*, 37.
- Kaufmann, J. M., & Schweinberger, S. R. (2004). Expression influences the recognition of familiar faces. *Perception*, *33*(4), 399-408.
- Kensinger, E. A., & Corkin, S. (2003). Effect of negative emotional content on working memory and long-term memory. *Emotion*, *3*(4), 378-393.
- Kissler, J., & Keil, A. (2008). Look–don't look! How emotional pictures affect pro-and anti-saccades. *Experimental Brain Research*, *188*(2), 215-222.
- Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7-49).
- Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers. Koehler, K., Guo, F., Zhang, S., & Eckstein, M. P. (2014). What do saliency models predict? *Journal of Vision*, *14*(3), 14.
- Kret, M. E., Roelofs, K., Stekelenburg, J., & de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience*, *7*, 810.
- Kret, M. E., & De Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia*, *50*(7), 1211-1221.
- Kret, M. E., & de Gelder, B. (2010). Social context influences recognition of bodily expressions. *Experimental Brain Research*, *203*(1), 169-180.
- Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General*, *144*, 339-365.
- Lai, V. T., Hagoort, P., & Casasanto, D. (2012). Affective primacy vs. cognitive primacy: Dissolving the debate. *Frontiers in Psychology*, *3*, 243.

- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: affect, activation, and action. In P.J., Lang, R.F., Simons, & M.F., Balaban, (Eds.), *Attention and Orienting: Sensory and Motivational Processes* (pp 97-135). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lazarus, R. S. (1984). On the primacy of cognition.
- LeDoux, J. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. Simon & Schuster, New York edition.
- LeDoux, J. E. (1995). Emotion: Clues from the brain. *Annual Review of Psychology*, 46(1), 209-235.
- Mack, A., Clarke, J., Erol, M., & Bert, J. (2017). Scene incongruity and attention. *Consciousness and Cognition*, 48, 87-103.
- Malcolm, G. L., Groen, I. I., & Baker, C. I. (2016). Making sense of real-world scenes. *Trends in Cognitive Sciences*, 20(11), 843-856.
- Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences*, 5(10), 443-450.
- Mavratzakis, A., Herbert, C., & Walla, P. (2016). Emotional facial expressions evoke faster orienting responses, but weaker emotional responses at neural and behavioural levels compared to scenes: A simultaneous EEG and facial EMG study. *Neuroimage*, 124, 931-946.
- Mazzietti, A., Sellem, V., & Koenig, O. (2014). From stimulus-driven to appraisal-driven attention: Towards differential effects of goal relevance and goal relatedness on attention? *Cognition and Emotion*, 28(8), 1483-1492.
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, 102(45), 16518-16523.
- Mermillod, M., Droit-Volet, S., Devaux, D., Schaefer, A., & Vermeulen, N. (2010). Are coarse scales sufficient for fast detection of visual threat? *Psychological Science*, 21(10), 1429-1437.
- Milner, D., & Goodale, M. (2006). *The visual brain in action*. Oxford, England, Oxford University Press.
- Mueller, C. J., Fritsch, N., Hofmann, M. J., & Kuchinke, L. (2017). Differences in the dynamics of affective and cognitive processing—an ERP study. *Brain Research*, 1655, 41-47.
- Mueller, S. C., Hardin, M. G., Mogg, K., Benson, V., Bradley, B. P., Reinholdt-Dunne, M. L., & Ernst, M. (2012). The influence of emotional stimuli on attention orienting and inhibitory control in pediatric anxiety. *Journal of Child Psychology and Psychiatry*, 53(8), 856-863.
- Mulckhuyse, M. (2018). The influence of emotional stimuli on the oculomotor system: A review of the literature. *Cognitive, Affective, & Behavioral Neuroscience*, 18(3), 411-425.
- Müller, M. M., Andersen, S. K., & Keil, A. (2007). Time course of competition for visual processing resources between emotional pictures and foreground task. *Cerebral Cortex*, 18(8), 1892-1899.
- Munneke, J., Brentari, V., & Peelen, M. (2013). The influence of scene context on object recognition is independent of attentional focus. *Frontiers in Psychology*, 4, 552.

- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92(3), 289-316.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9(3), 353-383.
- Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Visual Cognition*, 25(4-6), 563-574.
- Ngo, N., & Isaacowitz, D. M. (2015). Use of context in emotion perception: The role of top-down control, cue type, and perceiver's age. *Emotion*, 15(3), 292-302.
- Niu, Y., Todd, R., & Anderson, A. K. (2012). Affective salience can reverse the effects of stimulus-driven salience on eye movements in complex scenes. *Frontiers in Psychology*, 3, 336.
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2010). Semantic categorization precedes affective evaluation of visual scenes. *Journal of Experimental Psychology: General*, 139, 222-246.
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2009). Emotional scene content drives the saccade generation system reflexively. *Journal of Experimental Psychology: Human Perception and Performance*, 35(2), 305-323.
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion*, 6(2), 257-268.
- Nordström, H., & Wiens, S. (2012). Emotional event-related potentials are larger to figures than scenes but are similarly reduced by inattention. *Aquatic Biosystems*, 13(49), 1-10.
- O'Callaghan, C., Kveraga, K., Shine, J. M., Adams Jr, R. B., & Bar, M. (2017). Predictions penetrate perception: Converging insights from brain, behaviour and disorder. *Consciousness and Cognition*, 47, 63-74.
- Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130(3), 466-478.
- Okon-Singer, H., Stout, D. M., Stockbridge, M. D., Gamer, M., Fox, A. S., & Shackman, A. J. (2017). The interplay of emotion and cognition. In A. S., Fox, R. C., Lapate, A. J., Shackman & R. J., Davidson (Eds.). *The nature of emotion: Fundamental questions* (2nd edition). New York: Oxford University Press.
- Okon-Singer, H., Lichtenstein-Vidne, L., & Cohen, N. (2013). Dynamic modulation of emotional processing. *Biological Psychology*, 92(3), 480-491.
- Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, 41(2), 176-210.
- Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, 34(1), 72-107.
- Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biological Psychology*, 77(3), 247-265.
- Onat, S., Açıık, A., Schumann, F., & König, P. (2014). The contributions of image content and behavioral relevancy to overt attention. *PloS One*, 9(4), e93254.
- Palmer, S. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, 3(5), 519-526.

- Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurons responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. *Human Neurobiology*, 3(4), 197-208.
- Pessoa, L. (2018). Emotion and the Interactive Brain: Insights From Comparative Neuroanatomy and Complex Systems. *Emotion Review*, 10(3), 204-216.
- Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience*, 11(11), 773-783.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9(2), 148-158.
- Pessoa, L., Kastner, S., & Ungerleider, L. G. (2002). Attentional control of the processing of neutral and emotional stimuli. *Cognitive Brain Research*, 15(1), 31-45.
- Peyrin, C., Mermillod, M., Chokron, S., & Marendaz, C. (2006). Effect of temporal constraints on hemispheric asymmetries during spatial frequency processing. *Brain and Cognition*, 62(3), 214-220.
- Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, 17(4), 292-299.
- Pilarczyk, J., & Kunięcki, M. (2014). Emotional content of an image attracts attention more than visually salient features in various signal-to-noise ratio conditions. *Journal of Vision*, 14(12), 4-4.
- Pourtois, G., Schettino, A., & Vuilleumier, P. (2012). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology*, 92(3), 492-512.
- Pylyshyn, Z. W. (1980). Computation and cognition: Issues in the foundations of cognitive science. *Behavioral and Brain Sciences*, 3(1), 111-132.
- Rémy, F., Vayssière, N., Pins, D., Boucart, M., & Fabre-Thorpe, M. (2014). Incongruent object/context relationships in visual scenes: Where are they processed in the brain? *Brain and Cognition*, 84, 34-43.
- Reschke, P. J., Knothe, J. M., Lopez, L. D., & Walle, E. A. (2018). Putting “context” in context: The effects of body posture and emotion scene on adult categorizations of disgust facial expressions. *Emotion*, 18(1), 153-158.
- Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience*, 3, 1199-1204.
- Righart, R., & De Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience*, 8(3), 264-272.
- Rigoulot, S., D'Hondt, F., Honoré, J., & Sequeira, H. (2012). Implicit emotional processing in peripheral vision: behavioral and neural evidence. *Neuropsychologia*, 50 (12), 2887-2896.
- Rosch, E. (1978) Principles of categorization. In Rosch E. & Lloyd, B. B. (Eds.), *Cognition and Categorization*, Potomac, MD, Erlbaum.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145-172.

- Sabatinelli, D., Fortune, E. E., Li, Q., Siddiqui, A., Krafft, C., Oliver, W., Beck, S., & Jeffries, J. (2011). Emotional perception: Meta-analyses of face and natural scene processing. *Neuroimage*, *54*, 2524-2533.
- Schomaker, J., Walper, D., Wittmann, B. C., & Einhäuser, W. (2017). Attention in natural scenes: affective-motivational factors guide gaze independently of visual salience. *Vision Research*, *133*, 161-175.
- Schyns PG, Oliva A. (1999). Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*(3), 243-265.
- Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychological Science*, *5*(4), 195-200.
- Simola, J., Le Fevre, K., Torniainen, J., & Baccino, T. (2015). Affective processing in natural scene viewing: Valence and arousal interactions in eye-fixation-related potentials. *Neuroimage*, *106*, 21-33.
- Spotorno, S., Malcolm, G. L., & Tatler, B. W. (2015). Disentangling the effects of spatial inconsistency of targets and distractors when searching in realistic scenes. *Journal of Vision*, *15*(2):12, 1-21.
- Spotorno, S., Malcolm, G. L., & Tatler, B. (2014). How context information and target information guide the eyes from the first epoch of search in real-world scenes. *Journal of Vision*, *14*(2):7, 1-21.
- Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology*, *54*(1), 44-53.
- Storbeck, J., & Clore, G. L. (2007). On the interdependence of cognition and emotion. *Cognition and Emotion*, *21*(6), 1212-1237.
- Storbeck, J., Robinson, M. D., & McCourt, M. E. (2006). Semantic processing precedes affect retrieval: The neurological case for cognitive primacy in visual processing. *Review of General Psychology*, *10*, 41-55.
- Studer, T., & Hübner, R. (2008). The direction of hemispheric asymmetries for object categorization at different levels of abstraction depends on the task. *Brain and Cognition*, *67*(2), 197-211.
- Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, *400*(6747), 869-873.
- Sun, H.-M., Simon-Dack, S., Robert, D., & Wolfgang, A. (2011). Contextual influences on rapid object categorization in natural scenes. *Brain Research*, *1398*, 40-54.
- Tamietto, M., & De Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, *11*(10), 697-709.
- Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, *11*(5), 1-23.
- Teufel, C., & Nanay, B. (2017). How to (and how not to) think about top-down influences on visual perception. *Consciousness and Cognition*, *47*, 17-25.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520-522.
- Trapp, S., & Bar, M. (2015). Prediction, context, and competition in visual recognition. *Annals of the New York Academy of Sciences*, *1339*(1), 190-198.

- Truman, A., & Mudrik, L. (2018). Are incongruent objects harder to identify? The functional significance of the N300 component. *Neuropsychologia*, *117*, 222-232.
- Tye, C., Mercure, E., Ashwood, K. L., Azadi, B., Asherson, P., Johnson, M. H., & McLoughlin, G. (2013). Neurophysiological responses to faces and gaze direction differentiate children with ASD, ADHD and ASD+ ADHD. *Developmental Cognitive Neuroscience*, *5*, 71-85.
- Van den Stock, J., & de Gelder, B. (2014). Face identity matching is influenced by emotions conveyed by face and body. *Frontiers in Human Neuroscience*, *8*(53), 1-8.
- Van den Stock, J., & de Gelder, B. (2012). Emotional information in body and background hampers recognition memory for faces. *Neurobiology of Learning and Memory*, *97*(3), 321-325.
- VanRullen, R., & Koch, C. (2003). Visual selective behavior can be triggered by a feed-forward process. *Journal of Cognitive Neuroscience*, *15*(2), 209-217.
- VanRullen, R., & Thorpe, S. J. (2001a). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception*, *30*(6), 655-668.
- VanRullen, R., & Thorpe, S. J. (2001b). The time course of visual processing: from early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*(4), 454-461.
- Võ, M. L. H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision*, *9*(3), 1-15.
- Vuilleumier, P. (2015). Affective and motivational control of vision. *Current Opinion in Neurology*, *28*(1), 29-35.
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585-594.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, *30*(3), 829-841.
- Wangelin, B. C., Bradley, M. M., Kastner, A., & Lang, P. J. (2012). Affective engagement for facial expressions and emotional scenes: the influence of social anxiety. *Biological Psychology*, *91*(1), 103-110.
- Wenzler, S., Levine, S., van Dick, R., Oertel-Knöchel, V., & Aviezer, H. (2016). Beyond pleasure and pain: Facial expression ambiguity in adults and children during intense situations. *Emotion*, *16*(6), 807-814.
- Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review*, *5*(1), 72-77.
- Wieser, M. J., & Brosch, T. (2012). Faces in context: A review and systematization of contextual influences on affective face processing. *Frontiers in Psychology*, *3*(471), 1-13.
- Willis, M. L., Palermo, R., & Burke, D. (2011). Judging approachability on the face of it: The influence of face and body expressions on the perception of approachability. *Emotion*, *11*(3), 514-523.
- Wolfe, J., Võ, M., Evans, K. K., & Greene, M. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, *15*(2), 77-83.

- Wu, C. C., Wang, H. C., & Pomplun, M. (2014). The roles of scene gist and spatial dependency among objects in the semantic guidance of attention in real-world scenes. *Vision Research, 105*, 10-20.
- Yankouskaya, A., Rotshtein, P., & Humphreys, G. W. (2014). Interactions between identity and emotional expression in face processing across the lifespan: evidence from redundancy gains. *Journal of Aging Research, 2014*, 1-12.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151-175.

Chapitre II. Développement des capacités de catégorisation des informations perceptives, sémantiques et émotionnelles de la naissance à l'adolescence

Ce second chapitre présente une synthèse des connaissances sur le développement précoce des capacités de catégorisation, de la naissance à l'adolescence. Ce chapitre est constitué du résumé en français, puis de la reproduction, d'un article de revue de la littérature soumis pour publication le 25 juin 2019 :

Leroy, A., Spotorno, S., & Faure, S. (submitted). Object categorization from infancy to adolescence: Interplay between perceptual, semantic and emotional information.

II.1 Résumé français de l'article

La capacité de catégoriser est considérée comme une facette clé du fonctionnement cognitif humain : elle permet d'éviter une surcharge cognitive, de communiquer plus facilement et de déduire des connaissances en transposant des attributs d'un spécimen d'une catégorie à un autre spécimen de cette catégorie. La catégorisation peut se faire selon différents contenus (perceptif, sémantique et émotionnel) et à différents niveaux (catégorisation d'un objet isolé à un niveau local ou pour un ensemble d'informations contextuelles à un niveau global). Bien que cette capacité ait fait l'objet d'études approfondies chez les adultes, son développement n'est pas encore bien compris. Cette revue de littérature fournit une synthèse critique de la recherche sur le développement de la catégorisation perceptive, sémantique et émotionnelle des objets chez les nourrissons, les enfants et les adolescents. L'objectif de cet article est de résumer les principales connaissances sur le développement de la catégorisation jusqu'à l'âge adulte, mais également d'enrichir les débats actuels sur l'interaction entre les informations perceptives, sémantiques et émotionnelles. De plus, cette revue met en lumière certains aspects négligés qui méritent d'être approfondis, notamment la quasi absence de recherche sur les capacités de catégorisation et sur les associations entre les traitements local et global de l'information durant les périodes d'âge scolaire et à l'adolescence. L'amélioration de la compréhension des processus de catégorisation visuelle est cependant cruciale car, au-delà des implications théoriques fondamentales, cet axe de recherche représente un enjeu clé pour l'évaluation et la remédiation de certaines conditions pathologiques.

II.2 Article II. Object categorization from infancy to adolescence: Interplay between perceptual, semantic and emotional information

Object categorization from infancy to adolescence: Interplay between
perceptual, semantic and emotional information

Anaïs Leroy ^a

Sara Spotorno ^b

Sylvane Faure ^a

^a Laboratoire d'Anthropologie et de Psychologie Cliniques, Cognitives et Sociales (LAPCOS),
Université Côte d'Azur, LAPCOS, France.

^b School of Psychology, Keele University, United Kingdom.

Corresponding author:

Anaïs Leroy

LAPCOS, MSHS Sud Est

Université Côte d'Azur

Pôle Universitaire Saint Jean d'Angély

24 avenue des Diables Bleus

06357 Nice Cedex 4 (France)

Telephone: +33 (0) 4 89 88 14 60

E-mail: anais.leroy@univ-cotedazur.fr

Declarations of interest: none

Abstract:

The ability to categorize is considered as a key facet of human cognitive functioning: it allows one to avoid cognitive overload, to facilitate communication and to infer knowledge. Categorization may be based on different sources of information (perceptual, semantic and emotional) and occur at different levels (categorization of an isolated object at a local level or of contextual information at a global level). Although this ability has been extensively studied in adults, its development is not well understood. This review provides a synthesis of research on the development of perceptual, semantic and emotional object categorization in infants, children and adolescents. It also aims to contribute to the current debate concerning the interplay between these three sources of information, and to highlight poorly-studies areas, like categorization in school-aged and adolescence periods or the role of the relationships between local and global processing. Improving understanding of the categorization processes and their development has important theoretical implications concerning not only normal functioning but also several pathological conditions.

Keywords:

Visual categorization, infancy, childhood, adolescence, semantics, emotions.

Highlights:

- We review current knowledge on the early development of object categorization.
- Child can made perceptual, semantic and emotional categories from the start of life.
- Semantic and emotional object processing are influenced by surrounding information.
- This review highlights the interplay between perceptual, semantic and emotional information.
- Recent theories integrate the view of interacting levels of processing.

Introduction

From birth, human beings are surrounded by visual stimuli that far exceed their processing capabilities. To represent and give meaning to the world around them, they must recognize and group into categories the elements of their visual environment. Categorizing reflects the ability to group different elements that share some common properties based on perceptual features, such as grouping the sun and a lemon into a “yellow” category, or semantic features, such as grouping a lemon and an apple under a “fruit category” (Rosch, 1978). This is essential to avoiding cognitive overload (Harnad, 1987), to quickly make sense of the physical and social environment and to adapt to it as effectively as possible (Harnad, 2017). Categorizing also allows one to share information more efficiently and to infer that characteristics of one element extend to other elements of the same category. We thus recognize and define under the generic term "cup" different types of cups of different colors, shapes and materials, and thus, when we see a new cup, we can infer that it may contain liquid and be used to drink.

When we categorize an object, we have access to its perceptual characteristics and semantics but also to its emotional content, i.e., to the positive or negative emotional feeling it conveys (Barrett & Bar, 2009). In fact, we have the ability to quickly and effectively recognize and classify (i.e., categorize) emotional information (Hietanen & Astikainen, 2013; for a review, Brosch, Pourtois, & Sander, 2010), which allows us to understand the emotional state of our peers and to obtain important information about our environment. Moreover, emotional information is prioritized compared to "neutral" (without emotional content) information (Carretié, 2014; Pilarczyk & Kuniecki, 2014; Vuilleumier, 2015): for many authors, the presence of emotional content leads to an "emotional bias" in perception (for reviews, see: Mulckhuyse, 2018; Pourtois, Schettino, & Vuilleumier, 2012; Vuilleumier, 2015), which influences visual exploration and reinforces the early processing of emotional stimuli (Phelps, Ling, & Carrasco, 2006). For example, when you see a number of thin black lines nearby, you are likely to interrupt your current task to fixate this object in order to identify it and adjust your behavior depending on whether it is a spider or wires. Visual perception is, therefore, modulated by the presence of emotional content. This enables a faster and accurate processing of the emotional stimuli in order to adapt our behavior effectively, particularly to regulate approach versus avoidance responses (Barrett & Bliss-Moreau, 2009).

Over the past few decades, many studies have been carried out to better understand which kinds (i.e., perceptual, semantic or emotional) and levels (i.e., object/detailed versus contextual/more global) of information are prioritized in processing, and how they interact to enable rapid visual recognition. A large body of research has now shown that, in adults, object

categorization is carried out in a few hundred milliseconds (e.g., Thorpe, Fize, & Marlot, 1996; for a review, see: Fabre-Thorpe, 2011) and automatically at an optimal level of performance (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001). To understand how this ability develops, and how children perceive and represent the visual world around them, it seems necessary to analyze how perceptual, semantic and emotional sources of information are processed and integrated during object categorization in early development.

For this purpose, we review studies exploring object categorization in infants, children and adolescents using images of isolated objects or of complex visual scenes (i.e., objects included in contextual information). The review is structured as follows. First, it examines the development of perceptual and semantic (i.e., conceptual) categorization processes during infancy, childhood and adolescence. Second, it examines the influence of contextual (global) information on object (local) perceptual and semantic categorization. Third, the development of emotional information categorization. The aims are 1) to summarize the main findings on the development of categorization during infancy, childhood and adolescence ; 2) to enrich the current debate on the interplay between perceptual, semantic and emotional sources of information; and 3) to highlight some neglected issues that must be further explored.

The development of object categorization: How perceptual and semantic information interact

The development of categorization skills in infancy

Numerous studies on the processing of visual information in children have focused on the first years of life to study the acquisition and/or development of new cognitive skills during infancy. These studies, mostly using preferential looking or novelty-preference methods, have shown that the ability to categorize occurs very early. Indeed, less than an hour after birth, newborns show a preference for stimuli that schematize a human face over stimuli with the same elements but in a scrambled configuration (Johnson, Dziurawiec, Ellis, & Morton, 1991). Toddlers can group visual stimuli into perceptual categories such as colors from three months of age (Franklin & Davies, 2004), after which categories are refined and gradually take into account perceptual and semantic characteristics together (Poulin-Dubois & Pauen, 2017).

Categorization may be performed at three different levels. People can categorize at a basic (e.g., cat, car, forest, bedroom), superordinate (e.g., animal, vehicle, outdoor environment, indoor environment) or subordinate level (e.g., Siamese cat, jeep, rainforest, children's bedroom). For adults, pioneering studies by Rosch (1978) showed that the basic category of

objects is accessed first. However, this finding has been challenged by several more recent ultra-rapid categorization studies suggesting instead that the superordinate level is accessed first (Macé, Joubert, Nespoulous, & Fabre-Thorpe, 2009; Poncet & Fabre-Thorpe, 2014; Praß, Grimsen, König, & Fahle, 2013; Vanmarcke, Calders, & Wagemans, 2016). Studies of children have also suggested that the superordinate level is available earlier. During development, processing would then progressively include a more local, detailed level of information. For instance, at two months of age, children can differentiate mammals from furniture, but not cats from other mammals, being therefore able to categorize at a superordinate level but not at a basic level (Quinn & Johnson, 2000). At three-four months, they can differentiate the cat category from the dog category (differentiation at a basic level), and finally from seven-eight months of age, they start to be able to differentiate animals breeds and thus to categorize at a subordinate level (Quinn, 2004).

Two major theoretical views are still largely debated to explain the early developmental trajectory of object categorization. On the one hand, according to similarity-based models (Rakison & Poulin-Dubois, 2001; Sloutsky & Fisher, 2004), children might create categories by grouping elements according to perceptual similarities. This requires children to extract statistical regularities from the perceptual shapes of objects. Thus, categories may first rely critically on visual input. Categorization would depend on a domain-general mechanism, i.e., the development of conceptual knowledge about the category would be determined by the maturation of cognitive abilities (like working memory, attention, inhibition) and the accumulation of knowledge (Badger & Shapiro, 2015; Sloutsky, 2010; Sloutsky, Kloos, & Fisher, 2007). On the other hand, for theory-based models (or naïve theory or knowledge-based approaches), categorization is domain-specific and depends on innate constraints (Booth & Waxman, 2002; Gelman, 2003; Keil, Carter Smith, Simons, & Levin, 1998). According to these latter models, children have innate, essentialist knowledge that allows them to elaborate categories independently of the visual input. Children might also infer some knowledge about one exemplar of a category from another even when they do not share perceptual features (category-based induction). Many studies, using change detection tasks or comparing categorization decisions concerning objects that can be grouped at a perceptual or semantic level, have been conducted to test these theories and have found discrepant results (for reviews: Gelman & Meyer, 2011; Owen & Barnes, 2019; Rakison & Yermolayeva, 2010; Sloutsky, 2010; Sutherland & Cimpian, 2017).

To go beyond this dichotomy on the origins of categorization, some authors have proposed a different perspective: perceptual and cognitive processes might be integrated.

Mandler (2004) proposed a dual-representation theory positing that infants from birth can create perceptual and conceptual categories, at the same time, for an object with which they interact. For Poulin-Dubois and Pauen (2017), the development of categorization skills starts from an implicit learning of perceptual categories (bottom-up processes) enriched by top-down processes through experience: this allows one to progressively increase the number of representations in memory. For example, from four months onwards, visual image exploration strategies differ according to the infant's prior knowledge of the image content: for instance, children who are in contact with pets focus their visual exploration more on the informative regions of a pet's image than children who are not (Hurley & Oakes, 2015). Previous experiences and knowledge stored in long-term memory thus influence and enhance semantic categorization abilities. In the same way, Fisher, Godwin and Matlen (2015) proposed the Perceptual and Representational Similarity (PaRS) account, according to which children can represent two types of feature similarity. Perceptual similarity is carried out when comparing features at one point (i.e., I saw a cat and a dog and I compared their features directly), while representational similarity corresponds to the comparison of a current feature with a mental representation elaborated through previous perceptual experiences with an object (directly or through other sources, such as books). Other researchers have shown that the type of categorization changes with task demands and/or time constraints. In Duh and Wang's (2014) study, 15-month-old infants performed a change detection task with eye-movement recording. Changes could be perceptual (e.g., change in the color or the form of the object) or semantic (i.e., conceptual, for example disappearance or modification of an element from the scene as change a parasol in a table). When the stimulus duration was short (500 ms), infants were able to detect perceptual changes but not semantic changes. When the duration was extended to 3000 ms, infants fixated more the object undergoing semantic changes than those with perceptual changes. This study, therefore, also showed that infants since 15 months can extract the global meaning ("gist") from a visual scene when they are given the necessary amount of time (Duh & Wang, 2014).

To sum up, although the debate regarding the processes underlying categorization in infants is still ongoing, numerous studies have shown that, from birth, human beings are able to categorize the elements of their visual world, and that this ability becomes progressively more refined.

The development of categorization skills in preschool children (two-six years of age)

Few studies have examined object categorization processes in school-aged children (Owen & Barnes, 2019), although this period is rich for the debate on the early development of categorization processes. Most studies on the preschool period are in favor of a preferential focus on perceptual features rather than on semantic information during categorization whatever the task involved. During the free viewing of natural scenes, eye scan paths of children between two and six years old are more guided by perceptual factors compared to those of children aged eight to ten years and adults (Helo, Pannasch, Sirri, & Rämä, 2014). Moreover, children aged three to five years show more difficulties in switching from perceptual categories (e.g., grouping an open red umbrella with a red fungus) to semantic categories (e.g., associating an open umbrella with a closed umbrella) than vice-versa (Fisher, 2011). Finally, without specific instructions, three and five years old children mainly categorize according to perceptual similarities (Diesendruck & Peretz, 2013). This initial dominance of perceptual over semantic information may be explained by the attentional capture of perceptual salience (colors, visual contrasts, etc.) but also by an increase in semantic knowledge that is progressively carried out as experiences enrich. Indeed, access to semantic grouping seems to be strongly related to a child's everyday experiences. Two-year-old children conducted semantic categories more easily for everyday objects (e.g., categorizing pets when they have a cat or a dog at home) than for less familiar objects, and tend to categorize previously unknown objects relying on their perceptual features (Mareschal & Quinn, 2001). Badger and Shapiro (2012) studied the development of categorization processes in children of three to nine years of age. They used pictures of insects and trained the children to categorize them at a semantic level and not based on their appearance. In each trial, children were presented with two insects (one corresponding to the same family of insect learnt during the practice and one distractor with more or less similar perceptual features but corresponding to another family) and had to choose the one belonging to the learnt family. This study highlights a shift from a perceptual to a semantic default preference that is not due to improved cognitive control (i.e., the ability to inhibit irrelevant information). Indeed, when the target and the distractor had very similar perceptual appearance, all age children were influenced in the same way. For Badger and Shapiro, this shift towards a predominantly semantic grouping of information is related either to an increase in semantic knowledge or to the maturation of analogical reasoning skills. In contrast, other authors have suggested that the use of categorization criteria (perceptual or semantic) depends on the stimuli (e.g., animals versus furniture) and the instructions (Diesendruck & Peretz, 2013).

The development of categorization skills in school-aged children and adolescents

Object categorization in school-aged children and adolescents has been scarcely studied even though it is pivotal for the development of other cognitive abilities, such as memory and reasoning, and thus for learning at school (O'Sullivan, 1996). Object categorization skills further increase through adolescence, likely building on the further development of general cognitive abilities and brain maturation (Bova et al., 2007). Only two studies (Batty & Taylor, 2002; Vanmarcke & Wagemans, 2015a), to our knowledge, have analyzed visual categorization abilities during school age and adolescence. Batty and Taylor (2002) compared categorization skills in children and adolescents (four age groups: 7-8, 9-10, 11-12, and 13-15 years of age) and adults. They proposed a superordinate categorization task performed during event-related brain potentials (ERP) recording, consisting in the rapid detection (stimulus duration: 80 ms) of an animal in natural visual scenes. Although performance, in terms of both accuracy and response times, increased with age, children, even the youngest, showed already very high accuracy (more than 96% correct responses). Moreover, all age groups made errors on the same images, suggesting that they used the same mechanisms during categorization. ERP recordings showed that the number of the cerebral areas recruited for the task progressively decreased as age increased, and this may be linked to the maturation of cognitive abilities. More recently, Vanmarcke and Wagemans (2015a) proposed three different categorization tasks to adolescents aged 11 to 16 years: shape (perceptual) categorization, detection of the presence of an animal (semantic categorization with yes/no responses), and emotional categorization (asking whether the scene was positive or not). Stimulus duration was very brief (33 ms) for the perceptual and semantic categorization tasks, while it was a little longer (83 ms) for the emotional task to allow enough time to process emotional information (see Vanmarcke & Wagemans, 2015b). Regardless of the categorization task used, there were no differences in performance between adolescents and adults.

Concerning the question of a prioritization of perceptual (bottom-up) or semantic (top-down) information in visual processing during school-aged and adolescence period, some studies have used eye-movement recording and found an advantage of processing perceptually highly salient elements compared to highly semantically informative but less salient ones. Açık et al. (2010) used natural images to compare visual exploration before a categorization task in children aged seven to nine years, in young adults (19-27 years) and in older adults (over 72 years). They found that the children's fixations showed a bottom-up bias, being mostly driven by perceptual factors. They also reported that, as age increased, processing of perceptual information decreased while processing of semantic information increased, suggesting a

progressive shift from bottom-up to top-down processing with experience. Other eye-movement studies using natural visual scenes in laboratory settings (Helo et al., 2014) or realistic stimuli such as paintings from the Vincent Van Gogh museum (Walker, Bucker, Anderson, Schreij, & Theeuwes, 2017) also highlighted a prominent role of bottom-up processing in children in contrast with predominant top-down guidance in adults.

Contextual effects on object categorization: How perception and cognition interact

Objects and context: Processing of local and global image levels

Most studies exploring categorization have used images containing an isolated real object. However, in daily life, we are never confronted with an isolated object, as objects usually co-occur and are always included within some contextual information. A natural (real-world) visual scene is characterized by a context (the background or global level; e.g., a desert or an office) and one or – more often – several objects corresponding to living or nonliving entities (details of the scene or local level; e.g., camels, palm trees, etc., or computers, desks, chairs, etc.). It has also been shown that a visual scene presents specific regularities and properties in terms of its overall structure, concerning physical constraints (overall spatial properties, relative placements of composing elements) and semantic constraints (what objects are expected to occur and co-occur in a given scene, see e.g., Biederman, Mezzanotte, & Rabinowitz, 1982). Categorization of an object is impaired when these scene's regularities are violated (Biederman et al., 1982). The processing of global or local stimuli was initially analyzed using the Navon's (1977) task in which a large letter (global level) is made up of small letters (local level). These global/local levels may be consistent or inconsistent. For example, a large "H" made up of small "h" letters in a consistent condition or of small "s" letters in an inconsistent condition. Subsequent studies have adapted this paradigm using natural stimuli such as animals (Studer & Hübner, 2008) or real-world visual scenes (Peyrin, Mermillod, Chokron, & Marendaz, 2006) and have shown a processing bias in adults in favor of the global level, perceived more quickly than the local level.

In children, several studies using geometric stimuli or hierarchical letters have highlighted a transition from a local level bias towards a bias for the global level that develops during the elementary school period (seven-ten years of age) and that becomes comparable to that of adults (Dukette & Stiles, 2001; Harrison & Stiles, 2009; Poirel, Mellet, Houdé, & Pineau, 2008). This switch from the local predominance to the global predominance has been related to neuroanatomical changes (i.e., a decrease in gray matter in the right parietal occipital areas)

occurring during cerebral maturation (Poirel et al., 2011). For some authors, however, global level processing is carried out earlier in the development, from the age of five, although this ability continues to develop to the end of adolescence (Krakowski et al., 2016; Mondloch, Geldart, Maurer, & de Schonen, 2003). In the same way, it has been shown that, to categorize at the local level, children from seven years of age and adults must inhibit the global level while only children must inhibit the local level to process the global one (Krakowski, Borst, Vidal, Houdé, & Poirel, 2018). The global bias thus develops slowly (Harrison & Stiles, 2009). In addition, an influence of stimulus density on the predominant categorization level (i.e., a local bias in the presence of a few spaced elements and a global bias for more cluttered stimuli) has been found to a greater extent in children than in adults (Dukette & Stiles, 2001; Kimchi, Hadad, Behrmann, & Palmer, 2005; Krakowski et al., 2018). Kimchi et al. (2005) proposed a dual developmental trajectory, as the categorization that is the most taxing for adults (i.e., grouping few large elements and individualizing many large elements) has a longer developmental trajectory than less taxing processing (i.e., grouping many small elements and individualizing sparse large elements).

Objects into context: Interactions between the local and global levels

The studies described above used geometric shapes or letter stimuli. However, processing of local and global information occurs continuously in everyday life with more complex, natural stimuli (e.g., seeing multiple trees or a forest). The semantic relationships between the components of a visual scene (i.e., the objects and the context) influence the ability to process them. Indeed, in adults, it has been shown that categorizing an object is faster and more accurate when the object is included in a context that is semantically consistent (i.e., according to expectations arising from experience, e.g., a computer in an office) compared to when it is included in an inconsistent context (e.g., a computer in a forest) (e.g., Davenport, 2007; Davenport & Potter, 2004; Munneke, Brentari, & Peelen, 2013; Palmer, 1975; Rémy, Vayssière, Pins, Boucart, & Fabre-Thorpe, 2014; Sun, Simon-Dack, Gordon, & Teder, 2011). On the other way round, it has been demonstrated that semantic consistency of the included objects facilitates categorization of the context (e.g., Davenport & Potter, 2004; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Munneke et al., 2013).

Few studies have analyzed the effect of object-context semantic associations on categorization in infants, and even fewer have done so for children and adolescents. However, examining how the semantic consistency effect develops is important to understand how

infants, children and adolescents represent their visual environments and how categorization skills change, especially with respect to the role that increased experience plays in this change. Using eye-movement recording, Bornstein, Mash and Arterberry (2011a, 2011b) analyzed object-context semantic associations in four-month-old children. They found different eye-movement patterns depending on the object-context association involved: the object was fixated more than the context in the case of a consistent association, whereas the opposite was found in the case of an inconsistent association. The authors suggest that infants do not process objects and contexts separately but integrate these different levels to form an unified representation of the visual world. The results of the change detection task presented by Duh and Wang (2014) described above suggest that, at 15 months, infants can detect semantic inconsistency in a real-world visual scene when they are given enough time. This study indeed showed that during rapid visual presentation infants can only detect a perceptual change, while with longer presentation durations they fixate more the regions undergoing a semantic change than those undergoing a perceptual change. This suggests that semantic processing may override perceptual processing in infants. A free viewing study of two-year-old children showed that objects semantically inconsistent with the scene's context were fixated for longer, but only when they were also highly perceptually salient (Helo, van Ommen, Pannasch, Danteny-Dordoigne, & Rämä, 2017). This study suggested, therefore, that children can detect semantic inconsistency in a visual scene insofar as their attention is attracted by perceptual information: only once the object is selected (on a perceptual basis), infants are able to detect the inconsistent association. A recent study has used ERP recordings of two year-old children during the presentation of semantically consistent or inconsistent scene-word pairs, and has showed larger N400 amplitudes when processing inconsistent associations than when processing consistent associations (Helo, Azaiez, & Rämä, 2017). The enhancement of the N400 component had already been described in adults, and interpreted as reflecting detection of semantic inconsistency during linguistic processing in the first studies and, more recently, also during processing of pictorial stimuli (Ganis & Kutas, 2003; for a review, see: Kutas & Federmeier, 2011).

Research on individuals from the age of 11 years old have reported that object categorization is facilitated by a consistent scene's context. Vanmarcke and Wagemans (2015a) used an animal detection task (two alternative forced choice) with fast visual presentations to adolescents aged 11 to 16 years. Animals or inanimate objects were presented in isolation or within a semantically consistent context. Animal stimuli presented in a context were better detected than those presented in an isolated manner. Another study (Thomas, Nardini, &

Mareschal, 2017) presented a Stroop-like paradigm with a combination of visual and auditory stimuli to children (six-seven and eight-nine years of age) and adults: stimuli could be presented in a semantically consistent condition (e.g., a picture of a cat with a meowing), a semantically inconsistent condition (e.g., a picture of a cat with a barking), or a neutral condition (e.g., a picture of a cat without sound). The effect of semantic associations changed with age. Compared to the neutral condition, younger children (six-seven years of age) showed a facilitating effect of consistency and no effect of inconsistency, whereas older children, like adults, but to a lesser extent, showed both facilitation due to semantic consistency and interference due to inconsistency. This suggests that processing of semantic associations becomes more robust with age, highlighting the importance of experience in strengthening this skill.

Top-down impacts in visual perception

The reinforcement of the influence of semantic associations on categorization during childhood, which might depend on life experience, raises the issue of the impact of top-down (cognitively driven) processing on bottom-up (perceptually driven) processing. More generally, a core question in visual perception concerns the relationships between perceptual and cognitive processing. According to the classical sensory information processing framework, perceptual information is processed independently of semantic knowledge and expectations (Pylyshyn, 1980). Moreover, processing occurs in a sequential way, with perceptual processes followed by higher level, cognitive processes (i.e., perception is encapsulated in cognition). A different approach claims instead that perception is influenced by our expectations and mental representations (i.e., there is a top-down effect acting directly on perceptual processes,). Much empirical work on adults supports this thesis of cognitive penetrability of perceptual experience (Brandman & Peelen, 2017; O'Callaghan, Kveraga, Shine, Adams, & Bar, 2017; Teufel & Nanay, 2017).

Numerous studies regarding childhood are in favor of this modulation of visual perception by top-down influences. As we have described in the previous sections, children's experiences, including repeated encounters with certain objects in their daily life, affect the way the visual world is explored and categorized. Several infant studies have also highlighted a preference for, and a more precise categorization of, female faces (Ramsey, Langlois, & Marti, 2005) that may be explained by greater contact with women than with men on a daily basis (for a review: Quinn, Lee, & Pascalis, 2019). Infants also more accurately categorize people of the

same sex or belonging to the same ethnicity (or, as some studies conceptualize it, “race”) as their caregivers (Rennels & Kayl, 2017). This impact of top-down processes on visual perception is also found during the school-age period. In addition, the influence of culture on the global or local processing of visual scenes has been shown (for a review, see: Nisbett, Peng, Choi, & Norenzayan, 2001) until the school-age period (Imada, Carlson, & Itakura, 2013; Köster, Itakura, Yovsi, & Kärtner, 2018). Therefore, top-down processing strongly influences perception, in various ways. A well-known study by Bruner and Goodman (1947) highlighted the importance of the symbolic value of an object for the perception of its size. In this study, ten-year-old children were required to adjust the opening of a slot to fit coins or cardboard discs of the same size into a box. The children with coins overestimated their size by up to 30%, unlike the children with the cardboard discs. Moreover, the children of low socioeconomic status overestimated the size of the coins more than the children of higher socioeconomic status.

Taken together, the studies reviewed in this section show that experience influences perception, sometimes generating perceptive bias but mainly improving categorization performance, allowing for faster response times and access to more precise categories.

Emotional categorization: Is emotional information special?

Development of emotional categorization

Categorization also occurs with emotional information, mainly by grouping emotional stimuli according to basic emotions (joy, anger, fear, sadness, disgust, and surprise; Ekman, 1992) or with respect to the continuum positive/ pleasant versus negative/unpleasant images (i.e., emotional valence) and to the continuum relaxed/calm versus excited/activated (i.e., arousal) (e.g., the dimensional approach developed by Russell, 2003).

From seven months of age, toddlers differentiate facial expressions (for review: Bayet, Pascalis, & Gentaz, 2014), first according to the valence and then gradually (throughout childhood and adolescence) also according to basic emotions (Widen, 2013). The earliest emotion categorized is joy (probably related to the unique perceptual pattern of the face for joy: smiling), followed by sadness, anger and, finally, fear. The ordering of categorization between sadness and anger can be reversed depending on the study (for a review, see: Widen & Russell, 2008). Negative emotions are often confused with each other, including sadness and fear, whose categorization benefits the most from contextual information (Theurel et al., 2016). According to some studies, disgust can be categorized quite early on after joy, which may be explained by the perceptive specificity of the nose frown (Theurel et al., 2016; Vicari, Reilly, Pasqualetti,

Vizzotto, & Caltagirone, 2000). However other studies have suggested that disgust may be categorized later on, at approximately six-ten years, as surprise (Pochedly, Widen, & Russell, 2012). The improvement of emotional categorization is considerable between five and ten years of age, reaching peak accuracy from between five and seven years of age (Conte, Brenna, Ricciardelli, & Turati, 2018; Vicari et al., 2000), while a decrease in response time occurs from between seven and ten years of age (De Sonnevile et al., 2002). Although at that age emotional categorization processes are not mature and, thus, are still developing, children as young as seven years old are able to evaluate images with emotional content in terms of valence and arousal in a similar way compared to adults (McManis, Bradley, Berg, Cuthbert, & Lang, 2001). Finally, it has also been shown that, even if the ability to categorize emotions does not fully mature until the end of adolescence, from the age of ten, the neural circuits activated by the perception of facial expressions are identical to those of adults. These networks activated different structures depending on emotion categories even during implicit processing of emotions (Lobaugh, Gibson, & Taylor, 2006).

Although most studies about emotional categorization have use expressive faces that might be differentiated at a perceptual level, others have analyzed a more complex type of emotional categorization that requires understanding and/or inferring a feeling from contextual information. Indeed, the emotion-understanding task requires to link an emotional face or the label of an emotion with a story drawn or told. The child must be able to understand the emotional feeling of the protagonist of the story and to choose the correct label or picture depicting this emotion. This ability improves during the preschool period, with faster development for girls (Boyatzis, Chazan, & Ting, 1993). Indeed, three-year-old girls show the same ability as five-year-old boys. For some authors, emotion-understanding skills are linked to the development of language skills (Beck, Kumschick, Eid, & Klann-Delius, 2012; Nook, Sasse, Lambert, McLaughlin, & Somerville, 2017; Vicari et al., 2000), while other authors relate them to the development of non-verbal, high-level cognitive functions, such as abstract reasoning and executive functions (Albanese, De Stasio, Di Chiaccio, Fiorilli, & Pons, 2010; De Stasio, Fiorilli, & Di Chiacchio, 2014; Martins, Osório, Veríssimo, & Martins, 2016; von Salisch, Haenel, & Freund, 2013).

The development of emotional categorization skills, therefore, presents different trajectories depending on the emotion at hand (Conte et al., 2018) and on the required level of categorization (Vicari et al., 2000). Indeed, perceptual categorization of emotions - being able to discriminate or group emotions depending on specific features (e.g., the smile), but not to name them - develops early in infancy (White et al., 2019). Deeper understanding and finer

categorization of emotions might need more time to mature, which is likely related to the time course of high-level cognitive functions and language development (Widen, 2013).

Particularities of emotional information processing

Even though emotional stimuli may be categorized like any kind of object, many studies of adulthood have shown that processing of emotional stimuli is particular compared to that of neutral stimuli.

Emotional stimuli attract attention more than neutral ones (Bradley, Codispoti, Cuthbert, & Lang, 2001; Pilarczyk & Kuniecki, 2014; for a review, see: Carretié, 2014). They are also processed faster (for review, Vuilleumier, 2015) and in an irrepressible way (Nummenmaa, Hyönä, & Calvo, 2006). Similarly, during scene free viewing, the first saccade is mainly directed towards emotional content regardless of its emotional valence, either positive or negative (e.g., Calvo & Lang, 2005; Kissler & Keil, 2008; for review: Mulckhuyse, 2018, but see Acunzo & Henderson, 2011, for discrepant results). During free viewing, emotional content is even more predictive of oculomotor behavior than perceptual salience (Niu, Todd, & Anderson, 2012; Schomaker, Walper, Wittmann, & Einhäuser, 2017).

Many authors defend the view of an automatic processing of emotional stimuli linked to their biological importance and adaptive value (Öhman, Flykt, & Esteves, 2001; Vuilleumier, Armony, Driver, & Dolan, 2001). For others, attentional capture related to emotional stimuli may be modulated by various aspects: task and motivational relevance, stimulus' valence and arousal (Mazzietti, Sellem, & Koenig, 2014; Okon-Singer, Lichtenstein-Vidne, & Cohen, 2013). Finally, it has also been shown that emotional content is better memorized than neutral content (Dolcos, LaBar, & Cabeza, 2004; Kensinger & Corkin, 2003). This has an essential adaptive value, enabling faster response in subsequent experience of the same (emotional) stimuli.

Studies that have analyzed pupillary response (change in diameter) in infants from the age of seven months have shown that facial emotional expression may be detected even from subliminal presentation. This suggests an automatic processing of emotional information (Jessen, Altvater-Mackensen, & Grossmann, 2016). Priority in processing of emotional stimuli compared to neutral stimuli has been shown in detection tasks. For instance, infants aged eight to 14 months detect more easily snakes than frogs, and angry faces than happy ones (LoBue & DeLoache, 2010). This prioritization of emotional information (especially negative) have been found throughout child's development (see Fu & Pérez-Edgar, 2019; LoBue & Rakison, 2013;

Vaish, Grossmann, & Woodward, 2008). Thus, at three years, children, like adults, detect snakes or spiders faster than flowers or mushrooms (LoBue & DeLoache, 2008), and angry or fearful faces faster than happy ones (LoBue, 2009). Lagattuta and Kramer (2017) recorded eye movements of children (four-five years, six-seven years, eight-ten years) and adults when two faces (positive expression versus negative or neutral expression) were presented simultaneously, either in a free viewing task or with instructions to look at the positive face. Regardless of age and task demands, the first fixation was mostly on the negative face, suggesting an automatic processing of negative information.

The effect of emotional interference on cognitive processes in adults is well established (Pessoa, 2009). Although less documented, some studies have shown the same effect during childhood as well. Children aged four to eight years performed an ERP flanker task using neutral, positive or negative faces, with the emotional expression being irrelevant to the task (Rahman & Wiebe, 2019). Accuracy was higher with positive than with neutral or negative faces. Negative faces increased response times and latency of N2, an ERP component that reflects cognitive control processes, particularly during conflict resolution. This interference of negative content has also been found in working memory (Augusti, Torheim, & Melinder, 2014), inhibition (Kramer, Lagattuta, & Sayfan, 2015) and attentional blink tasks (Heim, Benasich, & Keil, 2013), suggesting that emotional content captures attentional resources at the expense of the current task. For other authors, however, emotional content does influence ongoing cognitive processing only when it is relevant to the current task (Schel & Crone, 2013).

Object into context: The emotional consistency effect

Historically, and still predominantly, the stimuli used to assess the influence or recognition of emotions are facial expressions. However, even when seen under optimal conditions, facial expressions may be ambiguous (Aviezer, Trope, & Todorov, 2012; Wenzler, Levine, van Dick, Oertel-Knöchel, & Aviezer, 2016). Moreover, in daily life, we are rarely (or never) confronted with faces that appear in isolation, and we extract emotional content from various, co-occurring levels of information: we need to infer emotions not only from faces, but also from body attitudes, social interactions, nonhuman elements such as animals or objects, and from the scene's context.

Research on adults has established that when two stimuli convey the same emotion (condition of emotional consistency), emotional categorization is facilitated compared to when they convey different emotions (condition of emotional inconsistency) (de Borst & de Gelder,

2016; Kret, Stekelenburg, Roelofs, & de Gelder, 2013; Kret & de Gelder, 2010; Righart & De Gelder, 2008). Taking into account different sources of consistent emotional information from a visual scene improves the categorization of emotions. The emotional context (e.g., a party, a paradisiac beach, a forest in fire, a city after an earthquake, etc.) influences the categorization of facial expressions (Righart & De Gelder, 2008) and body attitudes (Hietanen & Astikainen, 2013; Kret & de Gelder, 2010). Body attitudes influence categorization of facial expressions (Meeren, van Heijnsbergen, & de Gelder, 2005; Nelson & Mondloch, 2017) and reciprocally, facial expressions influence categorization of body attitudes (Willis, Palermo, & Burke, 2011).

How does this emotional consistency effect develop in early life? A recent study has shown that six-month-old infants can integrate emotional information from faces and bodies (Heck, Chroust, White, Jubran, & Bhatt, 2018). Moreover, the integration of facial and body attitude was not only conducted between positive and negative emotions, but also for emotions of the same negative category (e.g., sadness and anger), suggesting that infants can discriminate some close characteristics. Another recent study using ERP has shown that from eight months of age, the neural discrimination of facial expressions is decreased by the prior presentation of a body with an emotional attitude in conflict, suggesting that infants detect the emotional inconsistency between these two elements (Rajhans, Jessen, Missana, & Grossmann, 2016). Later in development, from the age of five years, the accuracy of facial emotion categorization is improved when a face is presented with another expressing the same emotion compared to an isolated presentation (Theurel et al., 2016). Facilitation from the emotional consistency between a face and contextual information or body attitude has been described in six-year-olds (Mondloch, Horner, & Mian, 2013) and, at the age of eight, it appears even higher than in adults (Mondloch, 2012). The impairment of the categorization of a facial emotion presented with a body with an inconsistent emotional attitude seems to be even greater with dynamic stimuli (Nelson & Mondloch, 2017).

In sum, while still insufficiently studied, the facilitating effect of emotional consistency, highlighted in adults, seems to be present early in a child's development, suggesting that children simultaneously process different sources of information to improve their emotional understanding.

Conclusion and future directions

The main aim of this review was first to provide an overview of the current knowledge on the development of categorization (perceptual, semantic and emotional) in infancy,

childhood and adolescence. Very early in their development, children are able to group different elements into the same category and to exclude from this category those elements that do not share the key characteristics (Quinn & Johnson, 2000; Widen, 2013). Studies have thus demonstrated that infants can group into perceptual, semantic and emotional categories from the first few months of life. The categories then become progressively more detailed until the end of adolescence.

Despite the numerous studies conducted to date, it is still difficult to understand the exact nature of these categories, as it is complex to separate processing of perceptual, semantic and emotional contents. Recent theories aiming to explain the development of categorization abilities have increasingly posited interacting dimensions of processing. Several arguments in favor of this integrative hypothesis have emerged from the research described in this review.

First, concerning the interaction of perceptual, semantic and emotional sources of information, several studies using different experimental tasks - categorization, change detection and visual exploration – have highlighted a transition from a predominantly perceptual to a predominantly semantic processing that is modulated by the task instructions, the nature of the stimulus and the child's experience. Additionally, the emotional content appears to be processed automatically and may affect performance even when it is not relevant to the task at hand.

Second, the studies reviewed here have shown that visual object processing is influenced by surrounding information: infants, children and adolescent simultaneously process information from co-occurring object, scene's contextual information or even from another modality (e.g., auditory), and these different sources interact. Infants may thus detect, from a few months of life, the presence of semantic or emotional inconsistency. In adults, research has shown that associations between local (object) and global (context) information within visual scenes contribute to the rapidity of visual categorization by enabling predictions, limiting the ambiguity of perception (Trapp & Bar, 2015) and guiding visual scene inspection (Hwang, Wang, & Pomplun, 2011; Onat, Aık, Schumann, & Knig, 2014; Wu, Wang, & Pomplun, 2014). Indeed, the repetition of encounters and experiences throughout life enables the inference of co-occurrences between objects and contexts, which leads children to become progressively more efficient at categorizing their daily visual environments (Bayet et al., 2014; Widen, 2013). In this way, children give meaning to what they see, and how they feel. These skills are strongly linked to the knowledge stored in long-term memory about regularities of perceptual, semantic and emotional configurations that allow one to predict scene context-related objects and to anticipate possible actions (Biederman et al., 1982; Friedman, 1979;

Mareschal & Quinn, 2001). Semantic and emotional networks formed between representations of objects and their contexts are thus crucial to a child's adaptation to the environment.

Another aim of this review was to contribute to the debate concerning the interplay between perceptual, semantic and emotional information. Overall, the results presented are in agreement with the theory of cognitive penetrability of perceptual experience, according to which cognitive factors directly influence perceptual processing (Brandman & Peelen, 2017; O'Callaghan et al., 2017; Teufel & Nanay, 2017). Indeed, while numerous studies have shown that perceptual information plays a predominant role in visual categorization from birth to adolescence, other work has highlighted that, as a child grows up, experience influences his or her perceptions. In fact, the semantic and emotional consistency effects, which develop with experience, show that top-down processes impact visual perception. In sum, as observed in adults, children's visual perceptions are thus characterized by the constant integration of low-level information from the stimulus (bottom-up processing) and of higher level (semantic and emotional) information (top-down processing). Continuous interactions between these sources of information lead children to gather an efficient understanding of the environment (Bar, 2004; Barrett & Bar, 2009; Trapp & Bar, 2015).

Finally, the last aim of this review was to highlight some poorly studied issues that must be further examined. Most of the research reviewed here was conducted with a focus on infancy, mainly to explore if categorization processes fully operate from birth or evolve with cerebral maturation and life experiences (i.e., innate or acquired ability). Fewer studies have been conducted during the preschool period, and very few have focused on the school-age and adolescence periods. However, exploring object categorization in children and adolescents is crucial to better understand how they perceive their visual world and process the perceptual, semantic and emotional information that surrounds them. Studying object-context associations in visual scenes is useful to understand the development of object categorization, which is essential for the rapid visual recognition that characterizes adulthood and that is so important, from basic activities of daily life to complex social interactions. Future research on emotions will also need to use visual scenes with emotional contexts (e.g., photographs of the same seafront with a tsunami versus with people watching the sunset), in addition to faces, as visual emotional experience cannot be reduced to the processing of facial expressions. It will also be important to use the same experimental material to examine both semantic categorization and emotional content categorization, in order to better understand the links between semantic and emotional contents.

Besides its fundamental theoretical implications, research on object categorization may be highly relevant for understanding some pathological conditions, several of which are characterized by particularities in visual information processing (especially emotional information). For instance, some psychiatric conditions, such as phobia (Öhman et al., 2001), depression (Burkhouse et al., 2017) and anxiety disorder (Mueller et al., 2012), or neurodevelopmental disorders, like autism spectrum disorders (ASD) and attention deficit hyperactivity disorder (ADHD) (Tye et al., 2013). A more in-depth knowledge of how young, healthy people categorize visual information would improve understanding of the deficits present in these pathologies, and would, thus, facilitate their differential diagnosis. For example, weaker abilities to process the contextual semantics of a visual scene have been reported in individuals with ASD compared to a control population (Vanmarcke et al., 2016). It has also been shown that the ocular inspection of emotional faces may help to differentiate individuals with ADHD from those with schizophrenia (Marsh & Williams, 2006) and from those with ASD (Tye et al., 2013).

References

- Açik, A., Sarwary, A., Schultze-Kraft, R., Onat, S., & Konig, P. (2010). Developmental changes in natural viewing behavior: Bottom-up and top-down differences between children, young adults and older adults. *Frontiers in Psychology, 1*, 1–14. <https://doi.org/10.3389/fpsyg.2010.00207>
- Acunzo, D. J., & Henderson, J. M. (2011). No emotional “Pop-out” effect in natural scene viewing. *Emotion, 11*(5), 1134–1143. <https://doi.org/10.1037/a0022586>
- Albanese, O., De Stasio, S., Di Chiacchio, C., Fiorilli, C., & Pons, F. (2010). Emotion comprehension: The impact of nonverbal intelligence. *Journal of Genetic Psychology, 17*, 101–115.
- Augusti, E. M., Torheim, H. K., & Melinder, A. (2014). The effect of emotional facial expressions on children’s working memory: Associations with age and behavior. *Child Neuropsychology, 20*(1), 86–105.
- Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225–1229.
- Badger, J. R., & Shapiro, L. R. (2015). Category structure affects the developmental trajectory of children’s inductive inferences for both natural kinds and artefacts. *Thinking & Reasoning, 21*(2), 206–229.
- Badger, J. R., & Shapiro, L. R. (2012). Evidence of a transition from perceptual to category induction in 3- to 9-year-old children. *Journal of Experimental Child Psychology, 113*(1), 131–146. <https://doi.org/10.1016/J.JECP.2012.03.004>
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617–629. <https://doi.org/10.1038/nrn1476>
- Barrett, L. F., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1521), 1325–1334.
- Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology, 41*, 167–218.
- Batty, M., & Taylor, M. J. (2002). Visual categorization during childhood: an ERP study. *Psychophysiology, 39*(4), 482–490. <https://doi.org/10.1017/S0048577202010764>
- Bayet, L., Pascalis, O., & Gentaz, É. (2014). Le développement de la discrimination des expressions faciales émotionnelles chez les nourrissons dans la première année. *L'Année Psychologique, 114*(03), 469–500. <https://doi.org/10.4074/S0003503314003030>
- Beck, L., Kumschick, I. R., Eid, M., & Klann-Delius, G. (2012). Relationship between language competence and emotional competence in middle childhood. *Emotion, 12*(3), 503–514. <https://doi.org/10.1037/a0026320>
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*(2), 143–177. [https://doi.org/10.1016/0010-0285\(82\)90007-X](https://doi.org/10.1016/0010-0285(82)90007-X)
- Booth, A. E., & Waxman, S. (2002). Object names and object functions serve as cues to categories for infants. *Developmental Psychology, 38*(6), 948–957. <https://doi.org/10.1037/0012-1649.38.6.948>

- Bornstein, M. H., Mash, C., & Arterberry, M. E. (2011a). Perception of object-context relations: eye-movement analyses in infants and adults. *Developmental Psychology*, *47*(2), 364–375. <https://doi.org/10.1037/a0021059>
- Bornstein, M. H., Mash, C., & Arterberry, M. E. (2011b). Young infants' eye movements over "natural" scenes and "experimental" scenes. *Infant Behavior and Development*, *34*(1), 206–210. <https://doi.org/10.1016/J.INFBEH.2010.12.010>
- Bova, S. M., Fazzi, E., Giovenzana, A., Montomoli, C., Signorini, S. G., Zoppello, M., & Lanzi, G. (2007). The development of visual object recognition in school-age children. *Developmental Neuropsychology*, *31*(1), 79–102.
- Boyatzis, C. J., Chazan, E., & Ting, C. Z. (1993). Preschool Children's Decoding of Facial Emotions. *The Journal of Genetic Psychology*, *154*(3), 375–382. <https://doi.org/10.1080/00221325.1993.10532190>
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276–298. <https://doi.org/10.1037/1528-3542.1.3.276>
- Brandman, T., & Peelen, M. V. (2017). Interaction between scene and object processing revealed by human fMRI and MEG decoding. *Journal of Neuroscience*, *37*(32), 7700–7710.
- Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition and Emotion*, *24*(3), 377–400.
- Bruner, J. S., & Goodman, C. C. (1947). Value and need as organizing factors in perception. *The Journal of Abnormal and Social Psychology*, *42*(1), 33–44. <https://doi.org/10.1037/h0058484>
- Burkhouse, K. L., Owens, M., Feurer, C., Sosoo, E., Kudinova, A., & Gibb, B. E. (2017). Increased neural and pupillary reactivity to emotional faces in adolescents with current and remitted major depressive disorder. *Social Cognitive and Affective Neuroscience*, *12*(5), 783–792. <https://doi.org/10.1093/scan/nsw184>
- Calvo, M. G., & Lang, P. J. (2005). Parafoveal Semantic Processing of Emotional Visual Scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 502–519. <https://doi.org/10.1037/0096-1523.31.3.502>
- Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(4), 1228–1258.
- Conte, S., Brenna, V., Ricciardelli, P., & Turati, C. (2018). The nature and emotional valence of a prime influences the processing of emotional faces in adults and children. *International Journal of Behavioral Development*, *42*(6), 554–562. <https://doi.org/10.1177/0165025418761815>
- Davenport, J. L. (2007). Consistency effects between objects in scenes. *Memory & Cognition*, *35*(3), 393–401.
- Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*(8), 559–564. <https://doi.org/10.1111/j.0956-7976.2004.00719.x>
- de Borst, A. W., & de Gelder, B. (2016). Clear signals or mixed messages: inter-individual emotion congruency modulates brain activity underlying affective body perception.

- Social Cognitive and Affective Neuroscience*, 11(8), 1299–1309. <https://doi.org/10.1093/scan/nsw039>
- De Sonneville, L. M. J., Verschoor, C. A., Njokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial Identity and Facial Emotions: Speed, Accuracy, and Processing Strategies in Children and Adults. *Journal of Clinical and Experimental Neuropsychology*, 24(2), 200–213. <https://doi.org/10.1076/jcen.24.2.200.989>
- De Stasio, S., Fiorilli, C., & Di Chiacchio, C. (2014). Effects of verbal ability and fluid intelligence on children's emotion understanding. *International Journal of Psychology*, 49(5), 409–414.
- Diesendruck, G., & Peretz, S. (2013). Domain differences in the weights of perceptual and conceptual information in children's categorization. *Developmental Psychology*, 49(12), 2383–2395.
- Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the Amygdala and the Medial Temporal Lobe Memory System Predicts Better Memory for Emotional Events. *Neuron*, 42(5), 855–863. [https://doi.org/10.1016/S0896-6273\(04\)00289-2](https://doi.org/10.1016/S0896-6273(04)00289-2)
- Duh, S., & Wang, S. (2014). Infants detect changes in everyday scenes: The role of scene gist. *Cognitive Psychology*, 72, 142–161. <https://doi.org/10.1016/J.COGLPSYCH.2014.03.001>
- Dukette, D., & Stiles, J. (2001). The effects of stimulus density on children's analysis of hierarchical patterns. *Developmental Science*, 4(2), 233–251. <https://doi.org/10.1111/1467-7687.00168>
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience*, 13, 171–180. <https://doi.org/10.1162/089892901564234>
- Fabre-Thorpe, M. (2011). The characteristics and limits of rapid visual categorization. *Frontiers in Psychology*, 2, 243. <https://doi.org/10.3389/fpsyg.2011.00243>
- Fisher, A. V. (2011). Processing of perceptual information is more robust than processing of conceptual information in preschool-age children: Evidence from costs of switching. *Cognition*, 119(2), 253–264.
- Fisher, A. V., Godwin, K. E., & Matlen, B. J. (2015). Development of inductive generalization with familiar categories. *Psychonomic Bulletin & Review*, 22(5), 1149–1173.
- Franklin, A., & Davies, I. R. L. (2004). New evidence for infant colour categories. *British Journal of Developmental Psychology*, 22, 349–377.
- Friedman, A. (1979). Framing pictures: the role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, 108(3), 316–355. <https://doi.org/10.1037/0096-3445.108.3.316>
- Fu, X., & Pérez-Edgar, K. (2019). Threat-related attention bias in socioemotional development: A critical review and methodological considerations. *Developmental Review*, 51, 31–57. <https://doi.org/10.1016/J.DR.2018.11.002>
- Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*, 16(2), 123–144. [https://doi.org/10.1016/S0926-6410\(02\)00244-6](https://doi.org/10.1016/S0926-6410(02)00244-6)

- Gelman, S.A. (2003). *The essential child: Origins of essentialism in everyday thought*. (Oxford Uni). Oxford.
- Gelman, Susan A., & Meyer, M. (2011). Child categorization. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(1), 95–105. <https://doi.org/10.1002/wcs.96>
- Harnad, S. (1987). *Categorical perception: The groundwork of cognition* (Cambridge:).
- Harnad, S. (2017). *To cognize is to categorize: Cognition is categorization*. In Handbook of categorization in cognitive science (Elsevier, pp. 21–54).
- Harrison, T. B., & Stiles, J. (2009). Hierarchical forms processing in adults and children. *Journal of Experimental Child Psychology*, 103(2), 222–240. <https://doi.org/10.1016/J.JECP.2008.09.004>
- Heck, A., Chroust, A., White, H., Jubran, R., & Bhatt, R. S. (2018). Development of body emotion perception in infancy: From discrimination to recognition. *Infant Behavior and Development*, 50, 42–51. <https://doi.org/10.1016/J.INFBEH.2017.10.007>
- Heim, S., Benasich, A. A., & Keil, A. (2013). Distraction by emotion in early adolescence: affective facilitation and interference during the attentional blink. *Frontiers in Psychology*, 4, 580. <https://doi.org/10.3389/fpsyg.2013.00580>
- Helo, A., Azaiez, N., & Rämä, P. (2017). Word Processing in Scene Context: An Event-Related Potential Study in Young Children. *Developmental Neuropsychology*, 42(7–8), 482–494. <https://doi.org/10.1080/87565641.2017.1396604>
- Helo, A., Pannasch, S., Sirri, L., & Rämä, P. (2014). The maturation of eye movement behavior: Scene viewing characteristics in children and adults. *Vision Research*, 103, 83–91. <https://doi.org/10.1016/J.VISRES.2014.08.006>
- Helo, A., van Ommen, S., Pannasch, S., Danteny-Dordoigne, L., & Rämä, P. (2017). Influence of semantic consistency and perceptual features on visual attention during scene viewing in toddlers. *Infant Behavior and Development*, 49, 248–266. <https://doi.org/10.1016/J.INFBEH.2017.09.008>
- Hietanen, J. K., & Astikainen, P. (2013). N170 response to facial expressions is modulated by the affective congruency between the emotional expression and preceding affective picture. *Biological Psychology*, 92(2), 114–124.
- Hurley, K. B., & Oakes, L. M. (2015). Experience and distribution of attention: Pet exposure and infants' scanning of animal images. *Journal of Cognition and Development*, 16(1), 11–30.
- Hwang, A. D., Wang, H.-C., & Pomplun, M. (2011). Semantic guidance of eye movements in real-world scenes. *Vision Research*, 51(10), 1192–1205. <https://doi.org/10.1016/J.VISRES.2011.03.010>
- Imada, T., Carlson, S., & Itakura, S. (2013). East-West cultural differences in context-sensitivity are evident in early childhood. *Developmental Science*, 16(2), 198–208. <https://doi.org/10.1111/desc.12016>
- Jessen, S., Altvater-Mackensen, N., & Grossmann, T. (2016). Pupillary responses reveal infants' discrimination of facial emotions independent of conscious perception. *Cognition*, 150, 163–169. <https://doi.org/10.1016/J.COGNITION.2016.02.010>
- Johnson, M., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40, 1–19.

- Joubert, O. R., Rousselet, G., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, 47(26), 3286–3297. <https://doi.org/10.1016/j.visres.2007.09.013>
- Keil, F. C., Carter Smith, W., Simons, D. J., & Levin, D. T. (1998). Two dogmas of conceptual empiricism: implications for hybrid models of the structure of knowledge. *Cognition*, 65(2–3), 103–135. [https://doi.org/10.1016/S0010-0277\(97\)00041-3](https://doi.org/10.1016/S0010-0277(97)00041-3)
- Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, 31(8), 1169–1180. <https://doi.org/10.3758/BF03195800>
- Kimchi, R., Hadad, B., Behrmann, M., & Palmer, S. E. (2005). Microgenesis and Ontogenesis of Perceptual Organization. *Psychological Science*, 16(4), 282–290. <https://doi.org/10.1111/j.0956-7976.2005.01529.x>
- Kissler, J., & Keil, A. (2008). Look–don't look! How emotional pictures affect pro- and anti-saccades. *Experimental Brain Research*, 188(2), 215–222. <https://doi.org/10.1007/s00221-008-1358-0>
- Köster, M., Itakura, S., Yovsi, R., & Kärtner, J. (2018). Visual attention in 5-year-olds from three different cultures. *PloS One*, 13(7), e0200239.
- Krakowski, C.-S., Borst, G., Vidal, J., Houdé, O., & Poirel, N. (2018). Children inhibit global information when the forest is dense and local information when the forest is sparse. *Journal of Experimental Child Psychology*, 173, 155–167. <https://doi.org/10.1016/J.JECP.2018.03.020>
- Krakowski, C.-S., Poirel, N., Vidal, J., Roëll, M., Pineau, A., Borst, G., & Houdé, O. (2016). The forest, the trees, and the leaves: Differences of processing across development. *Developmental Psychology*, 52(8), 1262–1272. <https://doi.org/10.1037/dev0000138>
- Kramer, H. J., Lagattuta, K. H., & Sayfan, L. (2015). Why is happy–sad more difficult? Focal emotional information impairs inhibitory control in children and adults. *Emotion*, 15(1), 61–72. <https://doi.org/10.1037/emo0000023>
- Kret, M. E., Stekelenburg, J. J., Roelofs, K., & de Gelder, B. (2013). Perception of Face and Body Expressions Using Electromyography, Pupillometry and Gaze Measures. *Frontiers in Psychology*, 4, 28. <https://doi.org/10.3389/fpsyg.2013.00028>
- Kret, M. E., & de Gelder, B. (2010). Social context influences recognition of bodily expressions. *Experimental Brain Research*, 203(1), 169–180. <https://doi.org/10.1007/s00221-010-2220-8>
- Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, 62(1), 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>
- Lagattuta, K. H., & Kramer, H. J. (2017). Try to look on the bright side: Children and adults can (sometimes) override their tendency to prioritize negative faces. *Journal of Experimental Psychology: General*, 146(1), 89–101. <https://doi.org/10.1037/xge0000247>
- Lobaugh, N., Gibson, E., & Taylor, M. . J. (2006). Children recruit distinct neural systems for implicit emotional face processing. *NeuroReport*, 17(2), 215–219.
- LoBue, V. (2009). More than just another face in the crowd: superior detection of threatening facial expressions in children and adults. *Developmental Science*, 12(2), 305–313. <https://doi.org/10.1111/j.1467-7687.2008.00767.x>

- LoBue, V., & DeLoache, J. S. (2008). Detecting the Snake in the Grass. *Psychological Science*, *19*(3), 284–289. <https://doi.org/10.1111/j.1467-9280.2008.02081.x>
- LoBue, V., & DeLoache, J. S. (2010). Superior detection of threat-relevant stimuli in infancy. *Developmental Science*, *13*(1), 221–228. <https://doi.org/10.1111/j.1467-7687.2009.00872.x>
- LoBue, V., & Rakison, D. H. (2013). What we fear most: A developmental advantage for threat-relevant stimuli. *Developmental Review*, *33*(4), 285–303. <https://doi.org/10.1016/J.DR.2013.07.005>
- Macé, M. J.-M., Joubert, O. R., Nespoulous, J.-L., & Fabre-Thorpe, M. (2009). The time-course of visual categorizations: you spot the animal faster than the bird. *PloS One*, *4*(6), e5927. <https://doi.org/10.1371/journal.pone.0005927>
- Mandler, J. M. (2004). *The foundations of mind: Origins of conceptual thought*. Oxford: Oxford University Press.
- Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences*, *5*(10), 443–450. [https://doi.org/10.1016/S1364-6613\(00\)01752-6](https://doi.org/10.1016/S1364-6613(00)01752-6)
- Marsh, P., & Williams, L. (2006). ADHD and schizophrenia phenomenology: visual scanpaths to emotional faces as a potential psychophysiological marker? *Neuroscience and Biobehavioral Reviews*, *30*(5), 651–665.
- Martins, E. C., Osório, A., Veríssimo, M., & Martins, C. (2016). Emotion understanding in preschool children: The role of executive functions. *International Journal of Behavioral Development*, *40*(1), 1–10.
- Mazzietti, A., Sellem, V., & Koenig, O. (2014). From stimulus-driven to appraisal-driven attention: Towards differential effects of goal relevance and goal relatedness on attention? *Cognition and Emotion*, *28*(8), 1483–1492. <https://doi.org/10.1080/02699931.2014.884488>
- McManis, M., Bradley, M., Berg, W., Cuthbert, B., & Lang, P. (2001). Emotional reactions in children: Verbal, physiological, and behavioral responses to affective pictures. *Psychophysiology*, *38*(2), 222–231.
- Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, *102*(45), 16518–16523. <https://doi.org/10.1073/pnas.0507650102>
- Mondloch, C. J. (2012). Sad or fearful? The influence of body posture on adults' and children's perception of facial displays of emotion. *Journal of Experimental Child Psychology*, *111*(2), 180–196. <https://doi.org/10.1016/J.JECP.2011.08.003>
- Mondloch, C. J., Geldart, S., Maurer, D., & de Schonen, S. (2003). Developmental changes in the processing of hierarchical shapes continue into adolescence. *Journal of Experimental Child Psychology*, *84*(1), 20–40. [https://doi.org/10.1016/S0022-0965\(02\)00161-3](https://doi.org/10.1016/S0022-0965(02)00161-3)
- Mondloch, C. J., Horner, M., & Mian, J. (2013). Wide eyes and drooping arms: Adult-like congruency effects emerge early in the development of sensitivity to emotional faces and body postures. *Journal of Experimental Child Psychology*, *114*(2), 203–216. <https://doi.org/10.1016/j.jecp.2012.06.003>
- Mueller, S. C., Hardin, M. G., Mogg, K., Benson, V., Bradley, B. P., Reinholdt-Dunne, M. L., ... Ernst, M. (2012). The influence of emotional stimuli on attention orienting and

- inhibitory control in pediatric anxiety. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 53(8), 856–863. <https://doi.org/10.1111/j.1469-7610.2012.02541.x>
- Mulckhuyse, M. (2018). The influence of emotional stimuli on the oculomotor system: A review of the literature. *Cognitive, Affective, & Behavioral Neuroscience*, 18(3), 411–425.
- Munneke, J., Brentari, V., & Peelen, M. V. (2013). The influence of scene context on object recognition is independent of attentional focus. *Frontiers in Psychology*, 4, 552. <https://doi.org/10.3389/fpsyg.2013.00552>
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9(3), 353–383. [https://doi.org/10.1016/0010-0285\(77\)90012-3](https://doi.org/10.1016/0010-0285(77)90012-3)
- Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Visual Cognition*, 25(4–6), 563–574. <https://doi.org/10.1080/13506285.2017.1301615>
- Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychological Review*, 108(2), 291–310.
- Niu, Y., Todd, R., & Anderson, A. K. (2012). Affective salience can reverse the effects of stimulus-driven salience on eye movements in complex scenes. *Frontiers in Psychology*, 3, 336. <https://doi.org/10.3389/fpsyg.2012.00336>
- Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2017). Increasing verbal knowledge mediates development of multidimensional emotion representations. *Nature Human Behaviour*, 1(12), 881–889.
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion*, 6(2), 257–268. <https://doi.org/10.1037/1528-3542.6.2.257>
- O'Callaghan, C., Kveraga, K., Shine, J. M., Adams Jr, R. B., & Bar, M. (2017). Predictions penetrate perception: Converging insights from brain, behaviour and disorder. *Consciousness and Cognition*, 47, 63–74.
- O'Sullivan, J. T. (1996). Children's metamemory about the influence of conceptual relations on recall. *Journal of Experimental Child Psychology*, 62, 1–29.
- Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130(3), 466–478. <https://doi.org/10.1037/0096-3445.130.3.466>
- Okon-Singer, H., Lichtenstein-Vidne, L., & Cohen, N. (2013). Dynamic modulation of emotional processing. *Biological Psychology*, 92(3), 480–491. <https://doi.org/10.1016/J.BIOPSYCHO.2012.05.010>
- Onat, S., Açık, A., Schumann, F., & König, P. (2014). The Contributions of Image Content and Behavioral Relevancy to Overt Attention. *PLoS ONE*, 9(4), e93254. <https://doi.org/10.1371/journal.pone.0093254>
- Owen, K., & Barnes, C. (2019). The development of categorization in early childhood: a review. *Early Child Development and Care*. <https://doi.org/10.1080/03004430.2019.1608193>
- Palmer, S. (1975). *Visual perception and world knowledge: notes on a model of sensory-cognitive interaction*. In *Explorations in Cognition* (DA Norman, pp. 279-307). San Francisco : Freeman.

- Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, 13(4), 160–166. <https://doi.org/10.1016/j.tics.2009.01.006>
- Peyrin, C., Mermillod, M., Chokron, S., & Marendaz, C. (2006). Effect of temporal constraints on hemispheric asymmetries during spatial frequency processing. *Brain and Cognition*, 62(3), 214–220. <https://doi.org/10.1016/J.BANDC.2006.05.005>
- Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, 17(4), 292–299.
- Pilarczyk, J.; Kuniecki, M. (2014). Emotional content of an image attracts attention more than visually salient features in various signal-to-noise ratio conditions. *Journal of Vision*, 14, 1–19. <https://doi.org/10.1167/14.12.4>
- Pochedly, J. T., Widen, S. C., & Russell, J. A. (2012). What emotion does the “facial expression of disgust” express? *Emotion*, 12(6), 1315–1319. <https://doi.org/10.1037/a0027998>
- Poirel, N., Mellet, E., Houdé, O., & Pineau, A. (2008). First came the trees, then the forest: Developmental changes during childhood in the processing of visual local-global patterns according to the meaningfulness of the stimuli. *Developmental Psychology*, 44(1), 245–253. <https://doi.org/10.1037/0012-1649.44.1.245>
- Poirel, N., Simon, G., Cassotti, M., Leroux, G., Perchey, G., Lanoë, C., ... Houdé, O. (2011). The Shift from Local to Global Visual Processing in 6-Year-Old Children Is Associated with Grey Matter Loss. *PLoS ONE*, 6(6), e20879. <https://doi.org/10.1371/journal.pone.0020879>
- Poncet, M., & Fabre- Thorpe, M. (2014). Stimulus duration and diversity do not reverse the advantage for superordinate- level representations: the animal is seen before the bird. *European Journal of Neuroscience*, 39(9), 1508–1516.
- Poulin-Dubois, D., & Pauen, S. (2017). *The Development of Object Categories: What, When, and How?* In *Handbook of Categorization in Cognitive Science* (Elsevier, pp. 653–671).
- Pourtois, G., Schettino, A., & Vuilleumier, P. (2012). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology*, 92(3), 492–512.
- Praß, M., Grimsen, C., König, M., & Fahle, M. (2013). Ultra rapid object categorization: effects of level, animacy and context. *PLoS One*, 8(6), e68051.
- Pylyshyn, Z. W. (1980). Computation and cognition: issues in the foundations of cognitive science. *Behavioral and Brain Sciences*, 3(1), 111-132. <https://doi.org/10.1017/S0140525X00002053>
- Quinn, P. C. (2004). Development of subordinate- level categorization in 3- to 7- month- old infants. *Child Development*, 75(3), 886–899.
- Quinn, P. C., & Johnson, M. H. (2000). Global-before-basic object categorization in connectionist networks and 2-month-old infants. *Infancy*, 1(1), 31–46.
- Quinn, P. C., Lee, K., & Pascalis, O. (2019). Face processing in infancy and beyond: The case of social categories. *Annual Review of Psychology*, 70, 165–189.
- Rahman, A., & Wiebe, S. (2019). Valence matters: An electrophysiological study on how emotions influence cognitive performance in children. *Developmental Psychobiology*, 61, 290–303. <https://doi.org/10.1002/dev.21813>

- Rajhans, P., Jessen, S., Missana, M., & Grossmann, T. (2016). Putting the face in context: Body expressions impact facial emotion processing in human infants. *Developmental Cognitive Neuroscience, 19*, 115–121. <https://doi.org/10.1016/j.dcn.2016.01.004>
- Rakison, D. H., & Poulin-Dubois, D. (2001). Developmental origin of the animate–inanimate distinction. *Psychological Bulletin, 127*(2), 209–228.
- Rakison, D. H., & Yermolayeva, Y. (2010). Infant categorization. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(6), 894–905. <https://doi.org/10.1002/wcs.81>
- Ramsey, J. L., Langlois, J. H., & Marti, N. C. (2005). Infant categorization of faces: Ladies first. *Developmental Review, 25*(2), 212–246.
- Rémy, F., Vayssière, N., Pins, D., Boucart, M., & Fabre-Thorpe, M. (2014). Incongruent object/context relationships in visual scenes: Where are they processed in the brain? *Brain and Cognition, 84*(1), 34–43. <https://doi.org/10.1016/J.BANDC.2013.10.008>
- Rennels, J. L., & Kayl, A. J. (2017). *How experience affects infants' facial categorization*. In *Handbook of categorization in cognitive science* (Elsevier, pp. 637–652).
- Righart, R., & de Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience, 8*(3), 264–272. <https://doi.org/10.3758/CABN.8.3.264>
- Rosch, E. (1978). *Principles of categorization*. In *Cognition and Categorization* (In Rosch E). Potomac, MD, Erlbaum.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145–172.
- Schel, M. A., & Crone, E. A. (2013). Development of response inhibition in the context of relevant versus irrelevant emotions. *Frontiers in psychology, 4*, 383.
- Schomaker, J., Walper, D., Wittmann, B. C., & Einhäuser, W. (2017). Attention in natural scenes: Affective-motivational factors guide gaze independently of visual salience. *Vision Research, 133*, 161–175. <https://doi.org/10.1016/J.VISRES.2017.02.003>
- Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science, 34*, 1244–1286.
- Sloutsky, V. M., & Fisher, A. V. (2004). Induction and Categorization in Young Children: A Similarity-Based Model. *Journal of Experimental Psychology: General, 133*(2), 166–188.
- Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science, 18*, 179–185.
- Studer, T., & Hübner, R. (2008). The direction of hemispheric asymmetries for object categorization at different levels of abstraction depends on the task. *Brain and Cognition, 67*(2), 197–211. <https://doi.org/10.1016/J.BANDC.2008.01.003>
- Sun, H.-M., Simon-Dack, S. L., Gordon, R. D., & Teder, W. (2011). Contextual influences on rapid object categorization in natural scenes. *Brain Research, 1398*, 40–54. <https://doi.org/10.1016/j.brainres.2011.04.029>
- Sutherland, S. L., & Cimpian, A. (2017). Inductive generalization relies on category representations. *Psychonomic Bulletin & Review, 24*(2), 632–636. <https://doi.org/10.3758/s13423-015-0951-z>

- Teufel, C., & Nanay, B. (2017). How to (and how not to) think about top-down influences on visual perception. *Consciousness and Cognition*, 47, 17–25.
- Theurel, A., Witt, A., Malsert, J., Lejeune, F., Fiorentini, C., Barisnikov, K., & Gentaz, E. (2016). The integration of visual context information in facial emotion recognition in 5- to 15-year-olds. *Journal of Experimental Child Psychology*, 150, 252–271. <https://doi.org/10.1016/J.JECP.2016.06.004>
- Thomas, R. L., Nardini, M., & Mareschal, D. (2017). The impact of semantically congruent and incongruent visual information on auditory object recognition across development. *Journal of Experimental Child Psychology*, 162, 72–88. <https://doi.org/10.1016/J.JECP.2017.04.020>
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.
- Trapp, S., & Bar, M. (2015). Prediction, context, and competition in visual recognition. *Annals of the New York Academy of Sciences*, 1339(1), 190–198.
- Tye, C., Mercure, E., Ashwood, K. L., Azadi, B., Asherson, P., Johnson, M. H., ... McLoughlin, G. (2013). Neurophysiological responses to faces and gaze direction differentiate children with ASD, ADHD and ASD + ADHD. *Developmental Cognitive Neuroscience*, 5, 71–85. <https://doi.org/10.1016/j.dcn.2013.01.001>
- Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, 134(3), 383–403. <https://doi.org/10.1037/0033-2909.134.3.383>
- Vanmarcke, S., Calders, F., & Wagemans, J. (2016). The time-course of ultrarapid categorization: the influence of scene congruency and top-down processing. *I-Perception*, 7(5), 2041669516673384.
- Vanmarcke, S., & Wagemans, J. (2015a). The influence of age and gender on ultra-rapid categorization. *Visual Cognition*, 23(7), 894–916. <https://doi.org/10.1080/13506285.2015.1091801>
- Vanmarcke, S., Mullin, C., Van der Hallen, R., Evers, K., Noens, I., Steyaert, J., & Wagemans, J. (2016). In the Eye of the Beholder: Rapid Visual Perception of Real-Life Scenes by Young Adults with and Without ASD. *Journal of Autism and Developmental Disorders*, 46(8), 2635–2652. <https://doi.org/10.1007/s10803-016-2802-9>
- Vanmarcke, S., & Wagemans, J. (2015b). Rapid gist perception of meaningful real-life scenes: Exploring individual and gender differences in multiple categorization tasks. *I-Perception*, 6(1), 19–37. <https://doi.org/10.1068/i0682>
- Vicari, S., Reilly, J. S., Pasqualetti, P., Vizzotto, A., & Caltagirone, C. (2000). Recognition of facial expressions of emotions in school-age children: the intersection of perceptual and semantic categories. *Acta Paediatrica*, 89(7), 836–845. <https://doi.org/10.1111/j.1651-2227.2000.tb00392.x>
- von Salisch, M., Haenel, M., & Freund, P. A. (2013). Emotion understanding and cognitive abilities in young children. *Learning and Individual Differences*, 26, 15–19.
- Vuilleumier, P. (2015). Affective and motivational control of vision. *Current Opinion in Neurology*, 28(1), 29–35.

- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of Attention and Emotion on Face Processing in the Human Brain: An Event-Related fMRI Study. *Neuron*, 30(3), 829–841. [https://doi.org/10.1016/S0896-6273\(01\)00328-2](https://doi.org/10.1016/S0896-6273(01)00328-2)
- Walker, F., Bucker, B., Anderson, N., Schreij, D., & Theeuwes, J. (2017). Looking at paintings in the Vincent Van Gogh Museum: Eye movement patterns of children and adults. *PLoS ONE*, 12(6), e0178912. <https://doi.org/10.1371/journal.pone.0178912>
- Wenzler, S., Levine, S., van Dick, R., Oertel-Knöchel, V., & Aviezer, H. (2016). Beyond pleasure and pain: Facial expression ambiguity in adults and children during intense situations. *Emotion*, 16(6), 807–814. <https://doi.org/10.1037/emo0000185>
- White, H., Chroust, A., Heck, A., Jubran, R., Galati, A., & Bhatt, R. S. (2019). Categorical Perception of Facial Emotions in Infancy. *Infancy*, 24(2), 139–161. <https://doi.org/10.1111/infa.12275>
- Widen, S. C. (2013). Children's Interpretation of Facial Expressions: The Long Path from Valence-Based to Specific Discrete Categories. *Emotion Review*, 5(1), 72–77. <https://doi.org/10.1177/1754073912451492>
- Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development*, 23(2), 291–312. <https://doi.org/10.1016/J.COGDEV.2008.01.002>
- Willis, M. L., Palermo, R., & Burke, D. (2011). Judging approachability on the face of it: The influence of face and body expressions on the perception of approachability. *Emotion*, 11(3), 514–523. <https://doi.org/10.1037/a0022571>
- Wu, C.-C., Wang, H.-C., & Pomplun, M. (2014). The roles of scene gist and spatial dependency among objects in the semantic guidance of attention in real-world scenes. *Vision Research*, 105, 10–20. <https://doi.org/10.1016/J.VISRES.2014.08.019>

Chapitre III. Traitement du contenu émotionnel des scènes visuelles par des enfants présentant un trouble déficitaire de l'attention avec ou sans hyperactivité

Ce troisième, et dernier, chapitre de la partie théorique de ce manuscrit, présente une revue systématique des travaux empiriques qui ont permis d'étudier les capacités de traitement du contenu émotionnel des scènes visuelles auprès d'enfants et d'adolescents qui présentent un Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H). Ce chapitre est constitué du résumé en français puis de la reproduction d'un article de revue systématique de la littérature, en cours de révisions majeures pour resoumission :

Leroy, A., Spotorno, S., & Faure, S. (under review). Emotional Scene Processing in Children and Adolescents with Attention Deficit / Hyperactivity Disorder: A Systematic Review. *European Child and Adolescent Psychiatry*.

III.1 Résumé français de l'article

Des particularités de traitement du contenu émotionnel sont fréquemment notées dans le TDA/H, à la fois à un niveau volontaire de traitement (e.g., tâches de catégorisation émotionnelle) ainsi qu'à un niveau non volontaire (e.g., influence sur la réalisation d'une tâche de la présence d'un contenu émotionnel non pertinent). La plupart des recherches dans ce domaine ont utilisé des visages comme matériel expérimental. Cependant, dans notre quotidien nous ne nous retrouvons pas en présence d'expressions faciales isolées, nous devons analyser le contenu émotionnel porté par l'ensemble d'une scène visuelle pour pouvoir nous adapter le plus efficacement à notre environnement. Pour mieux comprendre l'influence des émotions sur le quotidien des enfants qui présentent un TDA/H, il est ainsi essentiel d'analyser le traitement du contenu émotionnel porté par des stimuli de nature différente. Dans cette revue systématique, qui utilise les lignes directrices PRISMA, nous nous sommes intéressé aux études empiriques sur le traitement de scènes visuelles, avec ou sans contenu émotionnel, chez les enfants et adolescents qui présentent un TDA/H. Cette revue montre que les difficultés de traitement de l'information émotionnelle, décrites dans les études qui utilisent des visages comme matériel expérimental, sont également observées lors de l'utilisation des scènes visuelles. De plus, ces difficultés semblent provenir à la fois d'un impact ascendant plus fort des stimuli émotionnels chez les enfants qui présentent un TDA/H, augmentant ainsi l'intensité de leur expérience émotionnelle par rapport aux enfants au développement typique, et des déficits fondamentaux du trouble, qui diminuent les ressources allouées au traitement global des scènes visuelles. Le traitement de l'information émotionnelle impose donc une charge cognitive importante aux

enfants et aux adolescents atteints d'un TDA/H. Ceci pourrait expliquer leurs difficultés socio-émotionnelles et leur faible engagement cognitif dans les tâches en cours lorsqu'un contenu émotionnel est présent dans leur environnement. Ces difficultés doivent ainsi faire l'objet d'un examen plus approfondi en clinique et être prises en compte dans les projets thérapeutiques.

III.2 Article III. Emotional Scene Processing in Children and Adolescents with Attention Deficit / Hyperactivity Disorder: A Systematic Review

Emotional Scene Processing in Children and Adolescents with Attention Deficit / Hyperactivity Disorder: A Systematic Review

Anaïs Leroy^{a, c}

Sara Spotorno^b

Sylvane Faure^a

^a Université Côte d'Azur, LAPCOS, France.

^b School of Psychology, Keele University, United Kingdom.

^c Centre Référent des Troubles des Apprentissages, Hôpitaux pédiatriques de Nice CHU-LENVAL, France.

Corresponding author:

Anaïs Leroy

LAPCOS, MSHS Sud Est

Université Côte d'Azur

Pôle Universitaire Saint Jean d'Angély

24 avenue des Diables Bleus

06357 Nice Cedex 4 (France)

Telephone: +33 (0) 4 89 88 14 60

E-mail: anais.leroy@univ-cotedazur.fr

Declarations of interest: none

Abstract

Impairments in the processing of emotional information are frequently reported in attention deficit hyperactivity disorder (ADHD) at a conscious, explicit level (i.e., emotion recognition) and at an unconscious, implicit level (i.e., emotional regulation/interference). Most of the research in the field has used faces as experimental material. However, it is essential to understand the emotional content of an entire visual scene to adapt to our environment. In this systematic review using the PRISMA guidelines, we were interested in empirical studies about processing of visual scenes, with or without emotional contents, in children and adolescents with ADHD. This review shows that the difficulties in processing emotional information reported for faces are also observed for visual scenes. In addition, these difficulties may stem from a stronger bottom-up impact of emotional stimuli in ADHD, increasing the emotional experience, and from core deficits of the disorder, decreasing the overall processing of visual scenes. Processing of emotional information imposes a substantial cognitive load to children and adolescents with ADHD. This may explain their socio-emotional difficulties and their low cognitive engagement in the current tasks when an emotional content is present in their environment. These difficulties must be further investigated and taken into account in interventional plans.

Keywords: ADHD; Children; Visual scenes; Emotion recognition; Emotional interference; Contextual information.

Introduction

Attention Deficit Hyperactivity Disorder

Attention deficit hyperactivity disorder (ADHD) is a neurodevelopmental disorder that affects 3% to 7% of school-aged children and persists into adulthood in 60% of the cases. ADHD is characterized by the presence of various associated symptoms, including impaired attention, impulsivity and hyperactivity [1]. Delay aversion and socio-emotional difficulties [2] are also reported, such as behavioural disorders and peer rejection, which are thought to be partly related to poor emotional regulation skills [3-4]. According to a review by Shaw and colleagues, 25% to 45% of children with ADHD present emotional dysregulation, defined as emotional lability and emotional reactions that are disproportionate with respect to social norms [5]. This emotional dysregulation may impair maintenance of goal-oriented and appropriate behaviour [6]. Considering the importance of the social, familial and school/academic repercussions caused by these emotional difficulties, many studies have recently focused on how individuals with ADHD process emotional information.

Various aspects subtend emotional regulation skills: recognition of emotions, emotional reactivity, emotional regulation and empathy traits. A recent meta-analysis has shown that the association between ADHD and recognition of emotions has a moderate effect size [7]. Indeed, while some studies have found a deficit in emotion recognition in children and adolescents with ADHD [8-12], some others have not reported these difficulties [13-14]. The influence of several individual differences (e.g., ADHD subtype, comorbidity) and differences in the methodology of the studies (e.g., task demands, type of emotions examined) may, in part, explain these inconsistencies in findings (for a critical review on facial emotional recognition, see 15).

The stimuli most commonly used in studies examining recognition and influence of emotions are facial expressions. However, in our daily life, we never see isolated faces and we need to extract emotional content from different co-occurring types of information. Moreover, even when seen under optimal conditions, facial expressions may be ambiguous, and the same expression may convey different emotions: for example, an individual may cry for either joy or sadness [16]. Various studies have accordingly emphasized the importance of context for understanding the emotion expressed by an individual [17-19; for reviews, see 20-21]. In particular, we need to infer emotions from body attitudes, social interactions, and contextual information. The use of visual scenes with emotional content, therefore, seems more ecological than the use of isolated faces to study the processing of visual emotional information.

Particularities in the processing of emotional visual scenes

Each visual everyday scene (i.e. "complex scene" or "natural scene") is characterized by a context (background or global level; for example, a forest or a restaurant) and one or more "objects", corresponding to living or non-living entities (e.g., a deer or a glass). This local level concerns the perception of details on the scene. Many studies have shown a remarkable ability of humans to categorize these images, even from a very brief glance. This requires various complex processes being performed accurately, in a few hundred milliseconds [in adulthood: 22-24, in childhood: 25]. This ability allows us to attribute meaning and continuously adapt to the surrounding environment. However, when we perceive a visual scene, we have access not only to its meaning in a strict sense, but also to the emotions or feelings associated with it [26-27]. Scenes with emotional content (hereinafter referred to as "emotional scenes") are a particular category of complex visual scenes, which is closely linked to the quality of our subjective experience. These scenes include a context and objects (usual items, animals, people) that convey an emotion. The presence of emotional content in a visual scene modifies its inspection [28]. Emotional information has a special status because it is processed with priority over "neutral" information, that is to say information without emotional content [29-31]. According to many researchers, emotional content generates an "emotional bias" [32-33], which influences viewing and reinforces early processing of the stimulus [34]. Indeed, it is essential to extract very quickly and then accurately analyse emotional information from our environment to adapt our behaviour, and this particularly in relation to an approach or avoidance dimension [35].

Particularities in the viewing behaviour and processing of emotional visual information have been found in individuals with different psychopathologies, for example, in phobia [36], depression [37-38] and anxiety disorders [39]. Some studies have highlighted specific aspects in the visual exploration of emotional faces in people with ADHD and have highlighted their importance in differentiating ADHD from other disorders, such as schizophrenia [40] and autism spectrum disorder (ASD) [41]. However, specific patterns of exploration and understanding of visual scenes in various neurodevelopmental disorders have also been reported in the absence of emotional content. For example, in ASD, weaker abilities to process the overall semantic meaning of the scene have been found compared to a control sample [42]. Recently, Hayes and Henderson [43] have also shown that the scanpath (i.e., pattern of ocular inspection) of a visual scene might predict certain clinical conditions such as dyslexia, ASD and attentional deficit disorder (ADD).

The present study

In summary, many studies have highlighted difficulties in understanding and regulating emotions in children and adolescents with ADHD. However, most have used faces, whereas in everyday life, one has to continuously extract and understand contextual emotional information, and to react appropriately to it. The aim of this review was to examine how children and adolescents with ADHD explicitly and implicitly process emotional information in visual scenes compared to individuals with typical development. We also analysed which factors might influence processing of emotional information. For this purpose, we described previous studies of children and adolescents with ADHD that used visual scenes with or without emotional content. It is indeed important to consider the impact of the lack of general understanding of visual scenes (highlighted for example in ASD [42]), which would not be related to the emotional content per se. A better explanation of emotional recognition in children and adolescents with ADHD would help to refine diagnostic criteria, particularly by differentiating social and emotional deficits in ADHD from those presented in other neurodevelopmental disorders, including ASD [44,41]. This finer understanding of the specificities of emotional processing in ADHD would also contribute to improve interventional plans for this disorder [45-46].

Methods

In this review, we employed the PRISMA guidelines for the search strategy and synthesis of scientific literature [47].

Eligibility criteria

This review had considered studies published before March 2019. It included empirical articles using everyday visual scenes as stimuli and excluded studies using only faces. We reviewed studies on children and adolescents with ADHD with or without comorbidity, with the constraint that at least one ADHD group of participants in the study did not have comorbidity. We excluded from our work previous reviews, comment articles and book chapters. The diagnosis of ADHD had to be validated by well-established assessment tools, including the Diagnostic and Statistical Manual of Mental Disorders (DSM-5 and earlier [1]), the Child Schedule for Affective Disorders and Schizophrenia, Present and Lifetime version (K-SADSPL [48]) and Conners' Parent and Teacher Rating Scales [49].

Information sources

The systematic search was performed on PubMed, SCOPUS and ScienceDirect databases for articles written in English and published until March 2019. This review also included other sources, such as Google Scholar.

Search

Different combinations of keywords were used for the literature search: the terms “ADHD” or “attention deficit hyperactivity disorder” were combined with “scene perception”, “scene recognition”, “scene categorization”, “emotional scene”, “emotional picture” and “affective picture”. The first scan was made by analysing titles and reading abstracts. Studies with adults or without visual scene stimuli were excluded. Following the first scan, the full texts of the pre-selected studies were examined. This second scan excluded studies that did not meet the eligibility criteria.

Results

Study selection

The process of identification, screening and selection is depicted in the PRISMA flow diagram (see Figure 1). Seventeen studies are included in this review. Table 1 shows the 15 empirical studies using emotional visual scenes as stimuli for exploring emotional processes in children and adolescents with ADHD, and Table 2 shows the two studies using non-emotional visual scenes.

Among the 15 studies using emotional scenes, four found no difference in the processing of emotional information between children and adolescents with ADHD and children in control groups [50-53]. Seven studies highlighted behavioural differences (in response accuracy and/or response time) between children and adolescents with ADHD and controls during emotion recognition tasks [10, 54-55], memory tasks [56-57] or in the regulation of the interference due to emotional content [58, 14]. Two studies showed differences in oculomotor behaviour during the inspection of visual scenes [54, 59]. Three studies reported differences in the cerebral activity when processing emotional content, using event-related potentials (ERP [60]) or functional Magnetic Resonance Imaging (fMRI [61- 62]).

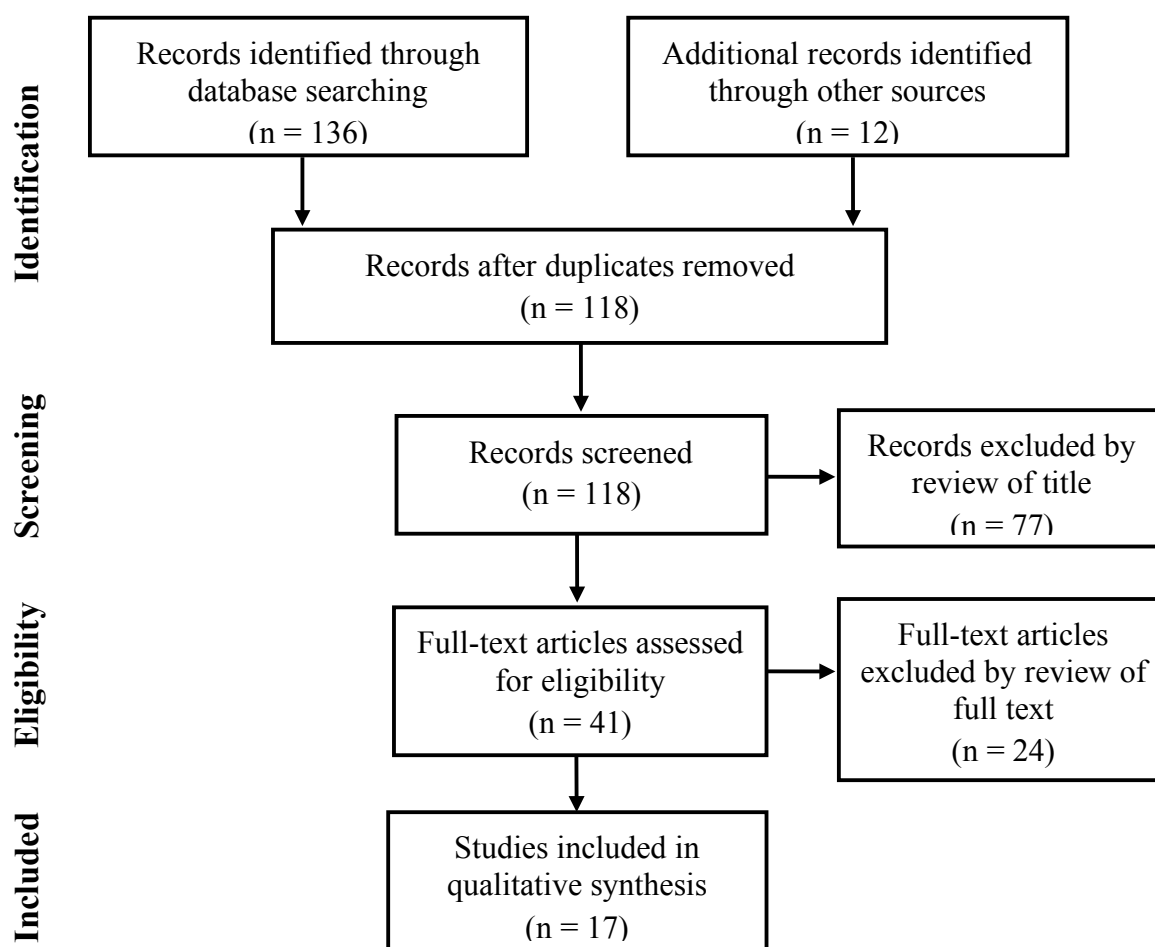


Figure 1. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

Types of images used by the different studies

Only three of the studies with emotional scenes did not use the International Affective Picture System (IAPS [63]), but used photography from popular French media magazines [10], found in the internet [54] or drawings [55]. In these three studies, faces expressing emotions were masked to evaluate the understanding of the emotional context of the scene. Five studies used, besides IAPS stimuli, images from the EmoMadrid affective picture database (<http://www.uam.es/CEACO/EmoMadrid.htm>; [14, 58, 60] or other non-specified images [56-57]. Considering the two studies using non-emotional scenes, one [64] used the images originally employed in [65], and the other [66] used black and white drawings of real-world events.

Table 1: Processing of emotional visual scene processing in ADHD individuals, compared to Typically Developing (TD) individuals.

Study	Participants (age: range or mean (SD), sample size, comorbidity)	Tasks, stimuli and measures	Findings
Da Fonseca et al. [10]	Age 5-15 27 ADHD-C (21 boys). Comorbidity: 6 ODD and 1 CD. 27 TD (21 boys).	Recognition of neutral objects, facial expression and contextual emotion (pictures from magazines); Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL); Conners' Parent and Teacher Rating Scale; WISC-III.	ADHD showed more errors than TD in recognizing facial and contextual emotion, but not in object recognition. No difference between the sub-groups of ADHD (with or without ODD and with or without CD comorbidity) was found.
Herpertz et al. [50]	Age 8-13 Only boys. 43 ADHD; 54 ADHD + CD; 21 CD; 43 TD	Self-rating of the International Affective Picture System (IAPS); Conners' Parent and Teacher Rating Scales; Child Behavior Checklist (CBCL); Anxiety Questionnaire for Pupils; Depression Inventory for Children and Adolescents.	Compared to TD, ADHD rated neutral pictures as been more arousing and negative pictures as been less arousing. CD children with or without ADHD showed lower autonomic responses to all pictures, independent of the emotional content.
Hwang et al. [61]	ADHD:14,53 (2.00), TD: 13.91 (2.13). 26 ADHD (11 ADD; 3 ADHD-IH; 9 ADHD-C). 35 TD.	Affective Stroop Task with IAPS pictures during fMRI; K-SADS; Conners' Parent Rating Scale.	No specific behavioural effect of emotional content was found. At fMRI, ADHD showed disrupted recruitment of cerebral regions implicated in executive attention, especially when the task demand increased (incongruent trials). Moreover, they presented stronger

			responses to positive stimuli.
Krauel et al. [56]	ADHD:12,9 (1.2); TD: 12.7 (1.7) All boys. 18 ADHD (14 ADHD-C; 4 ADD). Comorbidity: 11 with ODD/CD. 15 TD.	Incidental memory task (semantic or perceptual) with IAPS and similar pictures from other sources during ERP recording; K-SADS-PL; Youth Self-Report; CBCL.	Memory performance was impaired for “non-salient” (i.e., neutral) stimuli in ADHD compared to TD, but the initial perceptual and semantic processing of the images was preserved.
Krauel et al. [57]	ADHD:13,9 (1.3), TD: 13.9 (1.8) All boys. 14 ADHD; 16 ADHD + ODD/CD. 25 TD.	Incidental memory task (semantic or perceptual) with IAPS and similar pictures from other sources; K-SADS-PL; Youth Self-Report.	Memory for neutral pictures was lower in ADHD than in TD, while no differences were reported for positive and negative pictures. In ADHD + ODD/CD, compared to TD, memory was lower for positive and neutral pictures, but was the same for negative pictures.
López-Martin et al. [60]	Age 8-13 24 ADHD (19 boys).Comorbidity: 4 with ODD. 24 TD (17 boys).	Go/No-Go task with IAPS and EmoMadrid (CEACO) irrelevant distractor pictures during ERP recording; K-SADS-PL; ADHD Rating Scale-IV.	ADHD showed the same behavioural performance as TD, but an overactivation of inhibition neural mechanisms for emotional (positive and negative) contexts.
López-Martin et al. [58]	Age 8-13 All boys. 20 ADHD; Comorbidity: 3 with ODD. 20 TD.	Participants performed a digit categorization task with IAPS and EmoMadrid irrelevant picture distractors during ERP recording; K-SADS-PL; ADHD Rating Scale-IV.	Behavioural and ERP results showed ADHD had greater susceptibility to emotional distraction (independent of image’s valence) compared to TD.
Pishyareh et al. [59]	Age 6-11 All boys. 30 ADHD. 30 TD.	Eye-movement recording during free viewing pairs of IAPS pictures (positive-neutral; negative-neutral; positive-negative; neutral-	Compared to TD ADHD spent less time viewing positive pictures and more time viewing negative ones.

		neutral); Persian version of K-SADS-PL interview; Conner's Parent Rating Scale revised test.	
Serrano et al. [54]	Age 8-12 19 ADHD (78,9% boys). 26 TD (46,2% boys).	Eye-movement recording during an emotional face evaluation task, with Pictures of Facial Affect (POFA) and an emotional situation task, with pictures in which faces expressing emotions were masked; Impairment Rating Scale (IRS) for parents and teachers; Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II)	ADHD looked at the same relevant areas of the image as the TD. However, for some emotions (depending on the task), they spend less time on the images while been slower in recognizing the conveyed emotion.
Shin et al. [55]	Age 6-10 All boys. 42 ADHD (without comorbidity). 27 TD.	Emotional recognition test revised (ERT) with substest on facial emotion recognition and contextual understanding (CU); Continuous Performance Test; K-SADS-PL; WISC-III	No significant difference was found between ADHD and TD for recognition of facial emotions, but ADHD obtained significantly lower scores at the CU task.
Taskiran et al. [51]	Age 6-12 20 ADHD ED- (18 boys; 7 ADHD-C; 12 ADD; 1 ADHD-HI; 5 ODD). 28 ADHD ED+ (24 boys; 20 ADHD-C; 5 ADD; 3 ADHD-HI; 18 ODD). 22 TD (16 boys).	Autonomic reactivity (Electrodermal Activity: EDA; and Heart Rate Reactivity: HRR) recording during emotional evaluation (valence and arousal) of IAPS stimuli; K-SADS-PL; CBCL; Strengths and Difficulties Questionnaire (SDQ), McMaster Family Assessment Device (FAD), School-Age Temperament Inventory (SATI), and Conners' Parent Rating Scale	Compared to ADHD without ED and to TD, ADHD with ED were more oppositional, they had more social and emotional problems and their HRR did not change with different emotional content (positive, negative or neutral). Authors concluded that two different subgroups may be distinct in children with ADHD with and without ED.
Van Cauwenberge et al. [53]	Age 8-15 38 ADHD (19 boys; 11 ADHD-C; 13	Approach-avoidance paradigm with IAPS stimuli (positive and negative content); Diagnostic	ADHD and TD showed the same ability to regulate emotional

	ADD; 4 ADHD-HI; 11 ODD). 43 TD (26 boys).	Interview Schedule for Children—IV (DISC-IV); Abbreviated version of WISC-III-NL; Social Communication Questionnaire (SCQ); Disruptive Behavior Disorder Rating Scale (DBDRS).	responses to positive and negative pictures.
Van Cauwenberge et al. [52]	Age 8-15 29 ADHD (20 boys; 12 ADHD-C; 14 ADD; 4 ADHD-HI). Comorbidity: 11 ODD and 7 learning disorder. 38 TD (25 boys).	Emotional N-back task with IAPS background pictures; DISC-IV; SCQ; Abbreviated version of WISC-III-NL; Only for control group: DBDRS.	ADHD were distracted by all background pictures (emotional and neutral). This suggests a generic interference deficit and not a specific emotion regulation deficit.
Vetter et al. [62]	Age 11-17 All boys. 25 ADHD without comorbidity. 25 TD.	fMRI recording during an emotional attention task with IAPS images; CBCL; Parent- and self-rated ADHD scale of the comprehensive diagnostic system for mental disorders in childhood and adolescence (DISYPS II); Mini International Neuropsychiatric Interview for Children and Adolescents (M.I.N.I.-Kid); short version of WISC-IV.	ADHD committed more errors than TD with all type of pictures (emotional and neutral). They also showed an increased activation in the left inferior frontal gyrus and in the left amygdala when processing negative emotional picture, suggesting a hyper responsiveness to this type of stimuli
Villemonteix et al. [14]	Age 8-13 33 ADHD (28 boys; 26 ADHD-C; 7 ADD; 7 ODD). 24 TD (19 boys).	Emotional N-back test with IAPS and EmoMadrid irrelevant pictures; STAI-C; K-SADS-PL, ADHD Rating Scale IV; short version of WISC-IV.	ADHD showed longer response time for negative stimuli than TD.

ADHD-C: combined-type ADHD; ADHD-HI: predominantly hyperactive-impulsive; ADHD, ADD: predominantly inattentive; CD: conduct disorder, ED-: without emotional dysregulation; ED+: with emotional dysregulation; ODD: oppositional defiant disorder; TD: typically developing individuals; fMRI: functional Magnetic Resonance Imaging.

Table 2: Processing in non-emotional visual scene processing in children and adolescents with ADHD, compared to Typically Developing (TD) individuals.

Study	Participants (age: range or mean (SD), sample size, comorbidity)	Tasks, stimuli and measures	Findings
Karatekin, & Asarnow [66]	Age 9-20 30 ADHD (19 boys). Comorbidity: 8 ODD, 11 Anxiety disorder, 7 mood disorder, 13 schizophrenia. 27 TD (12 boys).	Participants responding to 3 questions (global meaning, focal information and counting) on 5 black and white drawings of complex visual scenes during eye-movement recording; K-SADS; Peabody Picture Vocabulary Test-Revised.	No significant difference was found between ADHD and TD participants in scanning or in inhibiting eye movements to irrelevant details. ADHD participants showed shorter fixation duration compared to TD.
Türkan et al. [64]	Age 8-11 24 ADHD (18 boys; 16 ADHD-C; 5 ADD; 3 ADHD-HI). 24 TD (18 boys)	Eye-movement recording during a change detection task with real-world scenes (originally used in Rensink et al., 1997); CBCL; Teacher report form; Child and Adolescent Behavior Disorders Screening and Rating Scale; Ishihara Color Blindness Test; Edinburgh Handedness Inventory	ADHD and TD participants preferentially directed attention towards semantically important areas of the scene. However, fixation duration on marginal regions was shorter in ADHD than in TD participants. This may be explained by the fact that ADHD individuals have more difficulty in disengaging attention in order to process information that is more marginal.

Explicit emotional scene processing

Among the studies we reviewed, seven asked participants an explicit judgement of emotions [10, 50-55]. Three reported weaker emotional recognition abilities in children and adolescents with ADHD compared to control children [10, 54-55]. These three studies used the same type of methodology, with two tasks proposed to the participants: a facial emotion

recognition task and a contextual understanding task, in which participants had to infer the emotion felt by a face-masked person presented in a visual scene. Participants needed, therefore, to understand the meaning of the scene and the relationship between the context and the target person (for example, at the award ceremony of a competition, there was a winning child, and next to him another child, with a masked face, who lost the competition and for whom the inferred emotion should be sadness). More precisely, Da Fonseca et al. [10], required participants to select the emotion felt by a face-masked person choosing one of four basic emotions: joy, anger, sadness and fear. Serrano et al. [54] and Shin et al. [55] also added disgust and surprise.

Four other studies [50-53] required to rate on a continuum the emotional valence and the arousal related to visual scenes. According to these studies, children and adolescents with ADHD rated valence in the same way as control children and adolescents. Differences were only reported for the rating of arousal. In Herpertz et al. [50], children and adolescents with ADHD, compared to controls, rated positive and neutral images as more arousing and negative images as a less arousing. The two studies of Van Cauwenberg and colleagues [52-53] also found a higher arousal rating of positive images in children and adolescents with ADHD than in controls, but without differences for neutral and negative images. One of these studies [53] used an approach-avoidance paradigm in which participants had to judge the emotional valence of an image. Two conditions were created. In the “consistency condition”, the action to approach or to avoid an image matched image’s emotional content (i.e., participants had to approach positive images and to avoid negative ones). In the “inconsistency condition”, the action to approach or to avoid an image mismatched image’s emotional content (i.e., participants had to approach negative images and avoid positive images). The authors found no difference in inhibition between ADHD and controls. Finally, Taskiran et al. [51] found a higher arousal rating for negative images, but only in children with ADHD with emotional dysregulation.

To sum up, the studies reviewed in this section have suggested that children and adolescents with ADHD show normal abilities when evaluating the affective valence (i.e., positive or negative) of the image, but are impaired when categorizing the basic emotion conveyed by the image.

It is worth noting that the images used in Taskiran et al.’s study, which did not find any emotional processing difference between ADHD and typically developing children, were exclusively IAPS stimuli. In contrast, the other three studies used a variety of images not from the IAPS database. The fact that different tasks and images are used might also, at least in part,

explain the mixed evidence in emotional processing reported when comparing participants with ADHD and controls. An additional aspect of discrepancy between studies may concern the type of emotional evaluation (basic emotion versus dimensional evaluation) or the evaluation level (evaluating the image itself versus inferring the emotional state of the person depicted in the image).

Implicit emotional scene processing

Eight studies examined the effects of implicit processing of the emotional content of visual scenes on a task performance.

Among these studies, one analyzed the influence of the emotional content on free viewing of a pair of images [59] and showed shorter fixation durations for positive images and longer fixation durations for negative images in children with ADHD than in control children. The seven remaining studies analyzed the effect of irrelevant emotional distractors on performance at various tasks.

Three studies highlighted behavioural differences related to the presence of emotional content in children and adolescents with ADHD during the performance of various tasks compared to those in control children. Villemonteix et al. [14] used an emotional N-back task. Letters were presented in the centre of a positive, negative or neutral distractor image, and children had to report the letter that was presented in the previous image (1-back condition) or two images prior (2-back condition). The authors showed a stronger distracting effect of negative images, resulting in longer response time, in children with ADHD compared to controls. In the digit categorization task of the Lopez-Martin et al.'s study [58], numbers were presented in the centre of a positive, negative or neutral distractor image. Participants had to respond with one hand for even numbers and with the other hand for odd numbers during the ERP recording. A higher distractibility due to emotional content was found in ADHD children compared to the controls. ADHD children, compared to controls, showed longer response times and a larger anterior N2 amplitude, a component that is enhanced by highly relevant stimuli [67]. This was true for all emotional images, either negative or positive. Finally, the study by Krauel et al. [56] used, first, an ERP implicit encoding task consisting of either a perceptual or a semantic judgement of the image and, second, an image recognition task. Children and adolescents with ADHD performed both encoding tasks similarly to controls, but showed lower image recognition performance for neutral images, for which ERP results suggested a weaker encoding. The presence of an emotional content, thus, appears to improve image encoding in

ADHD individuals. These findings are in line with previous reports of better memory for emotional stimuli than for neutral stimuli [68-69], result that has been partly related to an enhancement of visual analysis of emotional stimuli [31,33].

Only one study did not find any difference in the processing of emotional images between children and adolescents with ADHD and controls. Van Cauwenberge et al. [52] used an emotional N-back task and, contrary to Villemonteix et al. [14], they did not report any stronger distracting effect of emotional images compared to neutral ones. However, there were some differences in method between these two studies. Van Cauwenberge et al. [52] employed a 0-back condition, where participants had to respond to the current image, and a 1-back condition, where participants had to respond to the image seen just before. Villemonteix et al. [14] compared a 1-back condition to a 2-back condition, thus increasing the overall cognitive load of the task. Additionally, Villemonteix et al. [14] used neutral, positive and negative images as background images, while Van Cauwenberge et al. [52] used in addition a blank background. In this latter study, participants with ADHD showed slower response times when comparing the blank-screen condition to the neutral-image condition, while this pattern was reported for the controls. The authors, thus, described a general interference effect in ADHD due to the presence of a background image that is not specific to image's emotional content.

Finally, the last three studies we review, did not report any behavioural differences between children and adolescents with ADHD and controls, but found cerebral differences in the processing of emotional information. Lopez-Martin et al. [60] used a Go/No-Go task with positive, negative and neutral distracting background images, during ERP recordings. No difference in response accuracy or in response time was showed between the two groups of participants. However, the ERP results were interpreted as a hyper-activation of inhibition mechanisms (larger P3 amplitude) in children with ADHD compared to controls, particularly for positive and negative images with a high arousal. According to the authors, the children with ADHD seemed to make greater inhibition efforts to achieve the same behavioural outcomes as the control children, especially when an emotional context was presented. The two other studies were conducted with fMRI. Hwang et al. [61] proposed what they called an “emotional Stroop task”, in which participants saw a positive, negative or neutral image and were then presented with several numbers in a consistent (for example, the number three written three times) or in an inconsistent manner (for example, the number three written five times). Participants had to report how many numbers they saw: children and adolescents with ADHD were slower than controls for all images (neutral and emotional). Accordingly, at the cerebral level, ADHD participants presented decreased recruitment of regions related to executive

attention, compared to controls, especially when the cognitive load increased (i.e. in the inconsistent trials). ADHD participants had also stronger responses in the temporal regions for positive stimuli during inconsistent than during consistent trials, and weaker connectivity between the amygdala and the striatum during viewing of positive images compared to neutral ones. These results suggest amplified emotional cerebral responses in children and adolescents with ADHD for positive stimuli. Vetter et al. [62] used an emotional attention task in which participants saw two pairs of images simultaneously: an emotional pair and a scramble-image pair. They had to judge only one pair and decide whether its two image components were identical or not. ADHD participants made more errors than controls (but with the same response time) for both neutral and emotional images. At the cerebral level, however, negative images led to a stronger activation of the left amygdala and of the inferior frontal gyrus in ADHD participants than in controls. This hyper-activation of the amygdala was also associated with incorrect trials and would thus, according to the authors, be responsible for the ADHD hyper-responsiveness to negative stimuli, which was the major cause of error.

To sum up, all the studies reviewed in this section (except that of [52]) have found that emotional information conveyed by visual scenes is processed differently by individuals with ADHD compared to typically developing peers. This difference has been found at the behavioural level and at the cerebral level, either independently of the valence of the images [56, 58-60], or only for negative [14, 62] or positive images [61].

Overall scene understanding

Five of the reviewed studies analysed semantic processing of visual scenes in children and adolescents with ADHD.

In their study, Da Fonseca et al. [10] used two tasks, one requiring facial emotion recognition and the other requiring understanding of contextual information conveyed by the scene. In the latter task, participants had to infer the emotion felt by a face-masked person (for details, see “explicit emotional scene processing” section) or to recognize a masked object within a visual scene. This allowed to examine the general understanding of the scene and the ability to recognize the presence of non-emotional objects. Participants with ADHD performed worse than controls in recognizing facial and contextual emotions, whereas object recognition did not differ between the two groups. These results suggested preserved understanding of the general meaning of a scene and a deficit in ADHD specific to emotional material.

The first part of the study by Krauel et al. [56] consisted of an incidental encoding of images including a geometric shape positioned within an everyday scene. The second part of this study was a recognition task (for details see “implicit emotional scene processing” section). Encoding was realized through two tasks that required different levels of processing: a semantic task, where participants decided whether the visual scene contained humans or not, and a perceptual task, where participants decided if a shape embedded in the image was a circle or a rectangle. No performance differences at either encoding tasks were found between ADHD participants and controls. However, recognition was impaired for neutral images compared to emotional images in ADHD participants, although it was not the case in controls. This study, therefore, supported the claim of a preserved general understanding of the scene in children and adolescents with ADHD.

The other three studies used eye-movement recording to analyze the scene viewing pattern in children and adolescents with ADHD compared to controls. These studies differed in their task (change detection [64]; free viewing to complete questions on the content of the scene [66]; emotional judgement [54]) and in the type of scenes used (photographs [54, 64]; drawings [66]). However, they all found the same patterns of results. Participants with ADHD, compared to controls, showed similar fixated regions (rich in semantic or emotional information), but they made significantly shorter fixations. Their initial information gathering appeared preserved either for emotional [54] or non-emotional [64, 66] scenes. Nevertheless, shorter fixation durations might impair information gathering from the whole scene and would, according to the authors, lead to weaker performance which was reported for ADHD participants, compared to controls, in change detection [64], in emotion judgement [54], or in the analysis of semantic and perceptual information carried by the images [66].

All the studies reviewed in this section, therefore, supported the finding that information-processing skills, at least for initial information gathering, are preserved in children and adolescents with ADHD, allowing understanding of the general meaning of the scene. Children and adolescents with ADHD focus their attention on the same important areas for the meaning of the scene as controls do, and their ability to recognize scenes seems to be generally preserved. Deficits were reported only in fixation duration that is to say in the time allocated for analysing more precisely visual information.

Scene recognition, emotional processing and core deficits in ADHD

Some authors have related specificities in visual scenes processing in children and adolescents with ADHD to the core deficits of the disorder. Different theoretical models have been proposed to describe and explain ADHD's deficits (for a review of the main models, see [70]). These models differ, in particular, in the respective roles of the "cold" versus the "hot" executive functions [71-73]. Cold executive functions are required for resolving problems in the absence of emotional content. They are considered by top-down ADHD theories, for which the difficulties associated with ADHD would come mainly from deficits in controlling inhibition [74]. Hot executive functions are required for problems with emotional content, they are related to social cognition as emotional processing, theory of mind, and they are considered by bottom-up ADHD theories, for which ADHD also stem from a lack of motivation and an overestimation of emotional stimuli [72].

In this review, in line with the top-down theories, Türkan et al. [64] showed lower accuracies in change detection in children with ADHD compared to controls, but with the same response time. The authors explained the results by the characteristic impulsivity of ADHD. They also showed shorter fixation duration on the most informative areas of the scene resulting in less depth processing of this part of the image, associated with a lower disengagement from a priority informative area of the scene to a more marginally informative one in ADHD participants compared to controls. Changes in details of a priority informative area and changes in marginally informative areas were so less detected by ADHD participants than by controls. To summarize, for these authors, the behavioural pattern should be linked to low sustained attention capacities and ineffective attentional switching mechanisms.

Serrano et al. [54] linked the shorter fixation duration in children with ADHD compared to controls with the attentional deficit specific to the syndrome. However, they also reported that impairment in emotional recognition in children were more correlated with parents' impulsivity and hyperactivity rating than with inattention as rated in the impairment rating scale. Concerning emotional processing, specifically, according to Shin et al. [55], the deficits in emotion recognition in children with ADHD were related to attention deficits, assessed in their study by the ADHD diagnostic system (ADS), a computerized continuous performance test. Yet, their study included only emotional stimuli, thus it is difficult to draw conclusions about any specific effect of emotional information. Other studies related difficulties, regardless of the type of images (emotional or neutral), to deficits in working memory [14] or in inhibition [52], which are frequently found in ADHD.

However, the stronger interference effect in ADHD participants than in controls, caused by emotional images compared to neutral images would be, for several authors, associated with a stronger influence of these images and therefore with bottom-up attentional processes and “hot” executive function. Indeed, hyper-activation of the amygdala in children and adolescents with ADHD compared to controls has been shown in the presence of emotional stimuli [62]. The amygdala plays a primary role in processing emotional information and any stimulus relevant to the individual [75]. Thus, inhibiting irrelevant emotional information would have a higher cognitive load for children and adolescents with ADHD than for controls [58, 60-61]. The distractibility of children and adolescents with ADHD would then come from involuntary automatic responses to salient emotional information that need to be controlled by top-down inhibitory processes that are known to be less efficient in ADHD.

ADHD and comorbidities

Among the 17 studies included in this review, only two studies reported that their sample consisted of ADHD participants without comorbidity [55, 62], and five studies did not provide information [54, 56, 59, 61, 64]. Ten studies reported comorbid disorders associated with ADHD [10, 14, 50-53, 57-58, 60, 66]. These associated comorbid disorders were the oppositional defiant disorder (ODD) and the conduct disorder (CD), which are indeed the ones most frequently associated with ADHD [76].

The comorbidity between ADHD and ODD or CD has been analysed in seven studies. Three did not find any result difference when adding or removing data from ADHD participants showing either comorbidity [10, 52-53]. Among the four remaining studies, Villemonteix et al. [14] found a stronger distracting effect of negative images (larger effect size) in ADHD children with ODD than in ADHD children without ODD. This study also controlled for a possible effect related to anxiety, which was not significant. Krauel et al. [57] compared performance in a memory task between three groups of children and adolescents: with ADHD without comorbidity (“ADHD-only”), with ADHD associated with ODD (“ADD-ODD”) and with typical development (controls). The ADHD-only group showed impaired recognition of neutral images, compared to controls, but similar performance for positive and negative images. The ADHD-ODD group showed impaired recognition of neutral and positive images compared to controls, with a similar performance only for negative images. The authors, therefore, concluded that the emotional content improved image encoding in ADHD-only participants, whereas only negative images were salient enough to improve encoding in ADHD-ODD

individuals. Herpertz et al. [50] recorded the physiological responses (skin conductance and heart rate reactivity) of participants during an emotional judgement task. ADHD-only children showed the same physiological responses as controls, while those with comorbid CD had weaker physiological responses, regardless of whether the material was emotional or neutral. Taskiran et al. [51] compared heart rate reactivity of children with ADHD with and without emotional dysregulation (ED, as assessed using the Child Behaviour Checklist 6-18) during an emotional judgment task. Unlike control children and ADHD children without ED, ADHD children with ED showed the same level of heart rate reactivity independently of the type of emotional content of the experimental material. The authors concluded that ED might be a dimensional trait that would differentiate a particular subtype of ADHD.

Discussion

The main purpose of this review was to investigate how children and adolescents with ADHD explicitly and implicitly process emotional information in visual scenes compared to children and adolescents with typical development. Particularities in the explicit and implicit processing of emotional information were found in most of the studies reviewed here, although the results differed according to the type of task, the experimental method and the theoretical model underlying the concept of emotion the authors referred to (basic emotional models versus dimensional models).

Among the studies examining explicit emotion recognition, those based on dimensional models described identical abilities in children and adolescents with ADHD and in peer controls for emotion judgement according to a valence continuum (negative to positive), but differences for judgment of the level of arousal [50, 52-53]. This suggests that ADHD influences the subjective feeling of emotion intensity, rather than the type of the emotion experienced. Studies requiring emotional judgement at a basic level (basic emotions models) all reported impaired emotion recognition in children and adolescents with ADHD compared to controls [10, 54-55]. However, these studies did not require judging the emotion conveyed by the whole visual scene, but the emotion of a particular individual present in the scene. As besides the understanding of the general meaning of the scene, this task required to infer a person's specific feelings, it seems difficult to tease apart whether the performance impairment was due primarily to a lack of emotion recognition or it was secondary to other deficits that are characteristic of ADHD, such as those affecting executive functions [74, 77] or theory of mind [78].

To assess accurately the influence of emotions on task performance, it seems important to ensure that children and adolescents with ADHD have a preserved understanding of the general meaning of the visual scene in the absence of emotional content. However, few studies have focused on this aspect and on where ADHD individuals look within a scene. They all reported that ADHD participants and controls do not differ in the initial information gathering from the scene and in scene understanding [10, 54, 56, 64, 66]. However, some of these studies showed shorter fixation durations [54, 66] and a stronger difficulty in disengaging from the main information [64] in ADHD participants compared to controls, thus, resulting in a more superficial processing of the scene. These differences appeared to be related to the core symptoms of ADHD and might influence visual scene processing regardless the presence of emotional content. Other studies have reported longer response time in children and adolescents with ADHD than in controls in the case of emotional scenes [14, 58]. A processing cost in the presence of emotional content has also been reported at a cerebral level (e.g. ERP and fMRI measures), suggesting a need for stronger inhibitory control in children and adolescents with ADHD to achieve the same performance as control individuals [60-62]. Moreover, the distracting effect of negative emotional information appears further reinforced in the presence of a comorbid disorder, such as ODD or CD [14, 57] which are both frequently associated with ADHD.

Finally, it seems that the difficulties in processing emotional information carried by visual scenes in children and adolescents with ADHD are related to a stronger emphasis on emotional content reinforcing bottom-up processing and to top-down deficits related to the disorder, including attention deficit. The superior salience of emotional information might nevertheless sometimes be beneficial in helping children and adolescents with ADHD to improve their performance, especially in memory tasks [56-57]. However, this stronger influence of emotional content seems to require a higher control effort to children and adolescents with ADHD than to controls. This might explain why youth with ADHD present social and cognitive difficulties in emotional environments. However, a child or a teenager does frequently experience such environments, whether in the classroom, at home or during leisure activities. The effort they make to process (understand or inhibit) this emotional information may then lessen their attentional resources and cognitive engagement in the ongoing tasks. In addition, multiple, frequent contacts with emotional stimuli in their visual environment gradually reduces ADHD individuals' ability to regulate their emotions. This excessive emotional activation may then lead to behavioural disorders. Socio-emotional difficulties, including peer rejection, behavioural disorders and socio-cognitive deficits, are frequently

observed in children and adolescents with ADHD [79], and their presence increases the risk of developing subsequent psychopathologies [80]. Although the ability to recognize and regulate emotions is only one part of the mechanisms at work in emotional dysregulation [7], it seems highly relevant to take into account these aspects in interventions by targeting top-down control abilities, emotional recognition and regulation [45-46].

It should be noted that most of the studies examined in this review considered the different subtypes of ADHD in children and adolescents, but differences in performance between these ADHD subtypes were not analysed, primarily due to insufficient sample size. However, the different subtypes, because of the different underlying deficits, may lead to different emotional processing profiles [81]. It would be important for future studies to analyse separately the effect of each subtype on the processing of emotional information. Moreover, in the studies considered sex and age were not taken into account despite these variables may impact emotional processing [82]. It should also be noted that, to respect our eligibility criteria, only a limited number of studies could be included in this review, and the diversity of the tasks used by these studies makes the synthesis of the results, and their possible interpretations, difficult. Future research will have to further explore the processing of visual emotional scenes.

Conclusions

This review showed that the difficulties in processing emotional information frequently reported in children and adolescents with ADHD concerning facial expressions are also found when it comes to visual scenes. These difficulties may arise from a stronger bottom-up impact of emotional content in ADHD children and adolescents, which increases the intensity of their emotional experience, and from the core deficits of the disorder, which decrease their overall ability in everyday scenes processing. As a result, emotional information may generate an over-stimulation in children and adolescents with ADHD, and, therefore, may require higher inhibition abilities, which are, however, in turn, less effective in this population than in typically developing individuals. These mechanisms would explain why regulating emotions and maintaining effective cognitive activity in environments that may be emotionally charged –as it often happens in everyday life– is very difficult for children and adolescents with ADHD. However, few studies have used visual scenes so far. It is necessary that future research continues in this direction to improve understanding of the specificities of individuals with ADHD in processing emotional information in scenes. Moreover, it seems essential to take into

account both the cognitive and emotional regulation aspects in intervention plans for children and adolescents with ADHD.

The authors declare that they have no conflict of interest.

References

1. American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5)*. Washington, American Psychiatric Association
2. Sonuga-Barke EJS (2002) Psychological heterogeneity in AD/HD: A dual pathway model of behaviour and cognition. *Behav Brain Res* 130: 29–36
3. Barkley RA (2010) Why emotional impulsiveness should be a central feature of ADHD. *The ADHD Report* 18(4): 1-5
4. Bunford N, Evans SW, Becker SP, Langberg J M (2015) Attention-deficit/hyperactivity disorder and social skills in youth: A moderated mediation model of emotion dysregulation and depression. *J Abnorm Child Psychol* 43(2): 283-296
5. Shaw P, Stringaris A, Nigg J, Leibenluft E (2014) Emotion dysregulation in Attention Deficit Hyperactivity Disorder. *Am J Psychiatry* 171(3): 276-293
6. Thompson RA (1991) Emotional regulation and emotional development. *Educ Psychol Rev* 3(4): 269-307
7. Graziano PA, Garcia A (2016) Attention-deficit hyperactivity disorder and children's emotion dysregulation: A meta-analysis. *Clin Psychol Rev* 46: 106-123
8. Aspan N, Bozsik C, Gadoros J, Nagy P, Inantsy-Pap J, Vida P, Halasz J (2014) Emotion recognition pattern in adolescent boys with attention-deficit/hyperactivity disorder. *BioMed Res Int* <http://dx.doi.org/10.1155/2014/761340>
9. Corbett B, Glidden H (2000) Processing affective stimuli in children with attention-deficit hyperactivity disorder. *Child Neuropsychology* 6:144-155
10. Da Fonseca D, Seguier , Santos A, Poinso F, Deruelle C (2009) Emotion understanding in children with ADHD. *Child Psychiatry Hum Dev* 40(1): 111-121
11. Williams LM, Hermens DF, Palmer D, Kohn M, Clarke S, Keage H, Gordon E (2008). Misinterpreting Emotional Expressions in Attention-Deficit/Hyperactivity Disorder: Evidence for a Neural Marker and Stimulant Effects. *Biol Psychiatry* 63(10): 917–926. <https://doi.org/10.1016/j.biopsych.2007.11.022>
12. Yuill N, Lyon J (2007) Selective difficulty in recognising facial expressions of emotion in boys with ADHD. *Eur Child Adolesc Psychiatry* 16(6): 398-404
13. Pitzianti M, Grelloni C, Casarelli L, D'Agati E, Spiridigliozzi S, Curatolo P, Pasini A (2017) Neurological soft signs, but not theory of mind and emotion recognition deficit distinguished children with ADHD from healthy control. *Psychiatry Res* 256: 96-101
14. Villemonteix T, Marx I, Septier M, Berger C, Hacker T, Bahadori S, Acquaviva E, Massat I (2017) Attentional control of emotional interference in children with ADHD and typically developing children: An emotional N-back study. *Psychiatry Res* 254: 1-7
15. Borhani K, Nejati V (2018) Emotional face recognition in individuals with attention-deficit/hyperactivity disorder: a review article. *Dev Neuropsychology* 43(3): 256-277. <https://doi.org/10.1080/87565641.2018.1440295>
16. Aviezer H, Trope Y, Todorov A (2012) Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science* 338(6111): 1225-1229
17. Barrett LF, Mesquita B, Gendron M (2011) Context in emotion perception. *Curr Dir Psychol Sci* 20(5): 286-290

18. Righart, R., & De Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cogn Affect Behav Neurosci* 8(3): 264-272
19. Wenzler S, Levine S, van Dick R, Oertel-Knöchel V, Aviezer H (2016) Beyond pleasure and pain: Facial expression ambiguity in adults and children during intense situations. *Emotion* 16(6): 807-814
20. Aviezer H, Ensenberg N, Hassin RR (2017) The inherently contextualized nature of facial emotion perception. *Curr Op Psychol* 17: 47-54
21. Wieser MJ, Brosch T (2012) Faces in context: A review and systematization of contextual influences on affective face processing. *Front Psychol* 3(471): 1-13
22. Davenport J, Potter M (2004) Scene consistency in object and background perception. *Psycholog Sci* 15: 559-564
23. Fabre-Thorpe M, Delorme A, Marlot C, Thorpe S (2001) A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *J Cogn Neurosci* 13(2): 171-180
24. Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381: 520-522
25. Batty M, Taylor MJ (2002) Visual categorization during childhood: an ERP study. *Psychophysiology* 39(4): 482-490
26. Barrett LF, Bar M (2009) See it with feeling: affective predictions during object perception. *Philos Trans R Soc Lond B Biol Sci* 364(1521): 1325-1334
27. Malcolm G L, Groen I I, Baker C I (2016) Making sense of real-world scenes. *Trends Cogn Sci* 20(11): 843-856
28. Kuniecki M, Wołoszyn KB, Domagalik A, Pilarczyk J (2017) Effects of Scene Properties and Emotional Valence on Brain Activations: A Fixation-Related fMRI Study. *Front Hum Neurosci*. <https://doi.org/10.3389/fnhum.2017.00429>
29. Carretié L (2014) Exogenous (automatic) attention to emotional stimuli: a review. *Cogn Affect Behav Neurosci* 14(4): 1228-1258
30. Pilarczyk J, Kuniecki M (2014) Emotional content of an image attracts attention more than visually salient features in various signal-to-noise ratio conditions. *J Vis* 14(2014): 1–19. <https://doi.org/10.1167/14.12.4>
31. Vuilleumier P (2015) Affective and motivational control of vision. *Curr Op Neurology* 28(1): 29-35
32. Pourtois G, Schettino A, Vuilleumier P (2013) Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. *Biol Psychol* 92(3): 492-512
33. Vuilleumier P (2005) How brains beware: Neural mechanisms of emotional attention. *Trends Cogn Sci* 9(12): 585–594. <https://doi.org/10.1016/j.tics.2005.10.011>
34. Phelps EA, Ling S, Carrasco M (2006) Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychol Sci* 17(4): 292-299
35. Barrett LF, Bliss-Moreau E (2009) Affect as a psychological primitive. *Adv Exp Soc Psychol* 41: 167-218

36. Öhman A, Mineka S (2001) Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol Rev* 108(3): 483-522
37. Burkhouse KL, Owens M, Feurer C, Sosoo E, Kudinova A, Gibb BE (2017) Increased neural and pupillary reactivity to emotional faces in adolescents with current and remitted major depressive disorder. *Soc Cogn Affec Neurosci* 12(5): 783-792
38. Duque A, Vázquez C (2015) Double attention bias for positive and negative emotional faces in clinical depression: Evidence from an eye-tracking study. *J Behav Ther Exp Psychiatry* 46: 107-114
39. Mueller EM, Hofmann SG, Santesso DL, Meuret AE, Bitran S, Pizzagalli, DA (2009) Electrophysiological evidence of attentional biases in social anxiety disorder. *Psychol Med* 39(7): 1141-1152
40. Marsh PJ, Williams LM (2006) ADHD and schizophrenia phenomenology: visual scanpaths to emotional faces as a potential psychophysiological marker? *Neurosci Biobehav Rev* 30(5): 651-665
41. Tye C, Mercure E, Ashwood KL, Azadi B, Asherson P, Johnson MH, McLoughlin G (2013) Neurophysiological responses to faces and gaze direction differentiate children with ASD, ADHD and ASD+ ADHD. *Dev Cogn Neurosci* 5: 71-85
42. Vanmarcke S, Mullin C, Van der Hallen R, Evers K, Noens I, Steyaert J, Wagemans J (2016) In the eye of the beholder: Rapid visual perception of real-life scenes by young adults with and without ASD. *J Autism Dev Disord* 46(8): 2635-2652
43. Hayes TR, Henderson JM (2018) Scan patterns during scene viewing predict individual differences in clinical traits in a normative sample. *PloS One* 13(5): e0196654
44. Mikami AY, Miller M, Lerner MD (2018) Social functioning in youth with attention-deficit/hyperactivity disorder and autism spectrum disorder: transdiagnostic commonalities and differences. *Clin Psychol Rev*. <https://doi.org/S027273581830268X>
45. Southam-Gerow MA, Kendall PC (2002) Emotion regulation and understanding: Implications for child psychopathology and therapy. *Clin Psychol Rev* 22(2): 189–222. [https://doi.org/10.1016/S0272-7358\(01\)00087-3](https://doi.org/10.1016/S0272-7358(01)00087-3)
46. Suveg C, Southam- Gerow MA, Goodman KL, Kendall PC (2007) The Role of Emotion Theory and Research in Child Therapy Development. *Clinical Psychology: Sci Pract* 14: 358-371. doi:10.1111/j.1468-2850.2007.00096.x
47. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group (2010) Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Int J Surgery* 8(5): 336–341
48. Kaufman J, Birmaher B, Brent D, Rao U, Ryan N (1997) Schedule for affective disorders and schizophrenia for school-age children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the Am Ac Child Adolesc Psychiatry* 36(7): 980-988 doi:10.1097/00004583-199707000-00021
49. Conners CK (1997) *Conners' rating scales-revised: technical manual*. Multi-Health Systems Inc, New York
50. Herpertz SC, Mueller B, Qunaibi M, Lichtenfeld C, Konrad K, Herpertz-Dahlmann B (2005) Response to emotional stimuli in boys with conduct disorder. *Am J Psychiatry* 162(6): 1100-1107

51. Taskiran C, Karaismailoglu S, Cak Esen, HT, Tuzun Z, Erdem A, Balkanci ZD, Cengel Kultur SE (2018) Clinical features and subjective/physiological responses to emotional stimuli in the presence of emotion dysregulation in attention-deficit hyperactivity disorder. *J Clin Exp Neuropsychol* 40(4): 389-404
52. Van Cauwenberge V, Sonuga-Barke EJ, Hoppenbrouwers K, Van Leeuwen K, Wiersema JR (2015) “Turning down the heat”: Is poor performance of children with ADHD on tasks tapping “hot” emotional regulation caused by deficits in “cool” executive functions? *Res Dev Disabil* 47: 199-207
53. Van Cauwenberge V, Sonuga-Barke EJ, Hoppenbrouwers K, Van Leeuwen K, Wiersema JR (2017) Regulation of emotion in ADHD: can children with ADHD override the natural tendency to approach positive and avoid negative pictures? *J Neural Transm* 124(3): 397-406
54. Serrano VJ, Owens JS, Hallowell B (2015) Where children with ADHD direct visual attention during emotion knowledge tasks: Relationships to accuracy, response time, and ADHD symptoms. *J Atten Disord* 22(8): 752-763
55. Shin DW, Lee SJ, Kim BJ, Park Y, Lim SW (2008) Visual attention deficits contribute to impaired facial emotion recognition in boys with attention-deficit/hyperactivity disorder. *Neuropediatrics* 39: 323-327
56. Krauel K, Duzel E, Hinrichs H, Lenz D, Herrmann CS, Santel S, Rellum T, Baving L (2009) Electrophysiological correlates of semantic processing during encoding of neutral and emotional pictures in patients with ADHD. *Neuropsychologia* 47(8-9): 1873-1882
57. Krauel K, Duzel E, Hinrichs H, Rellum T, Santel S, Baving L (2009) Emotional memory in ADHD patients with and without comorbid ODD/CD. *J Neural Transm* 116(1): 117-120
58. López-Martín S, Albert J, Fernández-Jaén A, Carretié L (2013) Emotional distraction in boys with ADHD: neural and behavioral correlates. *Brain Cogn* 83(1): 10-20
59. Pishyareh E, Tehrani-doost M, Mahmoodi-gharaie J, Khorrami A, Rahmdar SR (2015) A Comparative Study of Sustained Attentional Bias on Emotional Processing in ADHD Children to Pictures with Eye-Tracking. *Iran J Child Neurol* 9(1): 64-70
60. López-Martín S, Albert J, Fernández-Jaén A, Carretié L (2015) Emotional response inhibition in children with attention-deficit/hyperactivity disorder: neural and behavioural data. *Psycholog Med* 45(10): 2057-2071
61. Hwang S, White SF, Nolan ZT, Williams WC, Sinclair S, Blair RJR (2015) Executive attention control and emotional responding in attention-deficit/hyperactivity disorder - A functional MRI study. *NeuroImage: Clin* 9: 545-554
62. Vetter NC, Buse J, Backhausen LL, Rubia K, Smolka MN, Roessner V (2018) Anterior insula hyperactivation in ADHD when faced with distracting negative stimuli. *Hum Brain Mapp* 39: 2972-2986
63. Lang PJ, Bradley MM, Cuthbert BN (2008) *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-8. University of Florida, Gaines, FL, US
64. Türkan BN, Amado S, Ercan ES, Perçinel I (2016) Comparison of change detection performance and visual search patterns among children with/without ADHD: evidence from eye movements. *Res Dev Disabil* 49: 205-215

65. Rensink RA, O'Regan JK, Clark JJ (1997) To see or not to see: The need for attention to perceive changes in scenes. *Psychol Sci* 8(5): 368–373
66. Karatekin C, Asarnow RF (1999) Exploratory eye movements to pictures in childhood-onset schizophrenia and attention-deficit/hyperactivity disorder (ADHD). *J Abnorm Child Psychol* 27(1): 35-49
67. Chong H, Riis JL, McGinnis SM, Williams DM, Holcomb PJ, Daffner KR (2008) To ignore or explore: Top-down modulation of novelty processing. *J Cogn Neurosci* 20: 120–134
68. Dolcos F, LaBar KS, Cabeza R (2004) Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron* 42(5): 855-863
69. Kensinger EA (2009) Remembering the details: Effects of emotion. *Emot Rev* 1(2): 99-113
70. Sergeant JA, Geurts H, Huijbregts S, Scheres A, Oosterlaan J (2003) The top and the bottom of ADHD: A neuropsychological perspective. *Neurosci Biobehav Rev* 27(7): 583–592. <https://doi.org/10.1016/j.neubiorev.2003.08.004>
71. Castellanos FX, Sonuga-Barke EJ, Milham MP, Tannock R (2006) Characterizing cognition in ADHD: Beyond executive dysfunction. *Trends Cogn Sci* 10(3): 117–123. doi:10.1016/j.tics.2006.01.011
72. Sonuga-Barke EJS (2005) Causal models of attention-deficit/hyperactivity disorder: From common simple deficits to multiple developmental pathways. *Biol Psychiatry* 57(11): 1231–1238. doi:10.1016/j.biopsych.2004.09.008
73. Zelazo PD, Müller U (2002) *Executive function in typical and atypical development*. Oxford, England: Blackwell.
74. Barkley RA (1997) Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychol Bull* 121(1): 65-94
75. Sander D, Grafman J, Zalla T (2003) The human amygdala: an evolved system for relevance detection. *Rev Neurosci* 14(4): 303–316
76. Maughan B, Rowe R, Messer J, Goodman R, Meltzer H (2004) Conduct disorder and oppositional defiant disorder in a national sample: developmental epidemiology. *J Child Psychol Psychiatry* 45(3): 609-621
77. Castellanos FX, Tannock R (2002) Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nat Rev Neurosci* 3(8): 617–628
78. Caillies S, Bertot V, Motte J, Raynaud C, Abely M. (2014) Social cognition in ADHD: irony understanding and recursive theory of mind. *Res Dev Disabilities*, 35(11): 3191-3198
79. Ros R, Graziano PA (2018) Social Functioning in Children With or At Risk for Attention Deficit/Hyperactivity Disorder: A Meta-Analytic Review. *J Clin Child Adolesc Psychol* 47(2): 213–235. <https://doi.org/10.1080/15374416.2016.1266644>
80. Nijmeijer JS, Minderaa RB, Buitelaar JK, Mulligan A, Hartman CA, Hoekstra PJ (2008) Attention-deficit/hyperactivity disorder and social dysfunctioning. *Clin Psychol Rev* 28(4): 692-708

81. Wheeler Maedgen J, Carlson CL (2000) Social functioning and emotional regulation in the attention deficit hyperactivity disorder subtypes. *J Clin Child Psychol* 29(1): 30-42
82. Kret ME, De Gelder B (2012) A review on sex differences in processing emotional signals. *Neuropsychologia* 50(7): 1211-1221

PARTIE EXPERIMENTALE

Certains résultats mis en évidence dans le cadre de cette recherche ont fait l'objet de communications :

Actes de colloques

- Leroy, A., Faure, S., & Spotorno S.** (2017). Interplay between semantic and emotional information in visual scene processing. *39th Annual Meeting of the Cognitive Science Society*, Londres (Royaume-Uni), 26-29 juillet. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Computational Foundations of Cognition*, 3761.
- Leroy, A., Faure, S., & Spotorno, S.** (2016). Integration of context and object semantic representations during rapid categorisation within and between the cerebral hemispheres. *39th ECVF, European Conference on Visual Perception*, Barcelone (Espagne), 29 août – 1er septembre. *Perception 45 ECVF Abstract Suppl.*, 163.

Communications orales

- Leroy, A., Faure, S., Fossoud, C., & Spotorno, S.** (2018). Traitement explicite et implicite du contenu émotionnel d'images présentées rapidement chez des enfants TDA/H. *5^{ème} colloque international de langue française sur le TDA/H*, 28 juin, Lausanne, Suisse.
- Leroy, A., Faure, S., & Spotorno, S.** (2017). Influence du contenu sémantique et émotionnel dans la catégorisation de scènes visuelles complexes présentées rapidement. *Journées de printemps de la Société de Neuropsychologie de Langue Française*, 2 juin, Grenoble, France.
- Spotorno, S., Leroy, A., & Faure, S.** (2016). Hemispheric contributions and context-object associations during categorisation of natural images. *International meeting of the Psychonomic Society*, 8 mai, Grenade, Espagne.

Communications affichées

- Leroy, A., Zeghari, R., Faure, S., & Spotorno, S.** (2017). Information selection and task relevance in emotional and neutral images: A comparison between oculomotor behaviour and subjective judgements. *20^{ème} conférence de la Société Européenne de Psychologie Cognitive (ESCoP)*, 3-6 septembre, Potsdam, Allemagne.
- Leroy, A., Faure, S., & Spotorno, S.** (2017). Mémoire à long terme d'images visuelles perçues brièvement : influence du contenu émotionnel et des niveaux global et local. *58^{ème} Congrès annuel de la Société Française de Psychologie*, 31 août, Nice, France.
- Leroy, A., Spotorno, S., & Faure, S.** (2016). Hemispheric Contributions During Categorisation of Natural Images. *Joint meeting of the British and French Neuropsychological Societies*, 16 mars, Londres, Royaume-Uni.
- Leroy, A., Spotorno, S., & Faure, S.** (2015). When two images make a scene: semantic and hemispheric contributions to the categorisation of visual contexts and objects. *Forum GDR Vision 2015*, 2-3 décembre, Grenoble, France.

Chapitre IV. Synthèse théorique, problématique et méthodologie générale

Ce chapitre propose une synthèse des principales données de la littérature présentées dans les trois revues de la littérature (articles I, II et III) de la partie théorique. Au regard de ces données, nous formulerons notre problématique de recherche et nos objectifs principaux, puis nous décrirons la méthodologie générale utilisée afin de tester ces objectifs.

IV.1 Synthèse théorique

L'analyse de la littérature présentée, grâce aux revues critiques, dans la partie théorique de ce document a permis de faire un tour d'horizon des travaux majeurs, mais également des axes de recherche délaissés et de certains débats toujours d'actualité sur le traitement des informations sémantiques et émotionnelles lors de la perception visuelle chez l'adulte (article I) et l'enfant (article II) au développement typique, puis dans le cas plus spécifique du TDA/H (article III).

Nous allons résumer succinctement les résultats majeurs qui ressortent de ces revues de littérature afin de nous permettre d'élaborer notre problématique.

- 1) L'être humain adulte a la capacité de catégoriser sémantiquement en 150 ms un objet ou un contexte présenté très rapidement (e.g., Thorpe et al., 1996). En outre, il a été montré que la mise en commun des informations du niveau local et du niveau global influence la vitesse et l'exactitude de catégorisation de ces deux niveaux. En effet, lorsqu'un objet et un contexte sont présentés en situation de cohérence sémantique (i.e., selon une association fortement cooccurrence comme un cochon et une ferme), les performances de catégorisation sémantique de l'objet et du contexte sont supérieures à celles obtenues lors d'une présentation en situation d'incohérence sémantique (e.g., un cochon et une banque). Cependant, la façon dont cet effet, dit de cohérence sémantique, se réalise et, plus particulièrement, la question du stade du traitement au sein duquel les informations visuelles qui proviennent du niveau local et du niveau global interagissent animent toujours des débats conceptuels.
- 2) Concernant les traitements émotionnels, la littérature met en évidence un effet de cohérence, cette fois-ci effet facilitateur de la cohérence émotionnelle, entre l'objet et le contexte de la scène. La présentation en situation de cohérence émotionnelle (i.e., lorsque les objets et le contexte de la scène sont porteurs de la même catégorie émotionnelle) facilite la reconnaissance de l'émotion par rapport à une situation d'incohérence émotionnelle (Righart & de Gelder, 2008). Toutefois, de manière

similaire aux questionnements sur les associations sémantiques, cette influence des liens émotionnels entre objet et contexte reste peu comprise. Notamment, le stade auquel se réalise l'intégration des traitements locaux et globaux, ici émotionnels, est toujours débattu. De plus, alors qu'il a été montré que le contenu émotionnel du contexte aide la catégorisation de l'émotion portée par l'objet, la réciproque, c'est-à-dire l'influence du contenu émotionnel de l'objet sur la catégorisation émotionnelle du contexte d'une scène visuelle, n'a pas encore été, à notre connaissance, explorée.

- 3) Un autre débat théorique majeur relevé par nos revues de littérature est relatif à la question des liens entre les traitements sémantiques et émotionnels. Pour certains auteurs, une réponse émotionnelle peut se produire directement sans nécessiter une évaluation cognitive délibérée et consciente, de façon rapide et automatique (e.g., LeDoux, 2003; Zajonc, 1984). Ainsi, face à un danger potentiel, nous ressentons automatiquement de la peur, puis dans un second temps nous catégorisons sémantiquement l'objet dangereux. Pour d'autres à l'inverse, l'évaluation cognitive d'une situation ou d'un objet est une précondition nécessaire à l'évaluation de l'émotion : nous ressentons de la peur parce que nous avons reconnu (et donc catégorisé sémantiquement) l'objet dangereux (e.g., Lazarus, 1984; Storbeck et al., 2006). Dans cette approche c'est l'interprétation que nous faisons du stimulus qui lui donne une valeur émotionnelle ou une autre (Sander, Grandjean, & Scherer, 2005). De nombreux travaux comportementaux et en neuroimagerie sont actuellement en faveur d'interactions constantes entre les traitements cognitifs et émotionnels mais dont les modalités restent néanmoins à préciser.
- 4) Le développement de la prise en compte et des influences réciproques entre les contenus sémantiques des objets et contextes des scènes visuelles a été peu étudié. Les quelques travaux réalisés suggèrent néanmoins la présence tôt dans le développement de l'enfant (4 mois) de la prise en compte de l'ensemble des informations d'une scène visuelle (Bornstein, Mash, & Arterberry, 2011) et un rôle essentiel des expériences dans le développement des effets d'association sémantique (Thomas, Nardini, & Mareschal, 2017). Les capacités de catégorisation visuelle rapide chez l'enfant, et notamment chez l'enfant d'âge scolaire, et chez l'adolescent n'ont pratiquement pas été explorées, malgré leur importance dans les processus d'apprentissage et dans l'adaptation à l'environnement physique et social. En outre, la période de scolarisation en primaire est la période durant laquelle se réalise le passage d'un traitement préférentiel du niveau

local à un traitement prédominant du niveau global (Poirel, Mellet, Houdé, & Pineau, 2008). Les traitements local/global ont été majoritairement étudiés avec des stimuli hiérarchiques composés par des figures géométriques ou des lettres. Leur étude à l'aide de stimuli plus complexes et représentatifs d'environnements réels semble maintenant nécessaire pour appréhender la façon dont l'enfant comprend son monde visuel.

- 5) Au niveau du traitement des informations émotionnelles, la détection automatique et non volontaire du contenu émotionnel a été mise en évidence dès la première année de vie et semble donc mature très tôt dans le développement (Jessen, Altvater-Mackensen, & Grossmann, 2016a). En revanche, le développement de la capacité à catégoriser volontairement une image selon une valence émotionnelle (i.e., positive ou négative) ou une émotion de base (e.g., joie, colère, peur, tristesse, dégoût) paraît se poursuivre jusqu'à la fin de l'adolescence. La capacité à détecter une incohérence émotionnelle entre l'émotion portée par un visage et une attitude corporelle a été mise en avant dès la première année de vie de l'enfant (Heck, Chroust, White, Jubran, & Bhatt, 2018; Rajhans, Jessen, Missana, & Grossmann, 2016). Il a ensuite été décrit, dès l'âge de cinq ans, que la performance de catégorisation de l'expression émotionnelle d'un visage est supérieure lorsque celui-ci est présenté avec un contexte ou un corps dans une attitude physique qui véhicule le même contenu émotionnel par rapport à un contenu émotionnel différent (Mondloch, 2012; Theurel et al., 2016). Comme chez l'adulte, la réciproque d'une influence d'un niveau local sur le jugement émotionnel d'un niveau global n'a pas, à notre connaissance, été étudiée.
- 6) Enfin, certaines psychopathologies et troubles neuro-développementaux sont caractérisés par des difficultés émotionnelles, qu'il s'agisse de déficits de reconnaissance et/ou de gestion des émotions. Le TDA/H est un trouble neuro-développemental caractérisé par différents dysfonctionnements, notamment des capacités attentionnelles et des fonctions exécutives. Ce trouble se manifeste également par une aversion au délai et des difficultés socio-affectives qui seraient, pour certains auteurs, en partie liées à un déficit de reconnaissance des émotions (Martel, 2009; Shaw, Stringaris, Nigg, & Leibenluft, 2014; van Stralen, 2016). Cependant, comme nous l'avons décrit dans notre troisième revue de littérature (article III), la majorité des études utilisent des visages émotionnels isolés comme stimuli, ce qui est important pour étudier l'identification des expressions faciales émotionnelles, ainsi que des aspects importants des relations sociales, mais qui ne suffit pas à caractériser l'ensemble des informations

émotionnelles que l'individu doit traiter dans son quotidien. De plus, très peu d'études ont analysé les capacités de traitement sémantique des informations visuelles auprès de cette population, ce qui peut rendre difficile l'interprétation des capacités de traitement des informations émotionnelles, telles qu'approchées au moyen d'épreuves de catégorisation émotionnelle de stimuli visuels, au moins si l'on se réfère à l'hypothèse de traitements émotionnels secondaires à des traitements sémantiques (e.g., Lazarus, 1984).

IV.2 Problématique

L'objectif général de cette thèse est de renforcer la compréhension des processus à l'œuvre lors de la perception visuelle rapide de scènes complexes, durant leur édification ainsi qu'à leur état mature. Le traitement rapide des informations visuelles est une capacité indispensable à l'adaptation et à la survie : l'être humain doit pouvoir réagir rapidement lors de la survenue d'un nouvel élément dans son environnement visuel, mais l'importance de cet élément doit être évaluée. En effet, le traitement de tous les stimuli environnants surchargerait les ressources cognitives et ne permettrait pas de poursuivre une tâche en cours. Un défaut de traitement de l'importance des stimuli est décrit dans le TDA/H : notamment il a été suggéré qu'un défaut d'attention soutenue, ou du contrôle inhibiteur, altérerait la hiérarchisation de l'importance des stimuli.

Notre problématique générale porte sur la question de l'existence d'interactions entre les traitements sémantiques et émotionnels lors de la présentation rapide d'une image d'objet et d'une image de contexte, hors de la vision fovéale, c'est-à-dire hors de la vision centrale où l'acuité visuelle est la plus fine. Notamment, il a été décrit une influence mutuelle des contenus sémantiques de l'objet et du contexte d'une scène visuelle sur leur catégorisation sémantique respective et une influence du contenu émotionnel du contexte sur la catégorisation émotionnelle de l'objet, mais nous ne savons pas si les associations objet contexte concernant les dimensions sémantiques et émotionnelles interagissent entre elles.

Comme certains auteurs le suggèrent (Lai, Hagoort, & Casasanto, 2012; Mueller, Fritsch, Hofmann, & Kuchinke, 2017), nous considérons que les questions de la dépendance entre les traitements émotionnels et les traitements sémantiques, et de la primauté d'un traitement par rapport à l'autre n'ont pas de réponse simple, notamment au regard des différences entre les traitements volontaires et non volontaires du contenu émotionnel. Nous faisons alors l'hypothèse générale que lors de l'apparition d'une scène ou d'un nouvel élément

dans l'environnement visuel, des traitements en parallèle des contenus émotionnel et sémantique se déclenchent et interagissent pour améliorer l'efficacité de la prise d'informations et, ainsi, l'efficacité de la perception visuelle. Ces traitements sont permis par des processus attentionnels, plutôt endogènes lorsqu'ils concernent les intentions de l'observateur, plutôt exogènes lorsqu'ils dépendent des caractéristiques du stimulus. Ces traitements initiaux, approximatifs, permettent l'évaluation rapide de la pertinence des stimuli pour l'individu, l'informent sur son environnement et permettent ainsi la formation de prédictions grâce aux représentations stockées en mémoire à long terme (Bar, 2007). Ils sont la base de l'efficacité de la perception visuelle, ils permettent de comprendre l'environnement physique et social et de s'y adapter rapidement. Les informations qui auront été jugées pertinentes en regard de l'activité en cours, des intentions ou intérêts de l'individu, seront ensuite traitées de façon plus détaillée pour permettre notamment la prise de décision.

Afin d'étudier cette problématique générale, nous la déclinons en différents sous-objectifs.

Premièrement, il s'agit d'identifier les mécanismes de traitement des associations sémantiques objet/contexte au sein des scènes visuelles.

Plusieurs propositions théoriques ont été avancées pour expliquer ces relations entre objet et contexte au sein des scènes visuelles : elles peuvent être regroupées dans trois modèles principaux qui se situent le long d'un continuum de la prise d'information visuelle jusqu'au stade de la prise de décision. Notre première étude empirique (article IV, chapitre V) a pour objectif d'éprouver ces différentes propositions théoriques afin de spécifier le stade de mise en commun des informations locales (objets) et globales (contextes) lors du traitement d'une scène visuelle complexe. Pour ce faire, nous analyserons si la représentation unifiée, générale d'une scène est maintenue lorsque ses constituants (objet et contexte) sont présentés simultanément mais séparément et ne font donc pas partie d'un unique percept. Grâce à cette modalité de présentation (cf. méthodologie générale de la thèse, chapitre IV, et celle de l'article IV pour plus de précision), nous pourrions tester l'hypothèse selon laquelle l'effet de cohérence sémantique dépend de l'étape de mise en relation d'un percept avec une représentation en mémoire à long terme. Cette hypothèse sera également éprouvée dans notre troisième étude empirique (article VI, chapitre VII) réalisée auprès d'enfants d'âge scolaire. Si l'effet de cohérence sémantique dépend des cooccurrences fréquentes entre les différents constituants

d'une scène visuelle, cet effet devrait émerger et se renforcer au fur et à mesure des expériences de vie.

Deuxièmement, deux autres études ont pour objectif de préciser les interactions entre les contenus émotionnels portés par l'objet et ceux portés par le contexte des scènes visuelles.

Les études présentées dans la partie théorique mettent en avant un effet facilitateur des informations contextuelles sur la catégorisation émotionnelle du niveau local (principalement visage). Il a également été montré une influence de l'émotion faciale (niveau local) sur la catégorisation émotionnelle d'une posture corporelle (niveau global). Cependant, à notre connaissance, l'influence du contenu émotionnel de l'objet sur la catégorisation émotionnelle du contexte dans le sens d'arrière-plan d'une scène visuelle (e.g., une forêt, une ville...), comme nous l'entendons dans ce travail de thèse, n'a pas été étudiée. Nous souhaitons tester l'hypothèse selon laquelle l'effet du contenu émotionnel du contexte sur la catégorisation émotionnelle de l'objet peut être généralisé à une influence mutuelle objet-contexte telle que celle observée dans l'effet de cohérence sémantique. Dans notre seconde étude empirique (article V, chapitre VI) chez l'adulte et dans notre troisième étude empirique (article VI, chapitre VII) chez l'enfant, nous analyserons ainsi l'effet de la présence d'un contexte émotionnellement cohérent sur la catégorisation du contenu émotionnel de l'objet, mais également l'effet du contenu émotionnel de l'objet sur la catégorisation émotionnelle du contexte. Enfin, afin de tester si les effets de cohérence émotionnelle et de cohérence sémantique entre les deux images – un objet et un contexte – présentées à chaque essai interagissent, nous manipulerons de façon orthogonale dans notre seconde étude (article V, chapitre VI) la cohérence sémantique et la cohérence émotionnelle entre l'objet et le contexte, de manière identique lors d'une tâche de catégorisation sémantique et une tâche de catégorisation du contenu émotionnel.

Troisièmement, il s'agira de différencier les traitements volontaires et non volontaires des dimensions sémantiques et émotionnelles.

Comme nous l'avons vu dans notre partie théorique, la question des relations entre les traitements sémantiques et émotionnels, et d'une primauté affective ou sémantique, est toujours en débat. Nous faisons l'hypothèse que ces traitements interagissent et que leurs influences mutuelles s'expriment différemment selon la dimension de traitement sur laquelle l'attention de l'observateur est orientée, sémantique ou émotionnelle. Pour tester cette hypothèse, nous

proposons dans notre seconde étude empirique (article V, chapitre VI), une tâche de catégorisation sémantique et une tâche de catégorisation émotionnelle orientant l'attention de l'observateur sur les informations sémantiques ou sur les informations émotionnelles respectivement, tout en manipulant dans chacune la dimension complémentaire. Nous souhaitons ainsi étudier l'influence du traitement non volontaire d'une dimension, car non pertinente pour la tâche, sur le traitement volontaire de l'autre dimension pertinente pour la tâche, sur laquelle le focus attentionnel de l'observateur est orienté.

Quatrièmement, il s'agira d'examiner s'il existe un déficit du traitement précoce du contenu émotionnel dans le TDA/H.

Pour cela, nous analysons les capacités d'enfants qui présentent un TDA/H à traiter le contenu émotionnel de scènes visuelles présentées rapidement. Au-delà du TDA/H, nous souhaitons avec cet axe de recherche déterminer l'intérêt d'utiliser une présentation rapide des scènes visuelles complexes pour spécifier les déficits dans certains troubles neurodéveloppementaux. Ainsi, dans notre dernière étude empirique (chapitre VIII) nous comparons les performances d'enfants qui présentent un TDA/H à celles d'enfants au développement typique en examinant l'influence du traitement non volontaire du contenu émotionnel de scènes visuelles sur les performances de catégorisation sémantique. Nous comparons aussi les deux groupes pour ce qui concerne les performances à la tâche de catégorisation du contenu émotionnel afin d'approcher le traitement volontaire du contenu émotionnel. Ceci nous permettra d'inférer si le traitement précoce des informations émotionnelles est altéré dans ce trouble.

IV.3 Méthodologie générale

IV.3.1 Considérations générales pour la partie expérimentale

Les travaux réalisés dans le cadre de cette partie expérimentale sont présentés sous format article, la méthodologie spécifique de chaque étude est ainsi décrite précisément dans l'article correspondant. Pour éviter des répétitions, nous ne présentons ici que certains aspects généraux concernant l'ensemble des études expérimentales. Chaque article est inclus dans un chapitre et dans chacun nous reprenons succinctement les points théoriques majeurs à l'origine de la problématique, nous présentons l'étude réalisée, nous incluons l'article, ou le projet d'article correspondant et enfin nous discutons brièvement les résultats. L'ensemble des résultats de cette thèse est repris plus en détail dans la discussion générale. Nous apportons dans

chaque chapitre, décrivant les études empiriques, seulement quelques précisions, principalement pour les études auprès d'enfants (au développement typique et présentant un TDA/H) mais nous ne reprenons pas l'ensemble des descriptions méthodologiques.

IV.3.2 Présentation des images en champ visuel divisé

Bien que nous n'étudions pas la spécialisation hémisphérique¹, nous utilisons dans les études expérimentales (chapitres V à VIII) de cette thèse le paradigme de présentation en champ visuel divisé (CVD) afin de contrôler les entrées perceptives reçues par le participant. Cette méthode s'appuie sur l'organisation anatomo-fonctionnelle croisée du système visuel humain (cf. Dimond & Beaumont, 1972). Ainsi, un stimulus présenté dans le champ visuel droit de l'observateur sera reçu initialement par l'hémisphère cérébral controlatéral, ici le gauche, et inversement, la stimulation dans le champ visuel gauche sera initialement reçue par l'hémisphère droit. Lors d'une expérimentation en CVD, l'observateur est face à un écran sur lequel apparaissent des stimuli latéralisés à plus de 2,5 degrés d'angle visuel à droite ou à gauche d'un point de fixation central, afin que le stimulus n'apparaisse pas en zone fovéale. Cette technique utilise donc une contrainte spatiale mais également une contrainte temporelle : le stimulus doit apparaître moins de 200 ms afin que le participant n'ait pas le temps de placer le stimulus en vision fovéale au moyen d'une saccade oculaire (cf. Bourne, 2006). Les contraintes spatiales et temporelles de la présentation en CVD nous permettent de nous assurer que les participants perçoivent de manière équivalente les deux images : par l'emplacement où est affiché le stimulus sur l'écran et par le contrôle de son temps de présentation. En outre, en plus de ces contraintes spatio-temporelles liées à la présentation en CVD, nous utilisons un masque afin de supprimer les persistances rétinienne.

IV.3.3 Tâches de catégorisation sémantique et de catégorisation émotionnelle

Afin de distinguer si des mécanismes différents sont impliqués dans les traitements des contenus sémantique et émotionnel selon que le traitement soit volontaire (processus *top-down*) ou non volontaire (processus *bottom-up*), nous proposons deux tâches différentes (en inter sujet) dans nos études empiriques portant sur les traitements sémantique et émotionnel (article V/chapitre VI, article VI/ chapitre VII et chapitre VIII). Selon les tâches expérimentales, par

¹ Les travaux présentés dans cette thèse n'intègrent pas l'étude de la spécialisation hémisphérique. Cependant l'influence du champ visuel de présentation a été initialement analysée dans les expériences 1 et 2 présentées dans la première étude empirique de cette thèse (article IV) et les résultats ont été présentés dans des communications orales et écrites (voir p.129). Cette dimension du traitement des scènes visuelles n'a ensuite plus été étudiée.

les consignes, nous focalisons différemment l'attention des participants afin que le contenu sémantique et le contenu émotionnel soient pertinents ou non.

La tâche de catégorisation sémantique est ainsi proposée afin d'analyser les capacités de catégorisation sémantique des composants d'une scène visuelle (contexte et objet), mais également afin d'étudier le traitement non volontaire du contenu émotionnel puisque celui-ci est non pertinent pour la tâche. Dans cette tâche, la consigne est de catégoriser sémantiquement l'image, à deux niveaux : le contexte, en tant que manufacturé ou naturel, pour la moitié des participant ; l'objet, en tant qu'animal, être humain ou objet inanimé, pour l'autre moitié. L'attention du participant est donc orientée sur le contenu sémantique de l'image ; si un effet du contenu émotionnel est mis en évidence, celui-ci suggérerait un traitement non volontaire du contenu émotionnel.

La tâche de catégorisation émotionnelle nécessite un traitement volontaire du contenu émotionnel : la consigne est de catégoriser le contenu émotionnel d'une image (contexte pour la moitié des participant, objet pour l'autre moitié) en tant que neutre, positif ou négatif. L'attention du participant est donc cette fois orientée vers le contenu émotionnel de l'image qui est pertinent pour la réalisation de la tâche ; si un effet du contenu sémantique est mis en évidence, celui-ci suggérerait un traitement non volontaire du contenu sémantique.

Chapitre V : Etude du rôle des liens sémantiques entre l'objet et le contexte lors de la perception rapide de scènes visuelles complexes

Ce chapitre présente la première étude empirique de ce travail de thèse. Il est constitué d'un rappel des principaux apports théoriques concernant les associations sémantiques objet-contexte au sein des scènes visuelles, d'une présentation de la méthodologie de l'étude, de son résumé en français, de l'article en cours de révision pour publication et enfin d'une discussion intermédiaire.

Leroy, A., Faure, S., & Spotorno, S. (in revision). When two images make a scene: Reciprocal semantic predictions drive categorization of contexts and objects even when they are separate. *Scientific Reports*.

V.1 Rappels des apports théoriques principaux

Plusieurs propositions théoriques ont été formulées pour expliquer le rôle des relations entre objet et contexte au sein des scènes visuelles, hypothèses qui se regroupent au sein de trois modèles principaux : le modèle du schéma perceptif (« perceptual schema model »), le modèle de sélection de l'objet ou modèle d'amorçage ou modèle de correspondance (« object selection model », « priming model », ou « matching model ») et le modèle d'isolation fonctionnelle (« functional isolation model »). Ces modèles se situent le long d'un continuum de la prise d'information visuelle et, ainsi, du traitement perceptif de ces informations, à l'accès aux représentations catégorielles de l'objet (i.e., sa « catégorisation », donc sa mise en correspondance avec une catégorie connue, et ainsi l'accès aux connaissances sur sa nature, ses fonctionnalités ...) jusqu'au stade de la prise de décision (cf. Figure 1). La perception visuelle ne se réalise pas uniquement selon une organisation strictement sérielle et ascendante mais pour plus de clarté nous nous focaliserons dans ce chapitre sur les traitements ascendants.

Le modèle perceptif (Antes, 1977; Biederman, Mezzanotte, & Rabinowitz, 1982; Boyce, Pollatsek & Rayner, 1989) situe ainsi l'effet de cohérence sémantique à un stade initial du traitement des informations visuelles (de la transcription rétinienne à la description structurelle de l'objet). Les informations contextuelles, notamment fournies par les basses fréquences spatiales ou par les relations spatiales entre différents constituants de la scène, constitueraient un « schéma perceptif » de la scène qui influencerait le traitement perceptif des objets inclus dans cette scène. Ainsi le traitement perceptif des objets cohérents avec la scène visuelle serait facilité, alors que celui des objets incohérents serait parasité.

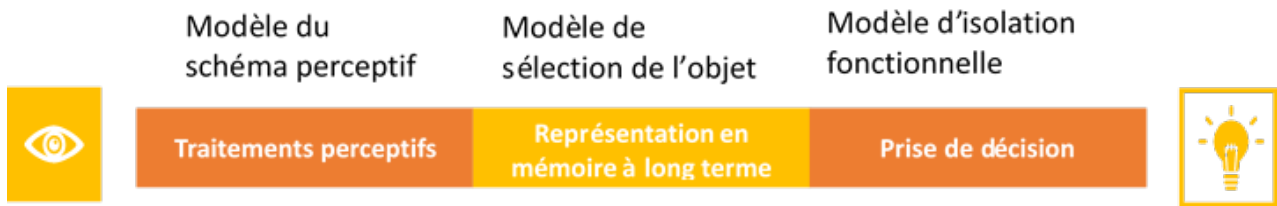


Figure 1: représentation schématique simplifiée des principaux modèles théoriques des associations objet/contexte

Ces modèles situent leurs hypothèses générales à différents stades du traitement des informations visuelles, sur un continuum de la prise d'information perceptive (à l'extrême gauche) jusqu'au stade décisionnel (à l'extrême droite).

Le modèle de sélection de l'objet et les modèles associés (d'amorçage ou de correspondance) situent les associations objet-contexte à un stade de plus haut niveau, c'est-à-dire dès qu'une correspondance est trouvée entre un percept et une représentation en mémoire à long terme (Bar, 2004; Palmer, 1975 ; Ullman, 1996). Pour ces modèles, dans un premier temps, un traitement rapide d'informations grossières, principalement contextuelles, se réalise et permet d'activer certaines représentations sur la scène formées préalablement et stockées en mémoire à long terme. Ces représentations faciliteraient ainsi la sélection des objets associés en diminuant la quantité d'informations perceptives nécessaires pour permettre la reconnaissance de l'objet cohérent sémantiquement dans cette scène (comparativement à un objet incohérent).

Enfin le dernier modèle, d'isolation fonctionnelle, défend une indépendance entre le traitement des informations contextuelles et locales au sein des scènes visuelles (De Graef, Christiaens & d'Ydewalle, 1990; Hamm, Johnson, & Kirk, 2002; Hollingworth & Henderson, 1998, 1999). Pour ce modèle, l'effet de cohérence sémantique n'existe pas à proprement parler, il serait en fait lié à des biais expérimentaux ou bien secondaire à des réponses devinées, imaginées, par les participants au moment de leur prise de décision, sans pour autant que cette décision se fasse de façon volontaire.

V.2 Présentation de l'étude

L'effet de cohérence sémantique entre l'objet et le contexte d'une scène visuelle, bien que souvent décrit dans la littérature sur la catégorisation visuelle rapide, reste néanmoins peu compris. Afin d'éprouver les prédictions issues des modèles explicatifs principaux décrits ci-dessus et d'améliorer la compréhension de l'effet de cohérence sémantique, nous avons créé

une tâche de catégorisation sémantique dans laquelle l'objet et le contexte d'une scène visuelle pouvaient être présentés ensemble, l'objet inclus dans le contexte (expérience 1), ou séparés en deux images distinctes. Dans ce dernier cas, une image représente l'objet inclus dans une configuration de bruit (i.e., « *pink noise* ») dont la distribution 1/f des fréquences spatiales reproduit les fréquences spatiales des scènes naturelles avec une surreprésentation des basses fréquences spatiales, l'autre image représente le contexte (expériences 2 à 5). Cinq expériences ont ainsi été réalisées afin d'étudier l'effet de la séparation et du positionnement spatial des images sur leur catégorisation. Les images de l'objet et du contexte pouvaient être présentées dans le même hémichamp visuel (Exp. 1 et 3) ou chacune dans un champ différent, sur la ligne horizontale médiane avec différentes excentricités (Exp. 2 et 5), ou enfin dans un quadrant en haut ou bas de l'écran (Exp. 3 et 4). Dans toutes les conditions expérimentales, la présentation des images se faisait rapidement (130 ms) et était latéralisée, c'est-à-dire hors de la vision fovéale, à droite et/ou à gauche du centre de l'écran. L'enregistrement des mouvements oculaires des participants à l'aide d'un oculomètre (EyeLink 1000, SR Research) a permis de nous assurer qu'ils fixaient le point central à chaque essai et qu'ainsi les deux images étaient vues de la même manière, toutes deux en vision extrafovéale. Pour une description plus précise de la méthodologie expérimentale, se référer à la partie « méthode générale » de l'article 4.

Cette méthodologie nous permet de tester l'hypothèse du modèle théorique « perceptif » selon laquelle l'effet de cohérence sémantique serait lié au schéma perceptif de la scène visuelle et donc nécessiterait que l'objet soit inclus dans le contexte. De plus, le fait de restreindre la quantité et la qualité des informations reçues avec la présentation des images à de plus grandes excentricités (expériences 3 à 5) et lors de présentations dans les champs visuels supérieur et inférieur (expériences 3 et 4) devrait également réduire voire annuler l'effet de cohérence sémantique. Ainsi, selon ce modèle théorique, nous devrions obtenir un effet de cohérence sémantique uniquement pour l'expérience 1 où l'objet et le contexte sont présentés ensemble dans une seule et même image.

Pour les modèles dits de sélection de l'objet ou d'amorçage et de représentation, la mise en commun des informations contextuelles avec celles des objets se réalise par une mise en correspondance du percept avec une représentation en mémoire à long terme et ne nécessite donc pas d'intégration perceptive de l'objet au sein du contexte de la scène. Selon ces modèles, nous devrions obtenir un effet de cohérence sémantique dans toutes les conditions de présentation des stimuli (Exp. 1 à 5).

Enfin, nous avons pris en compte les critiques énoncées par les auteurs du modèle d'isolation fonctionnelle sur la méthodologie de certaines études qui aurait conduit à des biais expérimentaux. Premièrement, nous présentons simultanément les images d'objet et de contexte, limitant le plus que possible l'intervention de l'amorçage sémantique. Deuxièmement, nous demandons aux participants de nommer les deux images qui leur sont présentées en répondant à l'oral et une variété de réponses est ainsi possible : les participants n'ont donc pas un choix limité de réponse, ce qui rend pratiquement impossible le fait de deviner une réponse correcte (e.g., si un participant a correctement reconnu un contexte, par exemple une ferme, même s'il donne au hasard une réponse pour l'objet, il a peu de chance de trouver la bonne réponse ; il peut ainsi répondre « une vache » ou « une chèvre » alors qu'il s'agissait d'un cochon). Troisièmement, nous avons enlevé de nos analyses statistiques toutes les réponses dites « d'intrusion », c'est-à-dire lorsque le participant donne une réponse erronée mais qui est cohérente sémantiquement avec la réponse attendue, comme dans l'exemple ci-dessus. En excluant ces éventuelles réponses devinées des analyses, nous réduisons le risque que de telles réponses biaisent les résultats. Il est néanmoins à noter que très peu d'intrusions sont commises (voir Tableau 1 dans l'article). Si l'hypothèse du modèle d'isolation fonctionnelle est valide, avec la mise en place de ces précautions méthodologiques, nous ne devrions pas obtenir d'effet de cohérence sémantique quelle que soit la version expérimentale.

V.3 Résumé de l'étude

Les performances de catégorisation visuelle sont améliorées lorsque les associations objet-contexte dans les scènes sont sémantiquement cohérentes, donc prévisibles, comparativement à des associations sémantiquement incohérentes. Cependant, il n'est pas encore bien établi à quel stade du traitement des informations visuelles ces associations se situent : à un niveau de traitement perceptif, dans l'appariement des représentations en mémoire ou dans les étapes ultérieures de la sélection des réponses. Nous avons testé ces trois propositions théoriques au cours de cinq expériences. A chaque essai, les participants devaient classer un contexte de scène et un objet présentés brièvement dans la même image (expérience 1), ou séparément, par la présentation simultanée de deux images (Expériences 2-5). Nous avons analysé les présentations unilatérales (expériences 1, 3) et bilatérales (expériences 2, 4, 5), et les présentations sur la ligne médiane horizontale de l'écran (expériences 1-2) et dans les champs visuels supérieur et inférieur (expériences 3, 4). Dans toutes les expériences, nous avons mis en évidence un avantage de la cohérence sémantique à la fois pour la catégorisation du contexte et

pour la catégorisation de l'objet. Ce résultat montre que la représentation en mémoire des associations sémantiques objet-contexte est activée que ces deux constituants de la scène soient ou non intégrés dans la même image. Ceci suggère que l'effet de facilitation de la cohérence sémantique sur la catégorisation se produit au stade cognitif de l'appariement d'un percept avec les connaissances antérieures stockées en mémoire à long terme. Cette étude est en faveur du modèle de la sélection de l'objet et étend ce cadre à une influence réciproque entre l'objet et le contexte sur les processus de mise en correspondance avec des représentations en mémoire à long terme.

V.4 Article IV. When two images make a scene: Reciprocal semantic predictions drive categorization of contexts and objects even when they are separate

Le texte ci-dessous est la reproduction d'un article soumis pour publication le 5 juillet 2019 à la revue *Scientific Reports*, en cours de révisions majeures depuis le 11 octobre 2019. Leroy, A., Faure, S., & Spotorno, S. (in revision). When two images make a scene: Reciprocal semantic predictions drive categorization of contexts and objects even when they are separate. *Scientific Reports*.

When two images make a scene: Reciprocal semantic predictions drive categorization of contexts and objects even when they are separate

Anaïs Leroy ^{a,*}, Sylvane Faure ^a, Sara Spotorno ^b

^a Université Côte d'Azur, LAPCOS, France.

^b School of Psychology, University of Keele, United Kingdom

*corresponding author:

Anaïs LEROY
LAPCOS, MSHS Sud Est
Université Côte d'Azur
Pôle Universitaire Saint Jean d'Angély
24 avenue des Diabes Bleus
06357 Nice Cedex 4 (France)
Telephone: +33 (0) 4 89 88 14 60
Email: anais.leroy@univ-cotedazur.fr

Abstract

Visual categorization improves when object-context associations in scenes are semantically consistent, thus predictable from schemas stored in long-term memory. However, it is unclear whether this is due to differences in perceptual processing, in matching of memory representations or in later stages of response selection. We tested these three concurrent explanations across five experiments. At each trial, participants had to categorize a scene context and an object briefly presented within the same image (Experiment 1), or separately in simultaneous images (Experiments 2-5). We analyzed unilateral (Experiments 1, 3) and bilateral presentations (Experiments 2, 4, 5), and presentations on the screen's horizontal midline (Experiments 1-2) and in the upper and lower visual fields (Experiments 3, 4). In all the experiments, we found a semantic consistency advantage for both context categorization and object categorization. This shows that the memory for object-context semantic associations is activated regardless of whether these two scene components are integrated in the same percept. Our study suggests that the facilitation effect of semantic consistency on categorization occurs at the cognitive stage of matching the percept with previous knowledge, supporting the object selection account and extending this framework to an object-context reciprocal influence on matching processes (object-context selection account).

Introduction

Viewers are able to categorize visual images from a brief glance, in a few dozens of milliseconds¹. This ability is a central dimension of cognition²: We make sense of the world fundamentally by relating and differentiating its components in terms of known groups, or categories^{3,4}. This allows us to process large amounts of information, while minimizing cognitive costs, and perceive common structures of elements⁴, even within the complex scenes that characterize everyday visual world. These scenes are composed of multiple objects, organized within a background (the scene's context) following semantic and spatial rules, which define the plausibility of object occurrence in the scene and of the scene's layout, respectively⁵. Object-context associations are at the core of the concept of "scene" and a key, omnipresent aspect in our visual experience. However, it is still unclear how they influence visual information processing, namely in which conditions and at what stage they operate.

Many studies have focused on the plausibility of object occurrence and have shown better categorization performance in the case of consistent, expected associations than inconsistent, unexpected associations. This semantic consistency advantage has been mainly found for categorization of the objects (e.g., "parasol", "notebook") included in the scene⁶⁻¹⁰ but also for the categorization of the scene's context^{7,11,12} in terms of its gist¹³ (e.g., "beach", "office").

Proposed explanations of the effect are contradictory, and situate it on a different time scale^{5,14-18}. Moreover, they focus on how the scene's context influences categorization of the embedded objects, without considering any impact of the objects on context categorization. Two accounts posit that we use predictions of likely (consistent) semantic associations to guide visual information gathering and that, therefore, object-context associations influence directly the categorization process. According to these two accounts, perception is cognitively penetrable¹⁹⁻²¹.

The perceptual schema model^{5,18,22} argues that semantic predictions impact the early stages of object perceptual analysis, facilitating processing of the local features that are likely to belong to expected objects in the given scene's context. This would facilitate categorization of those objects, while perhaps impairing categorization of inconsistent objects. These mechanisms are thought to be based on early inferences of the scene's gist based on coarse, low spatial frequency (LSF) information⁶, which would also enable activation of long-term knowledge related to the scene (schema¹³), containing indications about what objects to expect.

The object selection model (or priming model^{6,8,16,23-25}) also proposes that the categorization of the scene's context influences that of the embedded objects, but at a later stage. Instead of affecting their perceptual analysis, it would affect the viewer's criterion, i.e., how much perceptual evidence is needed to access the object's category by matching the perceived object with the object representation(s) selected in long-term memory. Semantically consistent object-context associations would facilitate object categorization by lowering the quantity of evidence required for a satisfactory matching.

A third account, the functional isolation model^{26,27}, claims that the scene's context has no effect on object categorization per se. Rather, contexts and objects are processed independently and integration of information takes place at a post-perceptual, decision-making stage of response selection. Any (apparent) facilitation of semantic consistency on categorization would come from an experimental bias promoting guessing responses. For instance, Biederman and colleagues⁵ presented the name of an object followed by a scene and asked participants to judge whether the object occurred in the scene. Hollingworth and Henderson²⁶ argued that cueing the target object's identity before the scene preactivates expectations and leads viewers to be more likely to respond positively about object occurrence within a consistent than an inconsistent context, and this even when the object is absent. More false alarms were indeed found in the consistent condition compared to the inconsistent condition in Biederman et al.'s. However, the two conditions when computing the sensitivity score (d') were pooled together, consequently overestimating the sensitivity rate in the case of consistency and underestimating it in the case of inconsistency. When Hollingworth and Henderson reran the study computing d' after separating false alarms for the consistent and inconsistent trials, or presenting first the scene and then two objects between which to choose the one included in the scene, no semantic consistency advantage was reported.

The general aim of the present study was to improve understanding of the object-context interplay in visual scene processing. For this purpose, we manipulated the semantic association (consistent versus inconsistent) between the scene's contexts and the objects (Fig.1) in five experiments using fast presentations and asking participants to categorize both the context and the object. Our main objective was to test predictions from the three theories described above, and to do so we introduced three fundamental aspects in our procedure (see General Method for more details).

First, we varied how object and context images were presented (Fig. 2). In Experiment 1, the objects were embedded into the scene's contexts, to set a baseline for the consistency

effect in our study. In all the other experiments, one object image and one context image were shown simultaneously but separately in different locations in the visual field, with the object embedded into a 1/f colored noise background (Fig.1, third column) reproducing the spatial frequency distribution of a natural scene. We used these separate and simultaneous presentations as a key manipulation to examine in what measure any semantic consistency effect requires low-level, perceptual relationships between the context and the object within the same image, or may arise exclusively from high-level guidance, independently from the actual and reciprocal organization of the context and the object in a unique percept. Moreover, and importantly, the simultaneous image presentation avoided any possible bias related to a priming effect (semantic preactivation) which may occur with sequential presentations (see also e.g.,^{7,17,25}). Previous research has reported a semantic consistency categorization advantage for two separate and simultaneous object images²⁸⁻³⁰, but no study has analyzed the semantic consistency effect between an object and a scene's context with such presentations.

Second, we did not constrain response within a preselected range of options. Consequently, the probability to give a correct guess response was arguably weak, as participants had to name the images choosing from a virtually unlimited set of labels at the entry-level category^{7,10}. Thus, even when they correctly recognized one of the images (e.g., a farm) and gave then a response for the other image guessing on the basis of a semantic association, they were quite unlikely to give the correct answer (e.g., they could respond “pig”, “goose”, etc., instead of the correct label “cow”).

Third, and again to minimize any impact of a guessing strategy, we removed from our analyses any trial with an intrusion response, which is an incorrect response given for one image and semantically coherent with the co-occurring image, for which the response has been correct⁷. For instance, when presented with the images of Figure 1, participants might correctly recognize the arena (context) and then guess the other image might be a toreador, while it is a bull or a laundry basket, or they might correctly recognize the bull and then guess that the context is a farm, while it is an arena or a bathroom. In addition, semantically associated responses might occur even when both the response for the context and that for the object are incorrect. Therefore, we also computed the number of consistent but both incorrect responses.

Different outcomes would support different theories.

According to the perceptual schema model, semantics influence the early stage of perceptual analysis (extraction of object's features from the context). Thus, we expected a semantic consistency effect only in Experiment 1, where the object was included into the

scene's context. Moreover, we expected the effect only for object categorization: this model does not easily account for any effect of object processing on context categorization, because the semantic effect would arise from gathering coarse information about the context, which would then affect processing of the local (object) information within the image.

According to the object selection model, the semantic consistency effect should occur in all our experiments and facilitate categorization of both the object and the context. In the classic version of this model, different possible perceptual representations of the same object are first accessed from coarse (contextual) information, and then the progressive accumulation of finer information enables selection of the best matching one³¹. However, we hypothesized that the same account may apply to the influence of objects on context categorization, as object processing may activate knowledge about the scene's context where the object is likely to appear⁶, lowering the viewer's criterion in matching the perceived context with a context representation in the case of semantic consistency.

Finally, according to the functional isolation model, we expected no semantic consistency effect because of the precautions (see above) taken in our study to limit the influence of a guessing strategy favoring performance in the consistent condition. Moreover, this model would predict many intrusion responses and many cases with two consistent but incorrect responses for the object-context pairs.

Besides providing a fundamental test for the theories concerning the impact of object-context semantic associations on visual categorization, our study offers other theoretical insights by examining categorization performance for both scenes' contexts and objects, and presenting the images lateralized in the left and/or right visual field (whereas previous research has used central presentations). In this way, we could analyze whether both context and object image levels can be quickly categorized without the contribution of foveal (central, high acuity³²) vision, and at different eccentricities (see General Method), therefore with limited involvement of detailed perceptual information (high spatial frequencies, HSF). Moreover, we could examine any impact of such extrafoveal presentations and of varying eccentricity on the semantic consistency effect for both contexts and objects.

Results

For each experiment, we ran a generalized linear mixed model (GLMM) on the proportion of correct responses with Semantic Association (Consistent, Inconsistent), Image Type (Context, Object) and their interaction as predictors, and participants and images as

random factors. All GLMMs had full random structure. For comparison between the experiments, we used independent-sample t tests. Prior to analysis, we removed the trials where participants gave incorrect responses classifiable as intrusions (Table 1). We present the results obtained after these removals, but we found the same patterns of results in each experiment also when including the trials with intrusions (see Supplementary Materials). Importantly, and despite the fact that we asked participants to categorize both the context and object, the amount of intrusions was overall very low ($M=2.82\%$, $SD=2.05$), suggesting little use of any strategy to guess semantically consistent responses.

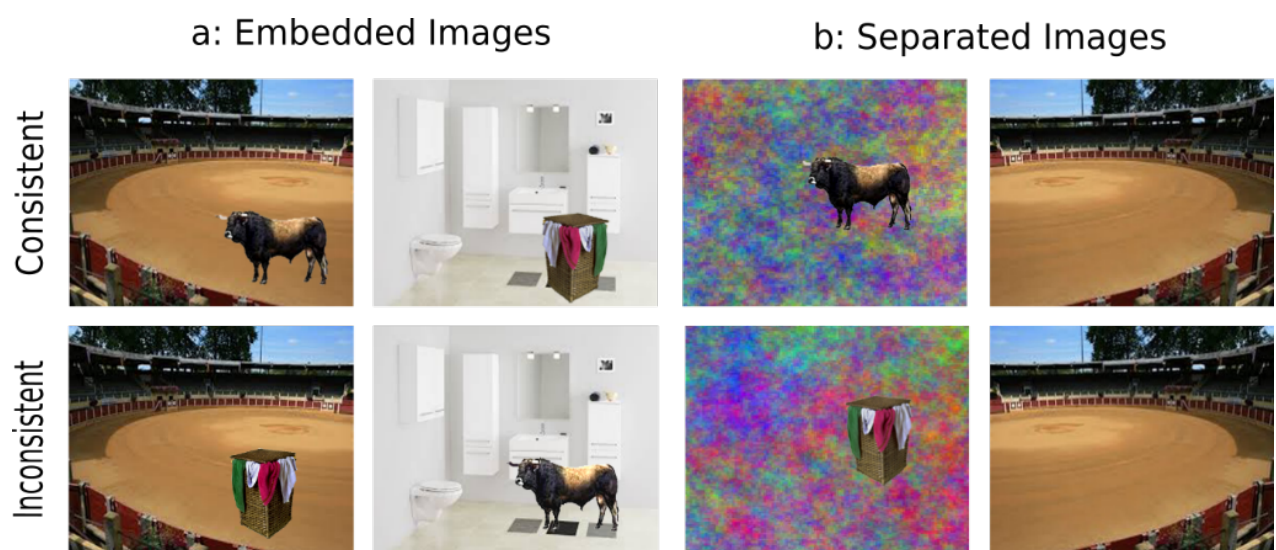


Figure 1: Examples of semantically consistent (top row) and inconsistent (bottom row) presentations, when either the object was included into the scene's context (a: Experiment 1) or the object and the context were shown simultaneously but separately, with the object pasted onto a $1/f$ colored noise background (b: Experiments 2-5).

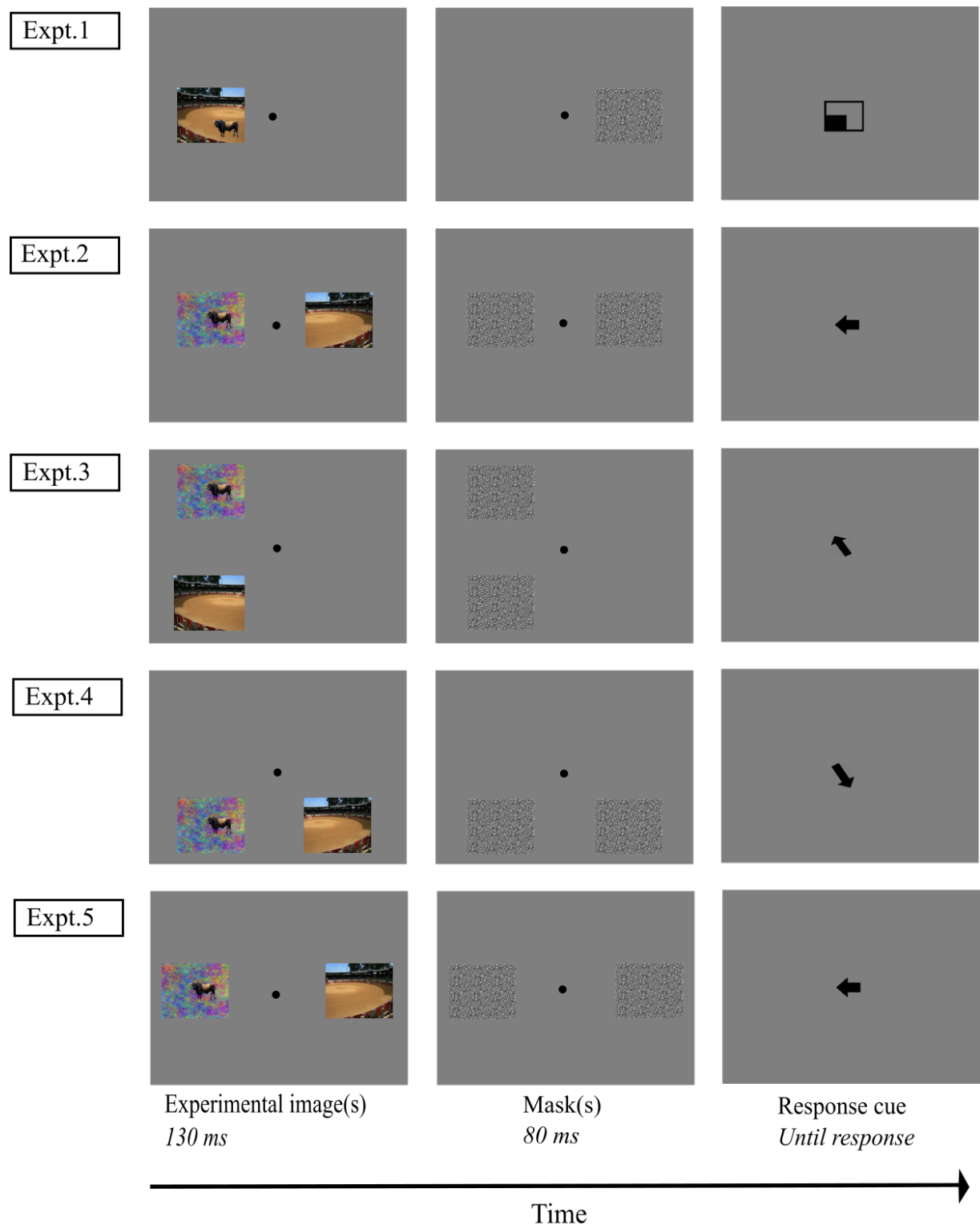


Figure 2: Time course of a trial in each experiment. The trial started with a single-point calibration check and a 90-ms black fixation cross (both in the center of the screen and here not depicted). Then, one image depicting an object included into a scene's context (Experiment 1) or one object image and one context image (Experiments 2-5) were briefly presented and then

masked. After this, a 150-ms medium gray blank screen was presented (here not depicted) and then a response cue (until response) indicated whether the participant had to categorize the object or the context first. Please note that the depictions here do not respect the original size proportions of stimuli presentation.

Number (%) of trials with:	Experiment					
	1	2	3	4	5	All
- no central fixation at image onset	35 (2.8)	45 (4.3)	24 (1.9)	58 (4.6)	34 (2.0)	196 (3.1)
- intrusion responses	144 (6.3)	79 (3.4)	18 (0.7)	29 (1.2)	76 (3.0)	346 (5.6)
- two incorrect but consistent responses	105 (4.1)	46 (1.8)	18 (0.7)	47 (1.8)	60 (2.3)	276 (2.2)

Table 1: Number and percentages (in brackets) of trials removed prior to analysis, for each experiment. Top row: trials with no central fixation at experimental image onset; percentages computed on the overall data. Middle row: trials with intrusion responses (i.e., with one incorrect response that was semantically consistent with the other, correct one); percentages computed on the data after the removals of trials with no central fixation at experimental image onset. Bottom row: trials with two incorrect responses that were semantically consistent with each other; percentages computed on the data after the removals of trials with no central fixation at experimental image onset.

Experiment 1

We presented the context and the object embedded in the same scene (Fig.2) to examine whether the semantic consistency advantage found in the literature for both context categorization and object categorization replicated with our set of stimuli and with lateralized presentations, and to establish a baseline for the consistency effect in our study.

We found a main effect of Semantic Association (Fig.3a), $\beta=-1.289$, $SE=0.227$, $z=5.69$, $p<0.001$, as images semantically consistent were categorized more accurately ($M=0.65$) than inconsistent ones ($M=0.46$), and of Image Type (Fig.3b), $\beta=-0.736$, $SE=0.258$, $z=-2.85$, $p=0.004$, as the context was categorized better ($M=0.62$) than the object ($M=0.51$). The interaction was not significant, $\beta=-0.015$, $SE=0.340$, $z<1$, $p=0.966$.

The advantage of semantic consistency for the categorization of both the context and the object (comparable for the two image types) is in agreement with several studies using central image presentation^{7,9-11}. We can, thus, conclude that our first experiment replicated previous findings and extended them to lateralized images perceived extrafoveally.

An important limitation of this approach is that having one level (the object) nested into the other (the context) does not allow us to disentangle, within the semantic consistency effect, the impact of low-level, perceptual relationships between the context and the object from that of high-level, cognitive guidance and representation. To examine this issue, Experiment 2 presented the context and the object separately, each in a different visual field with respect to the vertical midline (Fig.2).

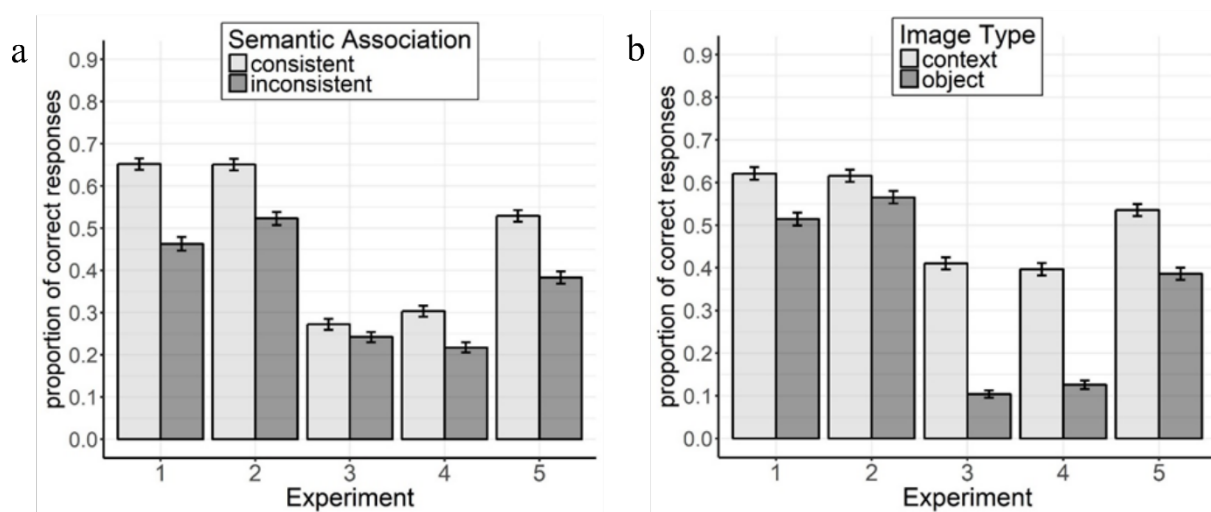


Figure 3: Categorization accuracy in Experiments 1-5 as a function of Semantic Association (Consistent vs. Inconsistent; a) and of Image Type (Context vs. Object; b). Error bars indicate one standard error around the mean.

Experiment 2

We found a main effect of Semantic Association, $\beta=-0.875$, $SE=0.228$, $z=-3.83$, $p<0.001$, (Fig.3a) as semantically consistent images were better categorized ($M=0.65$)

than inconsistent ones ($M=0.52$). The main effect of Image Type, $\beta=-0.375$, $SE=0.271$, $z=-1.38$, $p=0.166$, (Fig.3b) and the interaction, $\beta=-0.200$, $SE=0.431$, $z<1$, $p=0.642$, were not significant.

The consistency advantage clearly shows that viewers use semantic predictions about the likely semantic associations within scenes to guide categorization of both contexts and objects in a similar way, even when these two scene components do not perceptually belong to the same image. This facilitation, moreover, was comparable to when the context and the object belonged to the same image: the mean accuracy difference between semantically consistent and inconsistent trials did not significantly differ in Experiment 1 (0.19) vs. Experiment 2 (0.13), showing no significant difference, $t(30)=1.85$, $p=0.076$. These findings suggest that the influence of semantic consistency acts at a representational level, rather than an early perceptual level^{5,18,22}. As this aspect is crucial for our study, we will address it in detail in the General Discussion.

The absence of differences in categorization accuracy between context and object images seemed to arise from a slight performance increase for the objects in Experiment 2 ($M=0.57$) than in Experiment 1 ($M=0.51$), when they were included in the scene's context (although this difference was not significant, $t(30)=1.09$, $p=0.285$; Cohen's $d=0.039$), while accuracy for context categorization was virtually the same between these two experiments (both $M_s=0.62$, $t(30)<1$, $p=0.891$). Previous studies^{7,33} have reported an advantage for isolated objects compared to objects included within scene's contexts. However, in previous work the advantage might have arisen from easier figure/ground segregation, as the isolated objects were pasted into a blank background, while in our experiment the objects were embedded into 1/f noise backgrounds. We speculate that the relative facilitation of object categorization in Experiment 2 may rather stem from how we presented the simultaneous images. The short and separate presentation of the object and the context each in a different hemifield directed each image exclusively to the contralateral cerebral hemisphere³⁴.

Therefore, initial information processing was divided between the two hemispheres, instead of taxing the same hemisphere as in Experiment 1. Moreover, the advantage of separate, bilateral presentations is known to emerge mainly for the most difficult task condition^{35,36}, here object categorization.

To (a) generalize the semantic consistency effect between separate images to different views in the visual field and (b) examine the influence of image location in the visual field (same vs. different hemifield) on object categorization, Experiment 3 presented context and

object images separately, as Experiment 2, but within the same hemifield (left or right), thus both initially directed to the same hemisphere.

Experiment 3

We found a main effect of Semantic Association, $\beta=-0.379$, $SE=0.176$, $z=-2.15$, $p=0.031$, (Fig.3a) with better accuracy for consistent ($M=0.27$) than inconsistent ($M=0.24$) images, and of Image Type, $\beta=-2.599$, $SE=0.355$, $z=-7.31$, $p<0.001$, (Fig.3b) as contexts were better categorized ($M=0.41$) than objects ($M=0.10$). The interaction was not significant, $\beta=-0.506$, $SE=0.348$, $z=-1.46$, $p=0.145$.

The essential finding is that the advantage for semantically consistent object-context pairings that do not belong to the same perceptual configuration was replicated. Moreover, this effect was comparable for context and object images, as in Experiments 1 and 2, showing that it acts similarly when categorizing either global or more local image regions. Thus, we demonstrated that the consistency effect is not parochial to a specific image location, even though its amount may be modulated by the image arrangement in the visual field. We argue that the smaller difference between consistent and inconsistent trials compared to Experiment 2, $t(30)=4.72$, $p<0.001$ (and, by extension, Experiment 1) is likely to be due to a performance drop. Indeed, accuracy was much lower in Experiment 3 than 2 (and 1), and for both the context and (strikingly so) the object images (Fig.3b), both $ts(30)\leq 5.19$, $p<0.001$. There are two plausible explanations for this.

The first possibility arises from the unilateral presentation of two images that may cause an overload of the contralateral hemisphere, whereas in Experiment 1 only one image at the time was presented, and in Experiment 2 the two images were presented in an inter-hemifield condition, sharing the (initial) processing costs between the hemispheres. Previous research has suggested that independent, parallel processing of information in inter-hemifield presentations doubles visual short-term memory capacity compared to intra-hemifield presentations³⁷.

The second possibility stems from lower spatial resolution and contrast sensitivity, which may be due to the greater overall image eccentricity in Experiment 3 than Experiments 1-2 (see Procedure): While this was required to ensure lateralized processing, visual acuity decreases the further away from central vision³². Moreover, the horizontal-vertical anisotropy^{38,39} must be considered: Vertical eccentricities (greater in Experiment 3 than Experiments 1-2) are more disruptive for spatial resolution, contrast sensitivity⁴⁰, and information accumulation over time⁴¹ than horizontal eccentricities.

We then examined whether and how the semantic consistency effect between two separate images further generalizes to various arrangements in the visual field, while tackling in a systematic way the two possible explanations of intra- vs. inter-hemispheric processing and of decreased spatial resolution and contrast sensitivity. In Experiment 4, the two images were presented lateralized (as in Experiment 2) but both in the upper or in the lower visual field (as in Experiment 3). Comparing Experiments 3 and 4 allowed us to isolate the effect of intra- vs. inter-hemifield presentation, while keeping constant any other influence in our consistency vs. inconsistency categorization results.

Experiment 4

We found a main effect of Semantic Association, $\beta=-0.736$, $SE=0.164$, $z=-4.50$, $p<0.001$ (Fig.3a) with better accuracy for consistent ($M=0.30$) than inconsistent ($M=0.21$) images, and of Image Type, $\beta=-2.248$, $SE=0.316$, $z=-7.12$, $p<0.001$ (Fig. 3b) as contexts were better categorized ($M=0.40$) than objects ($M=0.13$). The interaction was not significant, $\beta=-0.235$, $SE=0.346$, $z<1$, $p=0.450$.

Apart from (crucially) finding again the advantage for consistent over inconsistent trials for both objects and contexts, as in the experiments 1-3, the results are very similar to those of Experiment 3 when considering accuracy for context images, $t(30)<1$, $p=0.687$, and object images, $t(30)<1$, $p=0.482$. This may suggest that the performance drop in Experiment 3 compared to Experiment 2 was mainly due to decreased spatial resolution and contrast sensitivity, rather than to the overload of a single hemisphere from the processing of two simultaneous images. Moreover, a categorization impairment due to decreased spatial resolution and contrast sensitivity would lead to a bigger impact on object categorization (involving more local, detailed information) than on context categorization (based mainly on global properties detectable from coarse processing), and this is exactly what we reported in Experiments 3-4.

The next and final experiment aimed to tease apart the role of overall image eccentricity from the screen's center and of horizontal-vertical anisotropy in the context vs. object effect and in any modulation of the influence of semantic consistency on categorization. To this purpose, Experiment 5 presented the two images centered on the horizontal midline, one to the left and the other to the right as in Experiment 2 (reducing any impact of vertical eccentricity), but with an overall image eccentricity that matched that in Experiments 3-4.

Experiment 5

We found a main effect of Semantic Association, $\beta=-0.944$, $SE=0.183$, $z=-5.14$, $p<0.001$ (Fig.3a), with an advantage for consistent ($M=0.53$) compared to inconsistent ($M=0.38$) images that was comparable to the one in Experiment 2, $t(30)<1$, $p=0.594$. We also found a main effect of Image Type, $\beta=-0.993$, $SE=0.259$, $z=-3.84$, $p<0.001$, with better categorization for contexts ($M=0.53$) than objects ($M=0.39$) (Fig. 3b). The interaction was not significant, $\beta=-0.307$, $SE=0.333$, $z<1$, $p=0.360$.

The replication of the semantic consistency advantage for both context and object images also in this experiment shows the robustness of our principal finding.

Experiment 5 differed from Experiment 2 only in the greater overall image eccentricity. Greater eccentricity should logically result in a steeper decrease in categorization accuracy for objects than contexts (finer vs. coarse processing, respectively), and we did show such a differential impact: Accuracy was lower in Experiment 5 than in Experiment 2 for objects, $t(30)=-3.62$, $p=0.001$, and only tended to be lower for contexts, $t(30)=-2.02$, $p=0.053$.

However, accuracy in Experiment 5 was higher than in Experiment 3 (+0.28 for objects, +0.12 for contexts; both $ts(30)\geq 3.65$, $p<0.001$) and in Experiment 4 (+0.26 for objects, +0.14 for contexts; both $ts(30)\geq 3.98$, $p<0.001$), which presented the images at the same overall eccentricity, but in the upper and lower quadrants. These improvements show the importance of differential modulation of categorization due to eccentricity in the vertical compared to the horizontal axis, especially for object categorization, for which, thus, the reduction in spatial resolution and contrast sensitivity associated with greater vertical eccentricity³² seems to have a bigger detrimental impact.

General Discussion

We studied rapid visual categorization comparing systematically performance for objects and scene's contexts. We focused on the role of object and context associations in scene processing, in terms of semantic expectations arising from previous experience regarding their co-occurrence. We tested predictions from the three principal theories (perceptual schema model, object selection model, functional isolation model) that have been used to account for the well-known, but still highly debated, finding of better categorization with semantically consistent (i.e., plausible) than inconsistent associations^{5,7-11}. To do so, in five experiments, we examined whether semantic consistency impacts categorization only when the context and object are integrated in a unique image, as it has been only studied, or even when the perceptual

scene configuration is not present in the visual field but only mentally constructed by the viewer.

Overall, showing a consistency advantage when an object image and a context image were presented simultaneously but separately, our study strongly suggests that a scene mental representation does not require a unified scene percept to influence categorization. We reported better categorization in the case of semantically consistent image pairs regardless of the specific image spatial arrangement in the visual field: When they were presented one in the left and one in the right hemifield, along the screen horizontal midline (at smaller, Experiment 2, or greater eccentricity, Experiment 5) or both in the upper or lower quadrants (Experiment 4), and when they were presented one in the upper and one in the lower quadrant in the same hemifield (left or right, Experiment 3). The replication of the consistency advantage, albeit reduced, even in Experiments 3 and 4, where performance was overall decreased with very low accuracy for object images, shows the robustness of our results.

Our study also shows that the object (local) level and the context (global) level influence each other^{7,11}, and that the learnt semantic associations between them impact their processing in a similar way. Indeed, we never found an interaction between the semantic consistency and type of image, object or context, categorized.

Both the influence of object-context semantic associations when the object is not embedded in the scene's context and the mutual interplay between the object and the context argue against the key assumption of the perceptual schema model^{5,18,22} that semantic expectations act directly on early perceptual processing of object features within the context.

Our results are also in disagreement with the functional isolation model^{26,27}, which proposes that any inconsistency effect is an epiphenomenon arising only during response selection, once all the stages of image categorization (perceptual processing and matching with a relevant long-term-memory representation) have been completed for the object and the context independently. Moreover, according to this model, the consistency advantage in response selection would be due to a guessing strategy promoted by an experimental bias²⁶. In our study, as responses were not restricted to a limited, preselected range of options, participants were unlikely to guess the correct answer. We also controlled for any guessing strategy by removing from analysis trials with intrusion (one response correct and the other incorrect but consistent). Few intrusions were made overall, and incorrect but semantically related responses for both the context and object images were rarely given (see Table 1): This suggests that our participants did not tend to guess.

The semantic consistency advantage we showed supports the object selection model^{6,8,23,24}, as it is compatible with the idea that object-context semantic associations influence the stage of matching the percept with previous knowledge and, therefore, with the representation of its category. Moreover, this model does not only account for the consistency effect on object categorization, but it may be extended to explain the reciprocal influence between objects and contexts we found, as the impact of semantic associations on the percept-to-representation matching mechanisms may apply to both the object and context levels. Inspired by the work of Bar and colleagues on object recognition within scene's contexts^{6,31,42}, we suggest that a first, coarse (LSF) percept is processed to a certain extent to allow partial recognition and used to generate a set of predictions about possible representations of the stimuli based on their general structure (object's silhouette, context's spatial statistics)⁴³. Then, viewers use knowledge of object-context co-occurrence within a type of scene to refine interactively these predictions and, finally, identify the stimuli with the subsequent integration of perceptual details (HSF)⁴⁴.

The overall superiority of context compared to object categorization in our study (except Experiment 2), especially with increased vertical eccentricity, is in agreement with a coarse-to-fine model of visual processing⁴⁵, where LSF, particularly important to recognize contexts, are processed earlier than HSF, which play a greater but not exclusive role in object recognition. However, both the finding that the objects were recognized in a considerable percentage of trials (approximately 30%-55%), at least when centered on the screen's horizontal midline, and the influence of semantic consistency on object categorization outside the high-acuity central vision (whatever the specific image location) are in agreement with previous research showing object recognition without fixation⁴⁶⁻⁴⁸. Our study is, thus, also relevant for a crucial debate in the literature, where several investigations have argued, instead, that recognition of an object, to a point that activates knowledge about the probability of occurrence in a given scene, would happen only in central vision (upon fixation)^{49,50}.

Finally, categorization of the objects and, to a lesser extent, the contexts varied considerably according to the image arrangement in the visual field, showing in particular a horizontal-vertical anisotropy^{38,39} with a huge performance decrease due to vertical eccentricity. This result highlights the need to carefully consider this dimension in any study of rapid visual categorization of objects and scene's contexts.

Conclusion

In five experiments, we showed that the facilitation of visual categorization of contexts and objects when they are semantically consistent compared to when they are inconsistent is a robust effect. Indeed, it emerges not only when the object is embedded in the context, belonging to the same image, but also when the object and the context are in separate images and in different locations in the visual field. Our results suggest that this effect is driven by percept-to-representation matching mechanisms, while it does not primarily act on processing of perceptual features (contrary to what the perceptual schema model claims) and does not mainly depend on guessing (contrary to what the functional isolation model claims). Thus, our study supports the object selection model and, finding a mutual interplay between the object and the context, extends it to explain categorization of the scene's context within the theoretical framework of an object-context reciprocal influence of matching processes (object-context selection account).

General Method

Participants.

Eighty individuals, 16 different in each experiment (Experiments 1-4: 8 females each, Experiment 5: 14 females), volunteered for no remuneration. All reported normal or corrected-to-normal visual acuity, normal color vision and no history of neurological disorders. Participants were aged 18 to 41 years old, with mean age (*SD*): 27.5 (6.7), 26.6 (5.1), 31.1 (5.6), 27.4 (5.6), 24.1 (4.4), for Experiments 1-5, respectively. All were right handed, with mean% of laterality quotient (*SD*) at the Edinburgh Laterality Inventory: 86 (7), 92 (7), 74 (31), 79 (24), 68 (31), for Experiments 1-5, respectively. This study was approved by the LAPCOS of the Université Côte d'Azur and carried out in accordance with the declaration of Helsinki. All participants gave informed written consent before participating.

Stimulus material and designs.

Eighty full-color photographs of animated and unanimated objects, plus six as practice, and 80 full-color photographs of indoor and outdoor scene's contexts, plus six as practice, were used. Objects were selected from Konkle et al.'s database⁵¹ or Google Images, contexts were selected from the SUN database⁵² or Google Images. "Object" images were defined by a unique object (e.g., bull, laundry basket). "Context" images were mainly defined by global, spatial information (e.g., arena, bathroom), even though they also contained local information. Eight

versions of each experimental scene used in Experiment 1 were created using Adobe Photoshop CS 9.0 (Adobe, San Jose, CA). First, a semantically consistent object was inserted within a context (e.g., the bull in the arena, the laundry basket in the bathroom), starting between 2° and 3° horizontally from the context's closest edge (the small variations of object placement were required to avoid spatial inconsistency within the context). Second, to create an inconsistent version of each scene, indoor and outdoor scenes were paired, and their consistent objects swapped (e.g., the bull was included in the bathroom and the laundry basket in the arena; Fig.1a). We repeated the first and second step inserting a mirror reversed object, to control for any effect of viewpoint typicality or eccentricity of its most diagnostic part (e.g., the handle in a cup). We then mirror reversed each scene to present it in each hemifield (left or right) with the object maintaining the same eccentricity from the central fixation point. For Experiments 2-5, the contexts and the objects were presented in separate images but paired and mirrored in the same way as in Experiment 1. Moreover, each object was pasted into a different colored 1/f noise background, created in Matlab R2015a (MathWorks, Inc., Natick, MA), which did not evoke any semantics. The objects had the same size as in Experiment 1. They were placed starting at 2° horizontally from the background's closest edge and centered vertically in the background. We also created, in Matlab, a white noise mask. All experimental images and the mask were shown at 10°x 8° (315 × 252 pixels). Moreover, we created visual response cues. In Experiment 1, the cue was a small white rectangle (3.24° x 2.76°, 102 x 87 pixels) with a black outline, where a black rectangular region indicated the scene's quadrant containing the object (this was done to prevent the participant to respond to another portion of the scene by mistake, thinking that it corresponded to the target object). In Experiments 2-5, the cue was a black arrow (2.5° x 1°, 79 x 32 pixels) pointing toward the screen's region where the image to categorize first had been presented.

Apparatus.

The experiments were conducted on a HP Compaq dc7100 computer running OS Windows XP. The stimuli were presented on a DELL 7004 CRT screen (32.5cm x 24.5cm, 1024 × 768 pixels, refresh rate: 85Hz). Stimulus presentation and response recording was controlled by Experiment Builder (SR Research, Canada). Eye movements were recorded using an EyeLink 1000 (SR Research, Canada) at a sampling rate of 1000Hz. Viewing was binocular, but only the dominant eye was tracked. A chinrest stabilized the eyes approximately 57cm away from the screen.

Procedure.

Participants were tested individually, seated in a quiet, dimly illuminated room, in front of the screen, and laying with their forehead and chin on the chinrest. They had to name both the context and the object as accurately as possible: Half had to name the context first, the other half the object first. They were instructed to be as specific as possible (e.g., “priest” rather than “person”, “lounge” rather than “house”), as in ⁷. Responses were given orally and transcribed by the experimenter. Prior to the experiment, participants underwent a randomized nine-point calibration procedure, which was validated to ensure that the average error was $< 0.5^\circ$ and the maximum error in one of the calibration points $< 1^\circ$. Recalibrations were performed during the task if necessary. Before each trial sequence, a single-point calibration check was applied while the participant fixated a dot in the center of the screen. The screen background color was medium grey (127,127,127). At the beginning of the experiment, participants performed six practice trials, followed by 80 experimental trials presented in random order. Each trial started with a 90-ms black central fixation cross (Times New Roman, bold, font size 40), followed by the experimental image(s) for 130ms, an 80-ms mask and a 150-ms medium grey blank screen. After this, a central cue indicated which item (i.e., whether the context or the object) the participant had to name first and was displayed until response. Examples of the presentations of the experimental scenes, mask and response cue in each experiment are depicted in Figure 2.

In Experiment 1 (object embedded in the context) and Experiment 2 (object and context presented separately; Figure 4) the images were centered on screen's horizontal midline. In Experiment 3, the object image and the context image were presented in the same hemifield one above the other (one in the upper left quadrant and one in the lower left quadrant, or one in the upper right quadrant and one in the lower right quadrant). All images (closest edges) appeared at an eccentricity from the screen's center of 2.5° horizontally and 2° vertically (therefore, the closest edge of each image was at 5.5° and the closest edge of the object to be categorized was at 7.5° overall eccentricity from the screen's center). In Experiment 4, the images were presented each in a different hemifield, both above or below the screen's horizontal midline (one in the left upper quadrant and one in the right upper quadrant, or one in the lower left quadrant and the other in the lower right quadrant), at the same eccentricity as in Experiment 3. In Experiment 5, the images were presented centered on the screen's horizontal midline, each in a different hemifield as in Experiment 2, but at a greater eccentricity (the

closest edge of each image was at 5.5° and the closest edge of the object to be categorized was at 7.5° from the screen's center), to have the same overall distance from the screen's center as in Experiments 3 and 4.

Each experimental image was presented only once during the experiment and all of the experimental factors were counterbalanced within- and between participants. Each experiment lasted approximately 40 minutes.

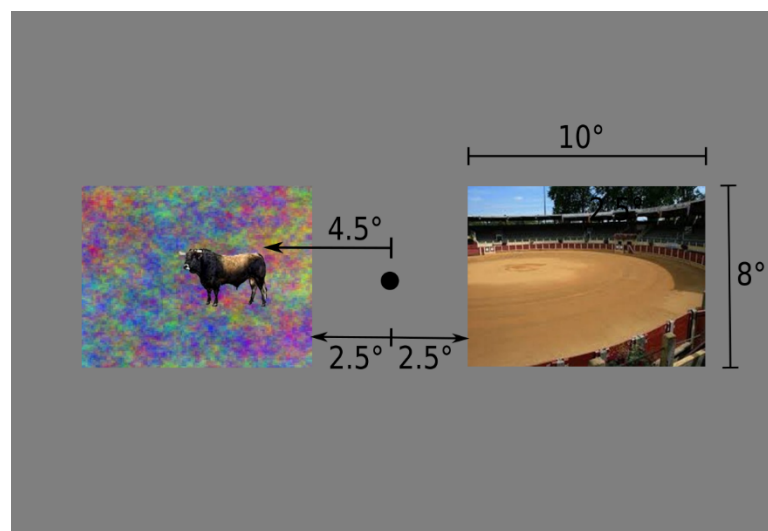


Figure 4: Image placement in Experiment 2. The images were placed at a horizontal eccentricity from the screen's center of 2.5° considering the closest edge of the image. Consequently, the closest edge of the object was placed at 4.5° of horizontal eccentricity from the screen's center. In Experiment 1 (object embedded in the context), only one image was presented in each trial, in the same location as one of the two images in Experiment 2.

Data analysis.

All results were scored blind to experimental condition. Coding was cross validated by two independent coders considering correct any response that named the image at the entry level (e.g., dog) or the subordinate level (e.g., Bernese Bouvier). Synonyms at an equal level of descriptiveness (at the entry level; e.g., puppy) or labels indicating an incorrect subordinate level (e.g., Shetland Sheepdog) but belonging to the correct entry level were also scored as correct. Responses at a superordinate level (e.g., "animal") were scored as incorrect (see ^{7,10} for a similar scoring procedure). We removed from analysis all trials in which the error in the single-point calibration check was $>1^\circ$ or the distance of gaze position from the screen's center

was $> 1^\circ$ when the experimental image(s) appeared, in order to ensure that participants maintained central fixation. Then intrusion responses (errors semantically consistent with the other, co-occurring image when this was correctly categorized, see 10), were removed from the analyses (see Table 1 for a summary of data removals in each experiment).

All analyses (Generalize Linear Mixed Models, independent-sample t tests) were conducted on the proportion of correct responses. GLMMs were run using the `lmer()` function of the `lme4` package⁵³ in the R programming environment (The R Foundation for Statistical Computing, Version 3.3.2); see Results for information about the models' structure. For each model, we reported the predictors' coefficients (β -values), SE-values, z-values, and associated p-values. Graphics were created using the `ggplot2` package⁵⁴. We used GLMMs because they have many advantages over traditional ANOVA models. They provide more statistical power considering all the trials without a priori grouping them into averages by subject and condition, and allow a simultaneous estimation of the variance due to subjects and items^{55,56}. In addition, GLMMs are very robust to data unbalance between experimental conditions (e.g., because of removals), allowing analyses to be performed preserving statistical power⁵⁷.

Data Availability Statement

The data from this study and the images used are available from the corresponding author on request.

References

1. Fabre-Thorpe, M. The characteristics and limits of rapid visual categorization. *Front. Psychol.* 2, 243 (2011).
2. Harnad, S. *To cognize is to categorize: Cognition is categorization.* in Handbook of categorization in cognitive science 21–54 (2017).
3. Murphy, G. . & Brownell, H. . Category differentiation in object recognition: typicality constraints on the basic category advantage. *J Exp Psychol Learn Mem Cogn* 11, 70–84 (1985).
4. Rosch, E., Mervis, C., Gray, W., Johnson, D. & Boyes-Braem, P. Basic objects in natural categories. *Cogn. Psychol.* 8, 382–439 (1976).
5. Biederman, I., Mezzanotte, R. J. & Rabinowitz, J. C. Scene perception: detecting and judging objects undergoing relational violations. *Cogn. Psychol.* 14, 143–177 (1982).
6. Bar, M. Visual objects in context. *Nat. Rev. Neurosci.* 5, 617–629 (2004).
7. Davenport, J. L. & Potter, M. C. Scene consistency in object and background perception. *Psychol. Sci.* 15, 559–564 (2004).
8. Palmer, T. E. The effects of contextual scenes on the identification of objects. *Mem. Cognit.* 3, 519–526 (1975).
9. Joubert, O. R., Fize, D., Rousselet, G. a & Fabre-Thorpe, M. Early interference of context congruence on object processing in rapid visual categorization of natural scenes. *J. Vis.* 8, 11.1-18 (2008).
10. Munneke, J., Brentari, V. & Peelen, M. V. The influence of scene context on object recognition is independent of attentional focus. *Front. Psychol.* 4, 1–10 (2013).
11. Joubert, O. R., Rousselet, G. a., Fize, D. & Fabre-Thorpe, M. Processing scene context: Fast categorization and object interference. *Vision Res.* 47, 3286–3297 (2007).
12. Mack, M. & Palmeri, T. Modeling categorization of scenes containing consistent versus inconsistent objects. *J. Vis.* 10, 1–11 (2010).
13. Oliva, A. & Torralba, A. Building the gist of a scene: The role of global image features in recognition. *Prog. Brain Res.* 155, 23–36 (2006).
14. Ganis, G. & Kutas, M. An electrophysiological study of scene effects on object identification. *Cogn. Brain Res.* 16, 123–144 (2003).
15. Mudrik, L., Shalgi, S., Lamy, D. & Deouell, L. Y. Synchronous contextual irregularities affect early scene processing: Replication and extension. *Neuropsychologia* 56, 447–458 (2014).
16. Truman, A. & Mudrik, L. Are incongruent objects harder to identify? The functional significance of the N300 component. *Neuropsychologia* 117, 222–232 (2018).
17. Guillaume, F., Tinard, S., Baier, S. & Dufau, S. An ERP Investigation of Object-Scene Incongruity. *J. Psychophysiol.* 32, 20–29 (2018).
18. Boyce, S., Pollatsek, A. & Rayner, K. Effect of background information on object identification. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 56–66 (1989).
19. Brandman, T. & Peelen, M. V. Interaction between scene and object processing revealed by human fMRI and MEG decoding. *J. Neurosci.* 37, 7700–7710 (2017).

20. O'Callaghan, C., Kveraga, K., Shine, J. M., Adams Jr, R. B. & Bar, M. Predictions penetrate perception: Converging insights from brain, behaviour and disorder. *Conscious. Cogn.* 47, 63–74 (2017).
21. Teufel, C. & Nanay, B. How to (and how not to) think about top-down influences on visual perception. *Conscious. Cogn.* 47, 17–25 (2017).
22. Palmer, S. Visual perception and world knowledge: notes on a model of sensory-cognitive interaction. in *Explorations in Cognition* 279-307 (1975).
23. Ullman, S. *High-Level Vision: Object Recognition and Visual Cognition.* (1996).
24. Kosslyn, S. *Image and Brain.* (1994).
25. Mudrik, L., Lamy, D. & Deouell, L. Y. ERP evidence for context congruity effects during simultaneous object–scene processing. *Neuropsychologia* 48, 507–517 (2010).
26. Hollingworth, A. & Henderson, J. M. Does consistent scene context facilitate object perception? *J. Exp. Psychol. Gen.* 127, 398–415 (1998).
27. Hollingworth, A. & Henderson, J. Object identification is isolated from scene semantic constraint: evidence from object type and token discrimination. *Acta Psychol. (Amst).* 102, 319–343 (1999).
28. Davenport, J. L. Consistency effects between objects in scenes. *Mem. Cognit.* 35, 393–401 (2007).
29. Gagne, C. R. & MacEvoy, S. P. Do simultaneously viewed objects influence scene recognition individually or as groups? Two perceptual studies. *PLoS One* 9, (2014).
30. Gronau, N., Neta, M. & Bar, M. Integrated contextual representation for objects' identities and their locations. *J. Cogn. Neurosci.* 20, 371–388 (2008).
31. Trapp, S. & Bar, M. Prediction, context, and competition in visual recognition. *Ann. N. Y. Acad. Sci.* 1339, 190–198 (2015).
32. Lu, Y. et al. Revealing Detail along the Visual Hierarchy: Neural Clustering Preserves Acuity from V1 to V4. *Neuron* 98, 417-428.e3 (2018).
33. Vanmarcke, S. & Wagemans, J. Individual differences in spatial frequency processing in scene perception: the influence of autism-related traits. *Vis. cogn.* 24, 115–131 (2016).
34. Bourne, V. *The divided visual field paradigm: Methodological considerations.* (2006).
35. Banich, M. T. The missing link: the role of interhemispheric interaction in attentional processing. *Brain Cogn.* 36, 128–157 (1998).
36. Davis, S. W. & Cabeza, R. Cross-hemispheric collaboration and segregation associated with task difficulty as revealed by structural and functional connectivity. *J. Neurosci.* 35(21), 8191-8200. 35, 8191–8200 (2015).
37. Delvenne, J. F. & Holt, J. L. Splitting attention across the two visual fields in visual short-term memory. *Cognition* 122, 258–263 (2012).
38. Carrasco, M., Talgar, C. P. & Cameron, E. L. Characterizing visual performance elds: Effects of transient covert attention, spatial frequency, eccentricity, task and set size. *Spat. Vis.* 15, 61–75 (2001).

39. Rousselet, G. a., Thorpe, S. J. & Fabre-Thorpe, M. Processing of one, two or four natural scenes in humans: The limits of parallelism. *Vision Res.* 44, 877–894 (2004).
40. Mackeben, M. Sustained focal attention and peripheral letter recognition. *Spat. Vis.* 12, 51–72 (1999).
41. Carrasco, M., Giordano, A. M. & McElree, B. Temporal performance fields: Visual and attentional factors. *Vision Res.* 44, 1351–1365 (2004).
42. Oliva, A. & Torralba, A. The role of context in object recognition. *Trends Cogn. Sci.* 11, 520–527 (2007).
43. Oliva, A. & Torralba, A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int. J. Comput. Vis.* 42, 145–175 (2001).
44. Kauffmann, L., Ramanoël, S. & Peyrin, C. The neural bases of spatial frequency processing during scene perception. *Front. Integr. Neurosci.* 8, (2014).
45. Gerlach, C. & Poirel, N. Navon's classical paradigm concerning local and global processing relates systematically to visual object classification performance. *Sci. Rep.* 8, 324 (2018).
46. Spotorno, S. & Tatler, B. W. The elephant in the room: Inconsistency in scene viewing and representation. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 1717–1743 (2017).
47. Underwood, G., Templeman, E., Lamming, L. & Foulsham, T. Is attention necessary for object identification? Evidence from eye movements during the inspection of real-world scenes. *Conscious. Cogn.* 17, 159–170 (2008).
48. Loftus, G. R. & Mackworth, N. H. Cognitive determinants of fixation location during picture viewing. *J. Exp. Psychol. Hum. Percept. Perform.* 4, 565–572 (1978).
49. Henderson, J. M. & Hollingworth, A. High-level scene perception. *Annu. Rev. Psychol.* 50, 243–271 (1999).
50. Võ, M. L. H. & Henderson, J. M. Object–scene inconsistencies do not capture gaze: evidence from the flash-preview moving-window paradigm. *Attention, Perception, Psychophys.* 73, 1742 (2011).
51. Konkle, T., Brady, T. F., Alvarez, G. A. & Oliva, A. Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *J. Exp. Psychol. Gen.* 139, 558–578 (2010).
52. Xiao, J., Hays, J., Ehinger, K., Oliva, A. & Torralba, A. *SUN Database: Large-scale Scene Recognition from Abbey to Zoo.* in IEEE Conference on Computer Vision and Pattern Recognition (2010).
53. Bates, D., Maechler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Statistical Softw.* 67, 1–48 (2015).
54. Wickham, H. *ggplot2: Elegant graphics for data analysis* (ed. Springer) (New York, 2009).
55. González, B., De Boeck, P. & Tuerlinckx, F. Linear mixed modelling for data from a double mixed factorial design with covariates: A case- study on semantic categorization response times. *J. R. Stat. Soc. Ser. C (Applied Stat.)* 63, 289–302 (2014).

56. Kliegl, R., Wei, P., Dambacher, M., Yan, M. & Zhou, X. Experimental effects and individual differences in linear mixed models: Estimating the relation of spatial, object, and attraction effects in visual attention. *Front. Psychol. 1*, (2011).
57. Baayen, R. H., Vasishth, S., Kliegl, R. & Bates, D. The cave of shadows: Addressing the human factor with generalized additivemixed models. *J. Mem. Lang. 94*, 206–234 (2017).

Acknowledgements

We thank Michelle Greene who kindly shared with us the script to generate the $1/f$ colored noise backgrounds.

Author contributions

AL and SS conceived the study, created the experimental materials, programmed the experiments, analyzed the data and drafted the manuscript. AL also collected the data. SF provided critical comments during study design, result interpretation and discussion, and edited the manuscript. All authors contributed to and approved the final version of the manuscript.

Additional Information

AL was supported by a doctoral funding of the French Ministry of Education and Research that had no role in the decision to submit the manuscript for publication. The authors declare no competing interests.

V.5 Discussion intermédiaire concernant les associations sémantiques objet/contexte

Dans une série de cinq expériences, nous avons étudié si l'effet de cohérence sémantique, largement décrit dans la littérature mais encore peu compris, se réalise à un stade de traitement perceptif, à un stade de mise en correspondance avec des représentations en mémoire à long terme ou à un stade plus tardif, décisionnel.

Cette recherche est la première, à notre connaissance, à analyser les associations sémantiques entre objet et contexte lorsque ces deux niveaux sont présentés séparément et simultanément via deux images. Notre étude a permis de montrer l'existence d'influences entre les traitements de l'objet et du contexte lors d'une présentation rapide, hors de la vision fovéale et sans que l'objet ne soit inclus dans le contexte de la scène.

L'effet de catégorisation sémantique a été mis en évidence dans l'ensemble des expériences, c'est-à-dire que l'objet soit inclus dans le contexte ou que ces deux niveaux soient présentés séparément, en présentation uni ou bilatérale ainsi qu'à diverses excentricités et positions sur l'écran. Nos résultats soutiennent donc les modèles de sélection de l'objet (ou modèles d'amorçage). En effet, la mise en correspondance des informations de l'objet et du contexte s'est réalisée même lorsque l'objet n'était pas inclus dans le contexte, ce qui infirme le modèle perceptif, et lorsque les biais expérimentaux décrits dans le modèle d'isolation fonctionnelle ont été limités au maximum. Ce contrôle d'hypothétiques biais a été réalisé en réduisant le risque que des réponses devinées correspondent aux réponses attendues et par l'exclusion des éventuelles réponses devinées (i.e., les intrusions). Pour ce faire, les réponses n'étaient pas restreintes à un choix prédéfini limité et étaient données verbalement de façon ouverte. En outre, la présentation rapide et simultanée des deux images a permis de limiter un éventuel effet d'amorçage.

L'effet facilitateur de la cohérence sémantique est retrouvé de façon identique pour l'objet et pour le contexte, ce qui suggère que des traitements rapides et « approximatifs » du contexte mais également de l'objet sont réalisés très rapidement et permettent de faciliter la reconnaissance des deux niveaux de l'image. Nos résultats sont donc en faveur d'une interaction mutuelle des deux niveaux (objet et contexte) et complètent les travaux de certains auteurs (Davenport & Potter, 2004 ; Joubert et al., 2007).

En conclusion, l'effet de cohérence sémantique est un effet robuste : il résiste aux dégradations perceptives dues à l'excentricité et il est maintenu même lorsque les constituants de la scène sont présentés de façon séparée et lorsque les éventuels biais expérimentaux qui

pourraient l'expliquer sont contrôlés. L'ensemble de nos résultats suggère que la représentation unifiée d'une scène visuelle, et la mise en relation des contenus sémantiques entre ses constituants, est issue du stade cognitif de mise en correspondance entre un percept et une représentation en mémoire à long terme.

Chapitre VI : Etude de l'interaction entre les traitements sémantiques et émotionnels lors de la perception rapide de scènes visuelles complexes chez l'adulte

Ce chapitre présente la seconde étude empirique de ce travail de thèse qui porte sur la question de l'interaction entre les traitements sémantiques et émotionnels dans les scènes visuelles. Il est constitué d'un bref rappel des principaux apports théoriques, d'une présentation de la méthodologie de l'étude, de son résumé en français, de la reproduction de l'article qui sera soumis pour publication au jour de la soutenance et enfin d'une discussion intermédiaire.

Leroy, A., Faure, S., & Spotorno, S. (submitted on the day of the defense). Interplay between semantic and emotional information in visual scene processing.

VI.1 Rappel des apports théoriques principaux

La première étude empirique de cette thèse (article IV) nous a permis de montrer que l'effet de cohérence sémantique est lié à l'activation de représentations en mémoire à long terme et dépend donc d'un traitement de haut niveau cognitif. De manière similaire aux questionnements sur les associations sémantiques, le stade auquel les contenus émotionnels de l'objet et du contexte s'associent est toujours débattu (Musch & Klauer, 2003 ; Spruyt, De Houwer, Hermans, & Eelen, 2007 ; Van den Stock & de Gelder, 2014). En outre, alors qu'il a été montré que le contenu émotionnel du contexte aide la catégorisation de l'émotion portée par l'objet, lorsque ces deux niveaux sont porteurs du même contenu émotionnel, la réciproque, c'est-à-dire l'influence du contenu émotionnel de l'objet sur la catégorisation émotionnelle du contexte, n'a pas encore été à notre connaissance explorée. Enfin, les éventuelles interactions entre les effets de cohérence émotionnelle et sémantique n'ont pas été étudiées malgré leur intérêt dans le débat relatif à l'indépendance/dépendance entre les traitements émotionnels et sémantiques.

Notre seconde étude empirique a ainsi pour but d'étudier l'interaction entre les traitements sémantiques et émotionnels lors de la perception visuelle rapide de scènes visuelles. Pour ce faire une image d'objet et une image de contexte, entre lesquelles sont manipulées les conditions de cohérences sémantique et émotionnelle, seront présentées selon le paradigme de champ visuel divisé durant deux tâches expérimentales différentes attirant l'attention du participant soit sur le contenu sémantique soit sur le contenu émotionnel.

VI.2 Présentation de l'étude

L'article IV (chapitre V) a permis de préciser que l'effet de cohérence sémantique est maintenu même lorsque l'objet et le contexte sont présentés séparément. Ce résultat, au-delà de son importance fondamentale, a un intérêt méthodologique. En effet, l'étude de la cohérence sémantique entre un objet et un contexte présentés simultanément se réalise, généralement, en incluant l'image d'un objet dans une image de contexte. Cependant, ceci peut être à l'origine de biais perceptifs, de scènes visuelles qui semblent étranges, qui perdent de leur aspect réel et ainsi, qui risquent d'influencer les résultats. A contrario, en présentant l'objet et le contexte de façon séparée, il n'y a pas de modification de la structure même de l'image. Comme l'effet de cohérence sémantique a été répliqué dans l'article IV du chapitre V avec une présentation séparée et simultanée des constituants de la scène visuelle, nous pouvons utiliser cette méthodologie pour étudier les relations entre contexte et objet sans avoir à modifier expérimentalement les images issues de photographies d'environnement ou d'objets réels.

Nous avons ainsi proposé à des participants de catégoriser, soit au niveau sémantique soit au niveau émotionnel, une image définie au sein d'une paire d'images (un objet et un contexte) flashée rapidement (100 ms), chaque image étant présentée dans un hémichamp visuel différent, à droite ou à gauche d'un point de fixation central. Comme pour l'étude précédente (article IV, chapitre V), nous avons enregistré les mouvements oculaires des participants afin de nous assurer de la fixation centrale lors de l'apparition des deux images et ainsi que les deux images soient vues hors de la vision fovéale. Par contre, à la différence de la méthodologie de cette même étude sur la cohérence sémantique (article IV, chapitre V) dans laquelle les réponses étaient données à l'oral et n'étaient donc pas contraintes par un nombre de réponses limité, nous demandons ici aux participants de répondre par choix forcé manuel (pour plus de précisions se référer dans ce chapitre à la description de la méthodologie dans l'article V). Lors de la tâche de catégorisation sémantique du contexte, les participants doivent choisir parmi deux réponses possibles : environnement « manufacturé » ou « naturel ». Dans la tâche de catégorisation sémantique de l'objet, ils doivent choisir entre « objet inanimé », « animal » ou « être humain ». Enfin, dans les tâches de catégorisation du contenu émotionnel (contexte ou objet) les choix possibles sont : image « neutre », « positive » ou « négative ». Les touches de choix de réponse sur clavier Azerty sont contrebalancées entre les participants.

La tâche de catégorisation sémantique et la tâche de catégorisation émotionnelle utilisent les mêmes images et même croisement des variables de cohérence sémantique et émotionnelle, ce qui permet d'éviter tout biais lié au matériel expérimental. Le matériel expérimental est

constitué de sept cent quatre-vingt-douze images naturelles (photographies en couleur) : 396 « contextes », donnant priorité au fond de l'image et donc aux informations ayant un caractère global (comme dans l'image d'une plage, d'une forêt, etc.), et 396 « objets », donnant priorité aux informations locales en premier plan (comme dans l'image d'une personne, d'un animal ou d'un objet inanimé, par exemple un livre). Une moitié de ces images est neutre sur le plan émotionnel et l'autre moitié porteuse de valence émotionnelle positive (50%) ou négative (50%). Les images émotionnelles proviennent majoritairement de *l'International Affective Picture system* (IAPS ; Lang, Bradley, Cuthbert, 2008) et du *Nencki Affective Picture System* (NAPS ; Marchewka, Żurawski, Jednoróg, Grabowska, 2014), qui sont des bases d'images validées selon la valence émotionnelle et le niveau d'activation physiologique. Pour la condition neutre, les images proviennent majoritairement de la base d'images SUN (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010) et de Google Images (images libres de droit). Les images d'objets émotionnels ne sont constituées que par des objets animés (50 % d'humains, 50 % d'animaux), car il a été montré que ceux-ci suscitent une activation émotionnelle plus forte que les objets inanimés (Carniglia, Caputi, Manfredi, Zambarbieri, et Pessa, 2006 ; Carretié, 2014). En outre, il a été mis en évidence, à la fois au niveau comportemental et à la fois au niveau des activations cérébrales, des patterns de résultats différents selon que les objets soient naturels ou manufacturés (Kalénine & Bonthoux, 2006 ; Kalénine *et al.*, 2009). L'utilisation d'objets naturels non manipulables pour toutes nos images d'objets émotionnels nous permet de limiter les biais d'interprétation de nos résultats. Les images d'objets neutres sont composées de 2/3 d'objets inanimés, 1/6 d'humains et 1/6 d'animaux. Les bases d'images IAPS et NAPS sont les deux les plus utilisées dans la recherche qui implique des images émotionnelles : le matériel a donc été mis en œuvre par de nombreuses études auprès de grands échantillons d'individus et porte sur des facteurs et phénomènes très variés (notamment les effets de valence émotionnelle et le niveau d'activation physiologique). Une étude pilote a de plus été réalisée auprès de juges indépendants, qui n'ont pas participé pas à la présente étude expérimentale, afin d'obtenir une évaluation française de la catégorisation sémantique et émotionnelle de toutes les images (voir Tableau 1 de l'article pour la description de la répartition des images et la comparaison des évaluations de notre étude pilote avec les évaluations de l'IAPS et de la NAPS).

Les paires d'images ont été créées initialement en condition de cohérence sémantique et émotionnelle (e.g., un skieur joyeux et un paysage de montagne enneigée un jour de beau temps ou un requin et un océan en pleine tempête) ; l'objet d'une paire a été ensuite associé au

contexte d'une autre paire afin d'obtenir différentes versions expérimentales : images cohérentes sémantiquement mais incohérentes émotionnellement (e.g., un skieur joyeux et une avalanche) ; images incohérentes sémantiquement et cohérentes émotionnellement (e.g., un skieur joyeux et une plage paradisiaque) ; images incohérentes sémantiquement et émotionnellement (e.g., un skieur joyeux et un océan en pleine tempête).

Cette étude est ainsi la première, à notre connaissance, à proposer des tâches de catégorisation sémantique et du contenu émotionnel d'un objet et d'un contexte présentés rapidement et séparément, pour lesquels les conditions de cohérences sémantique et émotionnelle ont été simultanément manipulées. Cette méthodologie nous permet d'étudier si d'éventuelles interactions se réalisent lors des traitements sémantiques et émotionnels, à savoir si le contenu sémantique de l'image cible et/ou de l'image associée influence l'évaluation du contenu émotionnel de l'image cible et, inversement, si le contenu émotionnel de l'image cible et/ou de l'image associée influence la catégorisation sémantique de l'image cible.

VI.3 Résumé de l'étude

Pour s'adapter à un environnement continuellement changeant, nous devons détecter et reconnaître les stimuli qui apparaissent dans notre champ visuel. Il a été montré que la catégorisation sémantique se réalise en quelques centaines de millisecondes et que le traitement du contenu émotionnel est souvent priorisé, mais nous ne savons pas si, et comment, ces deux processus interagissent pendant la perception rapide de scènes visuelles. En outre, la compétition attentionnelle entre deux images de même contenu émotionnel (e.g., les deux positives) ou de contenu émotionnel différents (e.g., une positive et une négative), n'a pas, à notre connaissance, été étudiée lorsque l'une des deux images seulement est pertinente pour la tâche. Notre objectif est d'analyser les processus sémantiques et émotionnels, et leurs interactions possibles, lors de la présentation rapide d'images d'objets et de contextes représentant des scènes visuelles du quotidien. Une tâche sémantique et une tâche émotionnelle impliquant les mêmes images positives, négatives ou neutres ont été créées. L'image cible est présentée simultanément avec une autre image, chacune dans un champ visuel différent et la cohérence sémantique et émotionnelle entre les deux images est manipulée. Dans la tâche de catégorisation sémantique, nous avons montré un effet d'interférence du contenu émotionnel négatif uniquement pour le contenu émotionnel de l'image cible et un effet de cohérence sémantique qui impact le temps de réponse des images de contexte. Dans la tâche émotionnelle, nous avons mis en évidence un effet de cohérence sémantique facilitant la catégorisation

émotionnelle des contextes négatifs et un effet de cohérence émotionnelle améliorant la catégorisation émotionnelle de l'objet et du contexte. La cohérence émotionnelle n'interagit jamais avec la cohérence sémantique et est donc probablement indépendante de celle-ci. Dans l'ensemble, nos résultats suggèrent que les contenus sémantique et émotionnel sont toujours traités durant la perception rapide des scènes visuelles, quelle que soit la tâche. En outre, l'absence d'effet de cohérence émotionnelle à la tâche sémantique, alors qu'un effet de cohérence sémantique est mis en évidence dans la tâche émotionnelle, suggère que les traitements émotionnels seraient plus dépendants de l'orientation attentionnelle de l'observateur que les traitements sémantiques.

VI.4 Article V. Interplay between semantic and emotional information in visual scene processing.

Le texte ci-dessous est la reproduction d'un article qui sera soumis au jour de la soutenance pour publication.

Leroy, A., Faure, S., & Spotorno, S. (submitted on the day of the defense). Interplay between semantic and emotional information in visual scene processing.

The interplay between semantic and emotional information in rapid visual scene processing

Anais Leroy*, Sylvane Faure* & Sara Spotorno[§]

* Université Côte d'Azur, LAPCOS, France

[§] School of Psychology, University of Keele, United Kingdom

Corresponding author:

Anaïs Leroy

LAPCOS, MSHS Sud Est

Université Côte d'Azur

Pôle Universitaire Saint Jean d'Angély

24 avenue des Diables Bleus

06357 Nice (France)

Phone: +33 4 89 88 14 60

Email: anais.leroy@univ-cotedazur.fr

Abstract:

To adapt to a continuously changing environment, we need to recognize emotional and non-emotional stimuli that appear in our visual field. It has been shown that semantic processing occurs briefly, and that emotional content is prioritized, but it is unclear how these two processes interact during scene viewing. Moreover, we do not know how two emotional scenes, with or without the same emotional content, compete for attentional resources when one of them is relevant for the task. We aimed to analyse semantic and emotional processes, and their possible interactions, during rapid presentation of object and context images representing everyday visual scenes. A semantic categorization task and an emotional categorization task, involving the same positive, negative or neutral images, were created. The target image was presented simultaneously with another image, each in a different visual field, and the semantic and the emotional consistency between them was manipulated. In the semantic categorization task, we found a negative interference effect only for the emotional content of the target image and a semantic consistency effect reducing response times for contexts images. In the emotional task, we found a semantic consistency effect improving emotional categorization of negative contexts and an emotional consistency effect improving both object and context emotional categorization. This effect of emotional consistency never interacted with, and thus is independent from, semantic consistency. Overall, our results suggest that semantic processing and emotional processing always occur during scene perception whatever the task and that emotional processing depends more on attentional allocation than semantic processing.

Keywords:

Visual scenes, rapid categorization, emotional content, semantic content, consistency

Introduction

Visual categorization occurs very briefly (Thorpe, Fize, & Marlot, 1996). A core aspect of this ability, necessary for adapting to our environment, concerns processing of the association between the scene's context (the background, structured around global perceptual properties, see Oliva & Torralba, 2001) and the foreground objects (local components) included within the context. Numerous studies (*e.g.* Davenport & Potter, 2004; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; see for review Fabre-Thorpe, 2011) have shown a better categorization performance, with greater accuracy and/or shorter response time, when the context and the objects are semantically consistent, that is to say frequently associated to each other in our everyday experience (*e.g.*, a computer in an office), compared to when they are semantically inconsistent (*e.g.* a computer in a forest). Therefore, the semantic information carried by the different components of an image allows us to recognise visual scenes and the objects included in it, in a fast, very efficient manner, within just a glance.

When we perceive a visual scene we not only have access to its meaning but also to the emotional feelings conveyed by it (*e.g.*, Barrett & Bar, 2009). Emotional scenes are characterized by the presence of a context and/or objects that convey an emotion (Sabatinelli et al., 2011). They are closely linked to the quality of our subjective experience and are, therefore, of great importance for the behaviours and interactions of everyday life. Attentional allocation is enhanced for emotional scenes compared to neutral ones and the emotional information is prioritised (see for review Vuilleumier, 2015). Emotional scenes attract attention more than neutral scenes regardless of the experimental task (Bradley, Codispoti, Cuthbert, & Lang, 2001; Pilarczyk, Joanna; Kuniecki, 2014) and even when participants are instructed not to look at them (Nummenmaa, Hyönä, & Calvo, 2006). Evolutionary psychological approaches claim that it is essential to extract and analyse very quickly emotional information from our environment in order to react appropriately, particularly with approach or avoidance responses (Barrett & Bliss-Moreau, 2009). However, it is not well understood how this prioritisation is carried out, and how emotional information and semantic information relate to each other during rapid scene processing.

Interplay between semantic information and emotional information in visual scenes.

Some authors argue that an emotional response may occur in a rapid and automatic way, that it is to say with no or few attentional resources, without requiring voluntary and conscious cognitive assessment (LeDoux, 2003; Zajonc, 1984). This Affective Primacy Hypothesis is

supported in particular by the phenomenon of simple exposure, which concerns an emotional reaction without consciousness and without semantic processing of the emotional stimulus. Thus, when faced with potential danger, viewers may first feel afraid and only subsequently be able to identify the potentially dangerous stimulus. This hypothesis argues, therefore, that emotional processing is either independent of, or takes place before, semantic processing. The independence of emotional and semantic processes is supported by evidence of independence of the underlying neural circuits (Vuilleumier, 2005). Support to affective primacy has also been provided by the emotional priming paradigm, where emotional judgement is facilitated when the target is preceded by a prime conveying the same emotion, compared to when the prime and the target convey different emotions. Further support has come from findings showing that emotional, task irrelevant stimuli presented subliminally influence task performance (see Tamietto & de Gelder, 2010).

From the seminal work of Lazarus (1984), an opposite approach, the Cognitive Primacy Hypothesis, has been claiming that cognitive processing of a stimulus is a necessary precondition to have access to the conveyed emotion (e.g., Nummenmaa, Hyönä, & Calvo, 2010; Pilarczyk, Joanna; Kuniecki, 2014; for a review, see Storbeck, Robinson, & McCourt, 2006). Therefore, semantic categorization, based on conscious recognition of the percept, would precede emotional categorization (Lähteenmäki, Hyönä, Koivisto, & Nummenmaa, 2015b). For these authors, emotional priming may indeed be a semantic priming effect (Storbeck et al., 2006). Moreover, processing of emotional information would not be automatic but require allocation of attentional resources (Luiz Pessoa, Kastner, & Ungerleider, 2002). The appraisal-based emotion theory (Ellsworth & Scherer, 2003; David Sander, Grandjean, & Scherer, 2005) argues that the stimulus does not have an intrinsic emotional value, but that value comes from the interpretation of the stimulus according to the relevance for the individual.

It has also been shown that the type task (semantic or emotional categorization), the nature of stimulus (images or words) and the setting of the task (emotional or not) all influence whether semantic or emotional information is processed first (Lai et al., 2012). For instance, response times for semantic categorization of visual scenes (e.g., indoor vs. outdoor) are longer in an emotional setting than in a non-emotional one, although the response times for emotional categorization of the same images (pleasant vs. unpleasant) are not affected by the emotional setting. Thus, emotional categorization is faster than semantic categorization in an emotional setting although the two types of categorization require nearly the same time in the non-

emotional setting. For words this difference is even more pronounced, with faster response times for emotional categorization in an emotional setting and faster semantic categorization in a non-emotional setting. It has also been suggested that affective primacy depends on the stimulus' properties, for some stimulus, emotional categorization occurs first but if the stimulus is visually degraded or does not match a highly learned representation, its emotional categorization is secondary to a semantic processing (Mueller et al., 2017).

Object-context associations during emotional processing

Our ability to quickly process emotional information allows us to understand others' emotional state and to react promptly to relevant information from our environment (*i.e.*, to detect an aggressive dog). When a context and an object of a visual scene co-occur in an emotionally consistent condition, that is they convey the same emotion, their emotional content is categorized more accurately and quickly than when the context and the object are presented in an emotionally inconsistent condition (de Borst & de Gelder, 2016; Kret & de Gelder, 2010; R. Righart & de Gelder, 2008).

Following several studies, contextual emotional information is automatically processed and integrated with finer, detailed emotional information, even when participants are asked to ignore the context (R. Righart & de Gelder, 2008). The emotional context, for instance, influences recognition of facial (R. Righart & de Gelder, 2008) and body (Kret & Gelder, 2010) emotional expressions, either when presented simultaneously or as a prime (Hietanen & Astikainen, 2013). Moreover, the body expression influences recognition of the facial expression, and this both with static (Meeren, van Heijnsbergen, & de Gelder, 2005) and dynamic (Nelson & Mondloch, 2017) stimuli. More specifically, the face expression is better recognized when the face is presented on a body with an emotionally consistent posture (*e.g.*, a happy face on a happy body) compared to an inconsistent one (*e.g.*, a happy face on an angry body). Recognition is further facilitated if the body itself is integrated into an emotionally consistent context (*e.g.*, a happy body in a party scene) compared to an emotionally inconsistent context (*e.g.*, a happy body in a cemetery) (Reschke, Knothe, Lopez, & Walle, 2018).

Most of these studies have focused on a small portion of the visual field, presenting visual scenes in central, high-acuity vision. However, in everyday life we need to detect and recognize information also in peripheral vision in order to respond efficiently to our environment. Some research has shown that fast emotional evaluation can be performed even when the images are presented peripherally, at up to 60 degrees of eccentricity for positive

emotional content and up to 30 degrees for negative emotional content (D'Hondt, Szaffarczyk, Sequeira, & Boucart, 2016). Calvo, Rodriguez-Chinea and Fernandez-Martin (2015) presented one emotional image and one neutral image, each in the right or left visual field, for 150ms, and asked participants to judge in which visual field the emotional scene appeared. They reported performance above chance. Moreover, it has been suggested that the emotional content is processed involuntary even when presented in the periphery. Indeed, a lateralized irrelevant emotional scene decreases performance at a task carried out in central vision, for example, a digit or letter discrimination task (Calvo & Nummenmaa, 2015) or direction judgement concerning left- or right-oriented arrows (D'Hondt et al., 2013).

It is, thus, admitted that when a neutral and an emotional scene are presented peripherally, attention is oriented toward the emotional scene. However, in our daily lives multiple emotional stimuli often appear simultaneously. Only a few studies have examined how two emotional images, carrying either the same or a different emotional content, compete for attentional resources when they are presented simultaneously, separately and each in a different visual field. For instance, in the study carried out by Fernández-Martín and Calvo (2016) irrelevant emotional scenes were presented in peripheral vision during a letter discrimination task in central vision with eye-movement recording. While irrelevant for the tasks, emotional scenes were fixated, showing that they attract attention and that processing of emotional information occurs even in covert attention with longer first fixation and total fixation duration for positive compared to negative scenes. Padmala, Sambuco, Codispoti and Pessoa (2018) presented two emotional scenes (positive or negative) or two neutral scenes as distractors during a letter discrimination task. They showed greater attentional interference when the two scenes were both emotional (regardless the type of pairing: positive/positive, negative/positive or negative/negative) than when one was neutral. However, in these studies, the emotional scenes were not the target image and the emotional content was task irrelevant. To our knowledge, no study has explored the attentional competition between two emotional scenes when one of them is the target image, with or without emotional content task relevant. We aim thus to explore the emotional categorization and the semantic categorization, of an emotional image presented rapidly, outside foveal (central) vision and simultaneously with another image of the same or different emotional content (positive or negative). Indeed, in our daily lives, we often have to recognize, in a fast, efficient manner, an emotion even if we are in an environment conveying a different emotional content. For example, we may have to recognize the emotion expressed by a dog that suddenly appears, or the emotion conveyed by a change in our environment, like

a tsunami, even when other elements in our visual field convey different emotional content. It is, thus, important to analyse if the emotional consistency effect occurs even when individuals do not have the time to explore the visual scene, during the early visual scene processing.

The present study

It has been mainly described the influence of semantic consistency between the context and the object of a visual scene to facilitate their semantic categorization, on the one hand, and the influence of emotional consistency on the categorization of emotional content on the other hand. However, we do not know whether semantic and emotional consistency effects interact during rapid scene perception and whether the emotional consistency effects occur between two separated and briefly presented images outside the fovea. This understanding is nevertheless essential to allow a global, integrative account of visual perception. Moreover, it also allowed us to contribute to the debate concerning the cognitive or affective primacy. Indeed, if processing of emotional information depends on semantic processing (Cognitive Primacy Hypothesis), the emotional consistency effect would be mostly influenced by semantic consistency and even, as it was the case for the emotional priming effect, the specific existence of this effect might be challenged.

In the present study, we aimed to examine semantic processing and emotional processing, and their possible interactions, during rapid categorization of a visual scene (composed by a context and an object) that quickly appeared outside the foveal vision. This early processing is essential, as it enables prompt adaptation and is the first step before a more precise analysis can begin. Moreover, this early processing enables understanding of the relevance of the stimulus for the individual. Our aim is, thus, to analyse which stimuli are processed as a priority and which stimuli are linked to another component of the visual scene to assist in the categorization process.

To this purpose, we proposed a semantic categorization task and an emotional categorization task during fast visual presentation of a pair of images (a scene context and an object), each in a different visual field, using the divided-visual-field paradigm (see Bourne, 2006). In each trial, two images were presented, one as the target image, and the other as distractor. The same stimuli were used for both tasks, and we manipulated orthogonally the semantic and emotional consistency between the two images. To approach reality, we used images depicting contexts and objects that form the visual scenes encountered in everyday life. These two experimental tasks directed the subject's attention to semantic or emotional

information respectively. We could thus study the influence of the involuntary processing of one dimension (task irrelevant) on the voluntary processing of the other (task relevant) and, thus, examine which stimulus content is processed as a priority.

The study of semantic and emotional object-context associations allowed us to analyse whether emotional processing depends on semantic processing. On the one hand, if emotional categorization requires semantic categorization (Cognitive Primacy Hypothesis), semantic inconsistency should reduce accuracy and/or lengthen the time needed for emotional categorization of the target on the emotional categorization task. On the other hand, if emotional categorization is carried out without requiring semantic categorization (Affective Primacy Hypothesis: Zajonc, 1980), the manipulation of semantic consistency should not influence emotional categorization performance.

Our result may enrich understanding of what kind of information (semantic or emotional) is processed first during fast visual scene processing, how the emotional contents of two emotional scenes compete for attentional resources, and how the different perceptual level of the image (global/context or local/object) interact, depending on task relevance.

General method

Apparatus

The experiments were conducted on a HP Compaq computer running OS Windows XP. The images were presented on a DELL 7004 CRT monitor (32.5cm x 24.5cm, 1024 × 768 pixels, refresh rate: 85Hz). Image presentation and response recording was controlled by Experiment Builder (SR Research, Canada). Eye movements were recorded using an EyeLink 1000 (SR Research, Canada) at a sampling rate of 1000Hz. Viewing was binocular, but only the dominant eye was tracked.

Materials

We used 144 full-colour photographs of real-word scene contexts (50% natural, 50% man-made), plus six as practice, and 144 full-colour photographs of objects (1/3 humans, 1/3 animals and 1/3 inanimate items), plus six as practice. Half of the images were neutral, the other half were emotional images (50% with positive content, 50% with negative content). They were selected (see Table 1) from the International Affective Picture System (IAPS database, Lang, Bradley, & Cuthbert, 2008), the Nencki Affective Picture System (NAPS database,

Marchewka, Zurawski, Jednoróg, & Grabowska, 2014), the Scene UNderstanding (SUN) database (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010) and Google Images. “Object” images were defined by a unique and central object (e.g., skier, shark, spoon) included into a background. “Context” images were mainly defined by global information (e.g., mountain, sea, kitchen), even though they also contained local information, and did not contain humans or animals. The emotional object images depicted only animate entities (50% human, 50% animal), as they have been shown to elicit stronger emotional activation than inanimate entities (Carniglia, Caputi, Manfredi, Zambarbieri, & Pessa, 2006; Carretié, 2014). The neutral object images were composed of 2/3 of inanimate entities, 1/6 of humans and 1/6 of animals.

Image modifications (e.g., changing the size, removal of animals, humans or inanimate objects in context images) were made using Adobe Photoshop CS 9.0 (Adobe, San Jose, CA). Image were presented at 315 x 252 pixels (10 cm x 8 cm).

We carried out an evaluation study of the experimental materials (see Table 1). Eight independent judges ($M=21.5$, $SD=1.75$), who did not take part in the main experiments, rated the emotional content of each image. The images from the IAPS and the NAPS had already been evaluated, but not with a French population; the images from the SUN and Google Images databases had not been formerly evaluated for their valence and their arousal level. The evaluation study was created and run in OpenSesame 3.1.9 (Mathôt, Schreij, & Theeuwes, 2012). Each image was presented in the center of the screen on a medium grey background (127, 127, 127) with the same size as in the main experiments, until the response. Each judge rated valence and arousal (i.e., level of physiological activation) of 144 randomly presented images on two Likert scales (from 1-very negative to 9-very positive, for valence and from 1-very calm to 9-very excited, for arousal). Each judge saw half of object images and half of context images. Half of these images were neutral, the other half were emotional (half positive, half negative). This study (see Table 1) showed that the valence of positive images was rated higher than the valence of neutral and negative images, and that the valence of neutral images was rated higher than the valence of negative images. This was true for both context (all $ts \geq 12.07$, all $ps < .001$) and object (all $ts \geq 14.70$, all $ps < .001$) images. Considering context images, the arousal level of negative images was rated higher than that of positive and neutral images, and the arousal level of neutral images was rated higher than that of positive images (all $ts \geq 4.77$, all $ps < .001$). Considering object images, the arousal level of negative images was rated higher than that of positive and neutral images (both $ts \geq 12.26$, both $ps < .001$), while the arousal ratings of positive and neutral images did not differ significantly ($t < 1$, $p = .991$).

In the main experiments, we orthogonally manipulated semantic consistency (consistent vs. inconsistent) and emotional consistency (consistent vs. inconsistent) only in trials in which the image pair was composed by two emotional (positive and/or negative) images. We also presented trials in which both images were neutral (always semantically and emotionally consistent) in order to prevent participants from habituation to the presence of emotional content and to inconsistency. Images were also mirror-reversed, thus the constituent elements of the image were located at the same distance from the center of the screen (and therefore from the participant's fixation) when the image was presented either in the right or left visual field. Eight versions of each experiment were created in order to counterbalance semantic consistency, emotional consistency and visual field within and between participants. Initially, each pair was created with an object image and a context image that were both semantically and emotionally consistent (a): for example, the image of a happy skier was paired with the image of a beautiful mountain context, and the image of a frightening shark was paired with the image of a sea storm. Object and context images were then swapped (see Figure 1) in order to obtain: (b) pairs that were semantically consistent but emotionally inconsistent (for instance, the happy skier was paired with a mountain context depicting an avalanche, and the frightening shark was paired with a paradisiac lagoon); (c) pairs that were semantically inconsistent but emotionally consistent (i.e. the happy skier was paired with the paradisiac lagoon and the frightening shark with the avalanche); (d) pairs that were both semantically and emotionally inconsistent (i.e. the happy skier was paired with the sea storm and the frightening shark with a beautiful mountain landscape). We also created, in Matlab, a white noise mask presented with the same size as the experimental images.

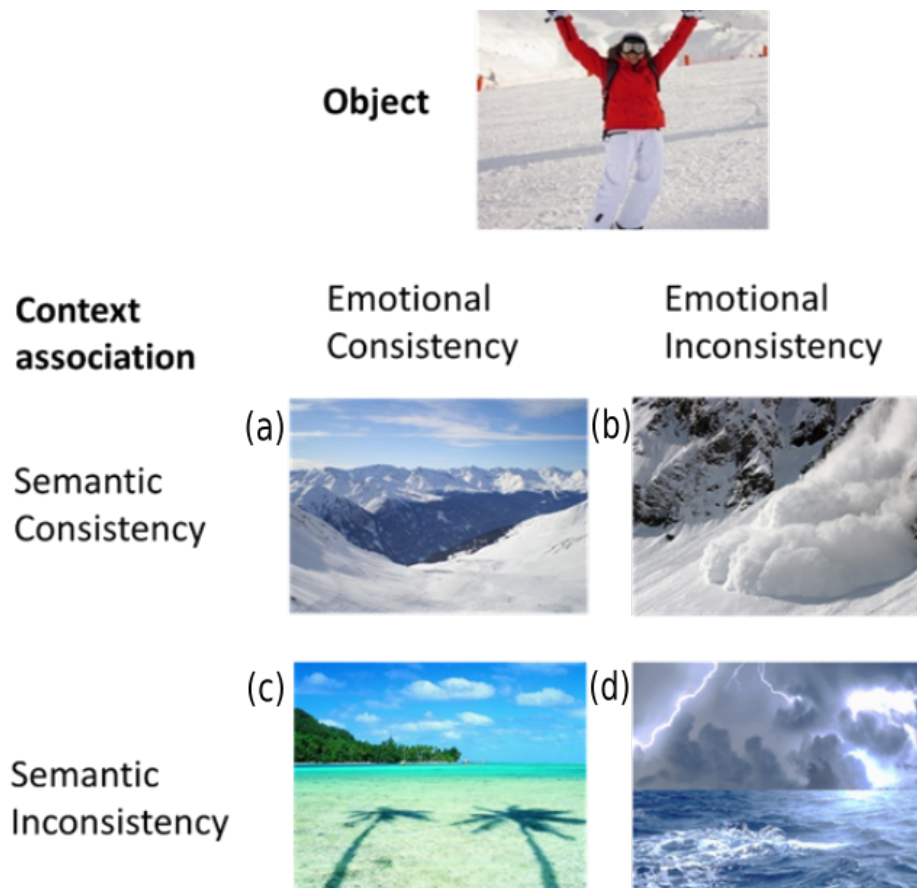


Figure 1: Example of image pairings in order to manipulate the semantic consistency and emotional consistency. First, (a) an object image is paired with a semantically and emotionally consistent context image (here the happy skier with the beautiful mountain). Then to create the different experimental conditions this object image is cross with the context image of another pairs of semantically and emotionally consistent images. Depending of the context image we obtain (b) a semantically consistent but inconsistent pair (here the happy skier with the avalanche), (c) a semantically inconsistent but emotionally consistent pair (the happy skier with a paradisiac lagoon) or even (d) a semantically and emotionally inconsistent pair (the happy skier with a rage sea).

		Object images				Context images			
		IAPS	NAPS	Google and SUN data base	Evaluation pilot study	IAPS	NAPS	Google and SUN data base	Evaluation pilot study
	N	24	17	31	72	7	17	48	72
Positives images	Valence	7.3 (0.5)	7.1 (0.4)		7.2 (1.2)	7.5 (0.3)	7.4 (0.5)		7.2 (1.4)
	Arousal	4.5 (0.6)	4.5 (0.7)		4.2 (2.3)	4.3 (0.6)	3.3 (0.4)		3.3 (2)
	N	42	17	13	72	14	10	48	72
Negatives images	Valence	3.2 (0.6)	3.4 (0.7)		2.5 (1.4)	3.1 (0.5)	3.3 (0.3)		3 (1.3)
	Arousal	5.6 (0.6)	5.7 (1.1)		7.1 (1.6)	4.8 (0.6)	5.4 (0.9)		6.5 (1.6)
	N	29	35	80	144	8	27	109	144
Neutral images	Valence	5 (0.2)	5.8 (0.6)		5 (0.9)	5.2 (0.5)	6.2 (1) 3.8		5.4 (1)
	Arousal	2.9 (0.5)	4.3 (0.6)		4.1 (1.6)	3.6 (0.6)	(0.8)		4.2 (1.6)

Table 1. Valence and arousal of the images. For each type of image (Object or Context), we report the number of positive, negative and neutral images from the IAPS, the NAPS or from the SUN and Google Images databases. We then report the mean score (*with standard deviation in brackets*) of valence and arousal for the images from the IAPS, the NAPS and for all the images by a pilot study.

Procedure

Participants were tested individually, seated in a quiet, dimly illuminated room, in front of the screen, laying with their forehead and chin on the chinrest and their eyes approximately 57 cm away from the screen. We recorded participant's eye movements to ensure central fixation. Prior to the experimental task, each participant underwent a randomized nine-point calibration procedure. Recalibrations were performed during the task if necessary. Before each trial sequence, a single-point calibration check was applied while the participant fixated a black dot in the center of the screen. At the beginning of the experiment, participants performed six practice trials, followed by 288 experimental trials presented in random order. Each trial (Figure 2) started with a 90ms black central fixation cross (Times New Roman, bold, font size 40), followed by a pair of images (one context image and one object image) presented for 100ms, and then by pair of white noise masks presented for 80ms. Each image appeared in one visual field (one in the left and one in the right) with respect to the screen's vertical midline, with the closest edge at 2.5 degrees horizontally from the center of the screen. The colour of the screen background was medium grey (127, 127, 127).

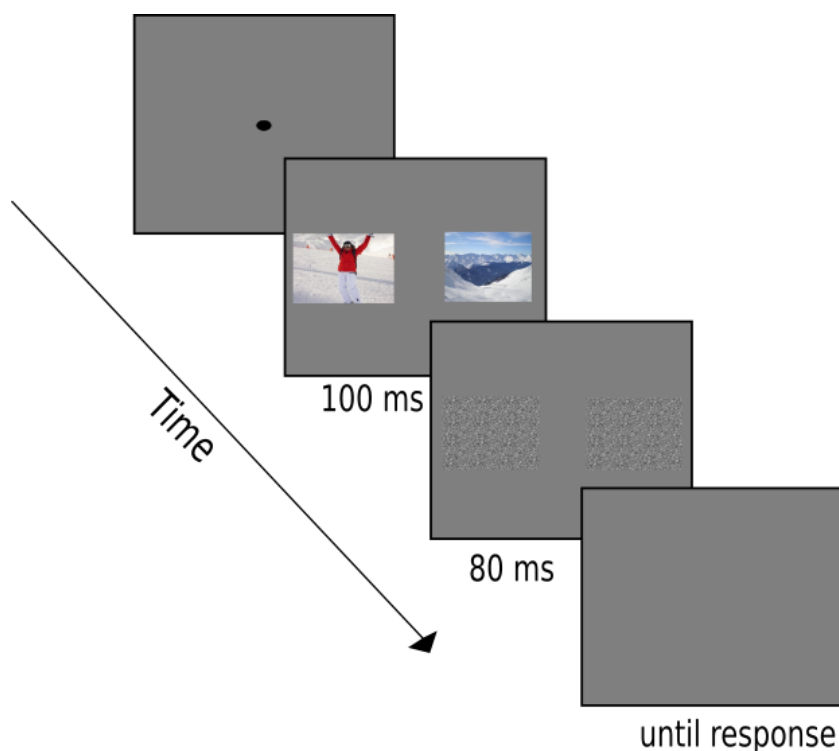


Figure 2: Trial example for both Experiment 1 (semantic categorization) and Experiment 2 (emotional categorization). An object image and context image were presented separately and simultaneously each in a different visual field (left vs. right) followed by two white noise masks. Then a grey screen stayed until participants' response.

Data analysis

Analyses were conducted using the `lmer` function of the `lme4` package (Bates et al., 2015) in the R programming environment (The R Foundation for Statistical Computing, Version 3.3.3, 2016). We used GLMM and LMMs because they have many advantages over traditional ANOVA models. Crucially, they perform item analysis together with subject analysis, and allow a simultaneous estimation of between-subject and between-item variances. In addition, they are known to be more robust than ANOVAs when a design is not fully balanced as a result of data exclusions (see Kliegl, Wei, Dambacher, Yan, & Zhou, 2011).

For each experiment, we ran one generalized linear mixed model (GLMM) on accuracy for each dependent measure (with binary responses: correct or incorrect), and one linear mixed model (LMM) on log-transformed response times of correct trials, with participant and target image as random factors. A total of 966 trials (4.42% of the trials for the semantic task and 6.14% of the trials for the emotional task) were excluded from the analyses because the error in this single-point calibration check was greater than 1 degree. Moreover, trials with a response time exceeding 3 *SD* from the mean were excluded as outliers (1.60% of the remaining trials for the semantic task and 1.10% for the emotional task). All models were first run with full random structure. When full model did not converge, we simplified it by removing correlations between the random slopes and the intercepts, and then if necessary by further reducing the complexity of the model (see Supplementary materials for the list of model structures).

First, we report the results of semantically consistent and emotionally consistent trial in order to compare the neutral target images with positive and with negative images, in order to have a baseline of the categorization performance. Emotional Content had three modalities, we contrasted the data from neutral images to those from positive images and those from negative images. This model has Emotional Content (Neutral vs Positive; Neutral vs. Negative), Image Level (Context vs. Object) and their interactions as predictors.

We then carried out a second analysis from which we excluded neutral images because they were always presented in a semantically and emotionally coherent condition. The GLMMs and LMMs for this analysis had thus Emotional Content (Positive vs. Negative), Emotional Consistency (Consistent vs. Inconsistent), Semantic Consistency (Consistent vs. Inconsistent), Image Level (Context vs. Object) and their interactions as predictors.

For each model, for significant effects ($p < .05$) we report the predictors' coefficients (β -values), the *SE*-values, the *z*-values (GLMMs) or the *t*-values (LMMs), and the associated *p*-values. *P*-values for the effects in LMMs were generated using the `lmerTest` library

(Kuznetsova, Brockhoff, & Christensen, 2017). For effects showing a tendency toward significance ($.05 \leq p < 0.1$), we reported z - (for GLMMs) or t - (for LMMs) and p -values but not β - and SE -values. Graphics were created using the `ggplot2` package (Wickham, 2009). Follow-up models were conducted in the case of a significant interaction.

Experiment 1: Semantic categorization

Method

Participants

Thirty-two participants (19 females), 19-36 years old ($M=24.41$, $SD=4.92$) volunteered for no remuneration, giving informed written consent before participating. All reported normal or corrected-to-normal vision and no history of neurological troubles. They all presented normal scores at several questionnaires used to ascertain the absence of autism, depression and anxiety: the French translation (Braun & Kempnaers, 2007) of the Autism-Spectrum Quotient (Baron-Cohen, Wheelwright, Skinner, Martin & Clubley, 2001), the French translation (Pichot & Lempérière, 1964) of the short form of the Beck Depression Inventory (Beck, 2006) and the French-Canadian adaptation of the revised version of Spielberger's State-Trait Anxiety Inventory (Gauthier & Bouchard, 1993). This study was approved by the LAPCOS of the Université Côte d'Azur and carried out in accordance with the Declaration of Helsinki.

Procedure

Participants had to categorize semantically one image of the pair, the context or the object depending of the task instruction. They had to respond as accurately and quickly as possible. Half of the participants had to categorise the context image (natural vs. man-made), the other half the object image (human, animal or inanimate object). Responses were given by pressing corresponding keys ("L" and "S" for the context images, "L", "S" and spacebar the object images) on an Azerty keyboard (the response keys were counterbalanced between participants). Participants used two index fingers, one on each button (S for the left hand and L for the right hand) and they used their two thumb fingers for the spacebar. Each experimental image was presented only once. All experimental factors, except for the Image Level to be categorized (whether context or object images), and the visual field (left vs. right) of image presentation were counterbalanced within- and between participants. The Image Level to be categorized was counterbalanced between participants.

Results

We first analyzed only semantically and emotionally consistent trials in order to provide a baseline (Figure 3). Concerning accuracy, the analyses showed a significant effect of Emotional Content when comparing neutral images with negative ones, $\beta=-1.22$, $SE=.21$, $z=-5.93$, $p<.001$, with an advantage for neutral images ($M=0.92$, $SE=0.01$) compared to negative images ($M=0.83$, $SE=0.01$). A tendency toward a significant effect of the Image Level, $z=1.71$, $p=.088$, was shown. The effect of Image Level differed significantly between neutral and negative images, $\beta=-.92$, $SE=.39$, $z=-2.35$, $p=.019$. Neutral objects ($M=0.95$, $SE<0.01$) were categorized with higher accuracy than neutral contexts ($M=0.88$, $SE=0.01$), $\beta=1.18$, $SE=.44$, $z=2.70$, $p<.001$, although no significant differences were found between the categorization of the negative objects ($M=0.84$, $SE=0.02$) and the negative contexts ($M=0.82$, $SE=0.02$), $z<1$, $p=.610$. We did not find other main or interaction significant effects (all $z<1$, all $ps \geq .384$).

Concerning response times, we reported a significant main effect of the Emotional Content, $\beta=.03$, $SE=.01$, $t=3.41$, $p=.001$, as response time for neutral images was shorter ($M=819$ ms, $SE=5$) than for negative images ($M=876$ ms, $SE=17$). No other main or interaction effects were shown (all $ts \leq 1.64$, all $ps \geq .101$).

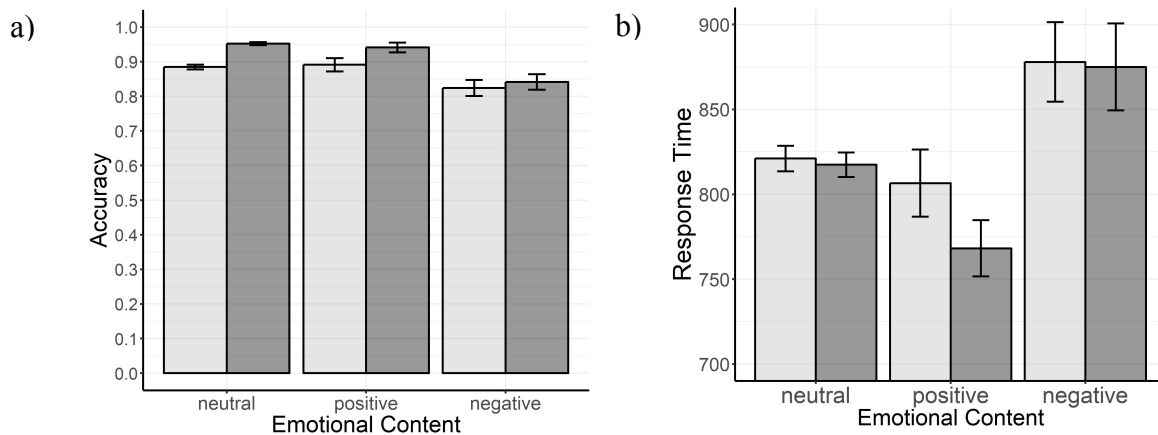


Figure 3: Categorization accuracy a) and response time b) as a function of Emotional Content (Neutral vs. Positive vs. Negative) and of Image Level (Context vs. Object). Error bars indicate one standard error around the mean.

In order to analyse the effects of semantic and emotional consistency, we then excluded the neutral trials. Concerning accuracy, we found a significant main effect for Emotional

Content, $\beta = -.85$, $SE = .16$, $z = -5.21$, $p < .001$, as positive images were categorized more accurately ($M = 0.90$, $SE = 0.007$) than negative ones ($M = 0.83$, $SE = 0.008$). A tendency toward a significant effect of Image Level, $z = 1.69$, $p = .091$, and a significant interaction between Image Level and Emotional Content were shown, $\beta = -.76$, $SE = .31$, $z = -2.43$, $p = .015$. Positive contexts ($M = 0.87$, $SE = 0.01$) were categorized less accurately than positive objects ($M = 0.93$, $SE = 0.008$), $\beta = -1.03$, $SE = .43$, $z = -2.42$, $p = .016$, although no significant differences were found between negative contexts ($M = 0.82$, $SE = 0.01$) and negative objects ($M = 0.84$, $SE = 0.01$), $z < 1$, $p = .503$. We did not find any other main or interaction significant effects (all $z_s \leq 1.36$, all $p_s \geq .174$).

Concerning response times we found a main effect of Semantic Consistency, $\beta = 0.01$, $SE = .005$, $t = 2.55$, $p = .012$ as semantically consistent trials were categorized faster ($M = 825$ ms, $SE = 8$) than semantically inconsistent ones ($M = 856$ ms, $SE = 8$). Semantic Consistency interacted with Image Level, $\beta = -.03$, $SE = .009$, $t = -2.78$, $p = .006$. Response times were slower for semantically inconsistent contexts ($M = 901$ ms, $SE = 12$) compared to faster for semantically consistent ones ($M = 841$ ms, $SE = 11$), $\beta = .03$, $SE = .007$, $t = 3.69$, $p < .001$, but response times for objects were not different depending on semantic consistency, $t < 1$, $p = .951$. We also found a main effect of Emotional Content, $\beta = .03$, $SE = .006$, $t = 5.47$, $p < .001$, with shorter response times for the positive images ($M = 808$ ms, $SE = 7$) than for the negative ones ($M = 875$ ms, $SE = 8$). We did not find any other main or interaction effects (all $t_s \leq 1.50$, all $p_s \geq 0.134$).

Discussion

This first experiment showed that semantic categorization of visual scenes is efficient (accuracy higher than 80%) even with a fast and lateralized presentation. The brief visual perception of an object and a context allowed their parallel processing across the visual field and the access to their semantic representations in long-term memory. Performance was higher for object images than for context images, except for negative images in which performance for contexts and objects was comparable. The difference of a forced choice between three possible responses for the semantic categorization of the object, compared to two responses for context images, thus, did not disadvantage objects categorization.

Interactions between the processing of context and object images were shown: response times were shorter when the two images were semantically consistent than when they were semantically inconsistent. The two images were thus processed in parallel and interaction occurred between the two processing. When the two images refer to the same semantic

representation, semantic categorization is improved, allowing faster responses than when the two images refer to different semantic representations, generating conflicts. This effect was shown mainly in favour of the context; the object, being already integrated in a context, perhaps also benefits from a semantic consistency effect within the target image.

Finally, this experiment highlights an involuntary processing of the emotional content of the target image. Compared to neutral and positive content, the presence of negative content impaired the semantic categorization of the image, decreasing accuracy and increasing response times. Moreover, this effect was greater for objects than for contexts. It has, indeed, been highlighted a prioritized processing of living emotional stimuli, which could potentially be more dangerous for the individual than non-living stimuli (Mermillod, Droit-Volet, Devaux, Schaefer, & Vermeulen, 2010). Our result suggests that there is competition for the allocation of attentional resources between processing of information relevant for the task (semantic information) and processing of emotional (irrelevant) information, which leads to an impairment in the completion of the ongoing task (Pessoa, 2009). However, it is important to note that the impact of emotion content occurred only for the emotional content of the target image. Indeed, the emotional content of the associated image did not influence the semantic categorization of the target image, and did not modulate processing of the emotional content of the target image. This result suggests, that although involuntary, processing of emotional content depends on attentional resources.

Experiment 2: Emotional categorization

Method

Participants

Thirty-two individuals (18 females), aged 19 to 40 years ($M=24.47$, $SD=5.48$) volunteered for no remuneration, giving written consent before participating. All reported normal or corrected-to-normal vision and no history of neurological troubles. They all presented normal scores at the same questionnaires used in Experiment 1 to ascertain the absence of autism, depression and anxiety. This study was approved by the LAPCOS of the Université Côte d'Azur and carried out in accordance with the Declaration of Helsinki.

Procedure

The procedure of Experiment 2 was identical to that of Experiment 1, except for the task: here the participants had to categorize the emotional content (neutral, positive, negative) of one image of the pair.

Results

We first analyzed only semantically and emotionally consistent trials in order to provide a baseline (Figure 4). Concerning accuracy, the analyses show a significant effect of Emotional Content when comparing neutral images with positive ones, $\beta=1.12$, $SE=.18$, $z=6.31$, $p<.001$, with an advantage for positive images ($M=0.78$, $SE=0.02$) compared to neutral images ($M=0.61$, $SE=0.01$). However, this effect interacted with Image Level, $\beta=-.76$, $SE=.32$, $z=-2.36$, $p=.018$. Accuracy for neutral images differed depending on Image Level, $\beta=.81$, $SE=.41$, $z=1.97$, $p=.049$, as objects ($M=0.68$, $SE=0.01$) were categorized with higher accuracy than contexts ($M=0.55$, $SE=0.01$), although no significant difference was found between the categorization of positive objects ($M=0.78$, $SE=0.03$) and contexts ($M=0.78$, $SE=0.03$), $z<1$, $p=.854$. We did not find any other main or interaction significant effects (all $z\leq 1.60$, all $p\geq .109$).

Response times for positive images were shorter ($M=1015\text{ms}$, $SE=22$) than for neutral images ($M=1123\text{ms}$, $SE=11$), $\beta=-.06$, $SE=.009$, $t=-6.55$, $p<.001$. A significant main effect of Image Level, $\beta=-.09$, $SE=.035$, $t=-2.46$, $p=.02$ was shown. Image Level interacted with Emotional Content ($\beta=-.07$, $SE=.017$, $t=3.94$, $p<.001$ when comparing neutral with positive images; $\beta=-.04$, $SE=.02$, $t=-2.11$, $p=.036$, when comparing neutral with negative images). For neutral images, $\beta=-.10$, $SE=.03$, $t=-2.73$, $p=.010$, and negative images, $\beta=-.14$, $SE=.04$, $t=-3.55$, $p<.001$, response times for objects ($M=1004\text{ms}$, $SE=13$ vs. $M=937\text{ms}$, $SE=24$, respectively) were faster than for contexts ($M=1276\text{ms}$, $SE=18$ vs. $M=1306\text{ms}$, $SE=50$, respectively), although that was not the case for positive images ($M=961\text{ms}$, $SE=25$ for objects vs $M=1068\text{ms}$, $SE=36$ for contexts), $t<1$, $p=.483$.

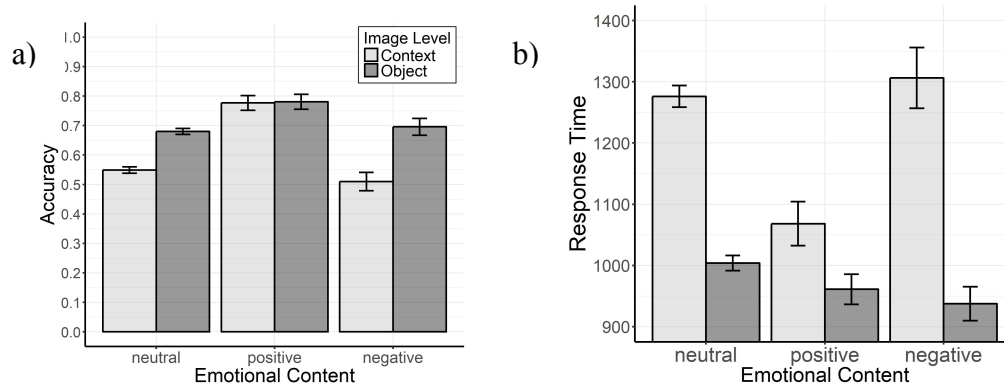


Figure 4: Categorization accuracy a) and response time b) as a function of Emotional Content (Neutral vs. Positive vs. Negative) and of Image Level (Context vs. Object). Error bars indicate one standard error around the mean.

We then excluded neutral trials to analyze semantic and emotional consistency.

Concerning accuracy we found significant main effects of Emotional Content, $\beta = -.82$, $SE = .08$, $z = -10.30$, $p < .001$, Emotional Consistency, $\beta = -.68$, $SE = .07$, $z = -9.62$, $p < .001$, Semantic Consistency, $\beta = -.18$, $SE = .07$, $z = -2.52$, $p = .012$, and Image Level, $\beta = .84$, $SE = .21$, $z = 3.99$, $p < .001$. Emotional Consistency interacted with Image Level, $\beta = .55$, $SE = .14$, $z = 3.9$, $p < .001$, but follow-up models showed that accuracy was higher for emotionally consistent trials compared to inconsistent ones, when categorizing either contexts or objects (all $z \geq 2.24$, all $p \leq .025$), with a higher advantage for contexts. Image Level also interacted with Emotional Content, $\beta = .94$, $SE = .16$, $z = 5.95$, $p < .001$, and with Emotional Content and Semantic Consistency, $\beta = .60$, $SE = .28$, $z = 2.10$, $p = .036$. In order to analyze these three-way interaction, we considered data for context images and object images separately. We found that accuracy improved for semantically consistent trials compared to semantically inconsistent ones, only for negatives contexts images ($M = 0.44$, $SE = 0.02$ vs. $M = 0.33$, $SE = 0.02$ respectively), $\beta = -.48$, $SE = .22$, $z = -2.17$, $p = .003$

Concerning response times we found a main effect of Emotional Content, $\beta = .05$, $SE = .006$, $t = 8.91$, $p < .001$ as positive images were categorized faster ($M = 1040$ ms, $SE = 13$) than negatives ones ($M = 1125$ ms, $SE = 16$). A main effect of Image Level was shown, $\beta = -.07$, $SE = .03$, $t = -2.08$, $p = .047$. Image Level interacted with Emotional Content, $\beta = -.07$, $SE = .01$, $t = -5.41$, $p < .001$: positive contexts ($M = 1111$ ms, $SE = 21$) were categorized faster than negative contexts ($M = 1294$ ms, $SE = 30$), $\beta = .09$, $SE = .01$, $t = 9.05$, $p < .001$, but response times for positive and negative objects ($M = 977$ ms, $SE = 15$ vs. $M = 1031$ ms, $SE = 17$, respectively) were

not significantly different, $t=1.18$, $p= .237$. We did not find other significant main effect or interaction (all $ts \leq 1.64$, all $ps \geq 0.111$).

Discussion

In this second experiment, we studied the emotional categorization of objects and contexts of visual scenes. This experiment showed that the amount of information acquired during a fast and lateralised image presentation was sufficient to evaluate the emotional content of an image and classify it according to an emotional criterion (i.e., positive, negative or neutral).

Our results suggest a facilitation of categorization of positive emotional images versus negative and neutral images. Moreover, for negative and neutral images, accuracy for emotional categorization of the object was higher than that for the context, which was not the case for positive images. Positive images were categorized more quickly and accurately than neutral and negative images independently of the image level. For some authors, when a stimulus is perceived, a rapid processing of its emotional content is achieved: In the case of detection of positive (safe) content, there would be no need for more specific processing, whereas detection of negative content, and thus potential danger, would require additional, more detailed processing (Ihssen & Keil, 2013). This would explain the faster and less costly processing of positive images compared to neutral and negative ones.

This experiment also highlighted an effect of emotional consistency without interaction with the level of the image to be categorized. This result suggests that contextual emotional information improve the emotional categorization of the object when these two levels carry the same emotional content compared to a situation of emotional inconsistency, but also and more important, that the emotional content of the object improved the emotional categorization of the context in a situation of emotional consistency.

Finally, we reported an effect of semantic consistency only for images of negative contexts, which were also categorized emotionally with the lowest accuracy. It seems, thus, that semantic consistency improved performance the most difficult images to categorize.

General Discussion

The aim of this study was to analyse the semantic and emotional processes, and their possible interactions, at work during perception of complex visual scenes that were presented briefly and lateralized. To this purpose, we proposed two different tasks involving the same

experimental material consisting of images with positive or negative emotional content and images without emotional content (i.e., neutral images). In these two tasks, the target image was presented simultaneously with another image, each in a different visual field. The semantic consistency and the emotional consistency between these two images were manipulated in the same way in both tasks for emotional images (i.e. positive and negative). This study, for the first time to our knowledge, presented pairs of emotional images in which the semantic and the emotional consistency were both manipulated. Our results suggest that 1) the relationship between emotional contents is evaluated even when those contents are conveyed in separate images seen just in a glance outside the fovea; 2) emotional processing depends more on attentional allocation than semantic processing; 3) emotional and semantic processing occur whatever the task but semantic consistency and emotional consistency never interact and seems to be independently processed.

Emotional consistency between two briefly images across the visual field

This study shows that the categorization of the emotional content of the target image is facilitated by the presence of an associated image with the same emotional content. We thus replicated the facilitating effect of emotional consistency (de Borst & de Gelder, 2016; de Gelder et al., 2006; Kret & de Gelder, 2010; Righart & de Gelder, 2008) with various stimuli.

The emotional consistency effect occurs even when the amount of information perceived is limited by the time and space constraints of our methodology. This suggests that when two images appear simultaneously at different locations in the visual field, the observer is able to extract their emotional content and relate them in order to improve understanding of the environment. In fact, this process has been described to improve the understanding of the general meaning of the scene, particularly by reducing ambiguities (Aviezer, Ensenberg, & Hassin, 2017; Aviezer, Trope, & Todorov, 2012; Kayyal, Widen & Russel, 2015; Wieser & Brosch, 2012).

Another important aspect is that the emotional consistency effect was shown for the categorization of both the object and the context. That is, the emotional content of a context facilitates the semantic categorization of an object if they carried the same emotional content, compared to a condition of emotional inconsistency; but in an innovative way this study shows that the emotional content of the object also facilitates the emotional categorization of the context. Emotional information of the object and of the context, therefore, seems to interact to facilitate their mutual recognition, as described for semantic categorization.

Attentional competition between two emotional stimuli

An important aspect of this study is the lack of emotional consistency effect in the semantic task. Indeed, while the emotional content of the target image interferes with its semantic categorisation (i.e., impaired performance for negative images compared to positive ones), the emotional content of the co-occurring image has no effect. Thus, although participants cannot know in advance in which of the two possible locations the target image will appear, and although the control of the ocular fixation (by eye-movement recording) ensures that they receive information from both images outside their foveal vision, only the emotional content of the target image affects semantic categorization. It seems that the emotional consistency effect occurs only when emotional processing task-relevant (i.e., emotional categorization task). In addition, the present study suggests that processing of emotional information seems to depend on the attentional resources available. In the semantic categorization task, emotional content is processed involuntarily only for the target image that benefits of the participant's attention. This result suggests a hierarchy of processing within the elements present according to the individual's expectations and the resources that remain available for stimuli-driven process. The involuntary emotional process of the target image is privileged over that of the co-occurring image since the participant's attention is oriented towards the target image. Indeed, when the task is to semantically categorize the image, attention is focused on the semantic content of the target image and the emotional content of that image, but not of the co-occurring one, is processed. On the other hand, the voluntary categorization of the emotional content of an image allocates the participant's attention to the emotional content and thus seems to improve the processing of all the emotional information present in order to help decision-making. Emotional processing seems, thus, to depend on attentional allocation.

In the emotional categorization task, the emotional consistency effect was found independently of the emotional content of the image to categorize (i.e., positive or negative). Indeed, negative images did not impair positive image evaluation more than positive images impaired evaluation of the negative images. This is in line with the attentional model derived from appraisal theories of emotion (Brosch, Scherer, Grandjean, & Sander, 2013; Sander, Grandjean, & Scherer, 2005), for which the focus of attention may be on the stimulus relevant for the action to be carried out regardless of its emotional (positive or negative) content.

Interplay between semantic and emotional processing

In the present study we aimed to analyze the possible interaction between semantic and emotional processing. Our results show that whatever the task proposed, semantic and emotional processing are always carried out. Indeed, during the semantic categorization task, emotional content of the target image, although not relevant for the task, is processed. During the emotional categorization task, the influence of semantic consistency between the two co-occurring images was shown on the emotional categorization of negative contexts, which are the images with the lowest emotional categorization performance. It seems, therefore, that emotional categorization, at least for negative contexts, occurs after semantic categorization. This result might be an argument in favour of the Cognitive Primacy Hypothesis (Lazarus, 1984; Storbeck et al., 2006) according to which emotional processing requires semantic processing. This result may also fit with the theories suggesting that the primacy of one level of processing over another is not pre-established but depends on constraints. For example, it is possible that emotional categorization occurs independently of semantic processing for the majority of the images and that semantic categorization may occur just for more complex images to decrease their ambiguity and improve their identification (Mueller et al., 2017).

The effects of semantic consistency in the semantic task and that of emotional consistency in the emotional task suggest interactions between the two co-occurring images to facilitate their respective categorization. The sharing of information between the context and the object, thus improves visual processing. However, although semantic and emotional processing take place regardless of the task, we did not find any interaction between semantic and emotional consistency. The processes at work in these two consistency effects, therefore, seem independent. This result suggests that the emotional consistency effect does not arise from and/or reflect a semantic consistency effect. This thus seems to disagree with the Cognitive Primacy Hypothesis.

Conclusion

Many studies have highlighted prioritization of the processing of the emotional content, as it attracts exogenous attention and interferes with the performance of the current task. However, our study suggests that this effect of prioritizing emotional content is modulated by endogenous attentional allocation. It also shows that semantic and emotional processing occur during visual scene processing to improve visual efficiency regardless of the task (semantic or emotional). Lastly, we found that emotional categorization depends on semantic processing, at least for the most difficult images for inferring an emotional content. However, semantic

consistency and emotional consistency never interact and thus seems to be independently processed. Our results, thus, do not agree with the Affective Primacy Hypothesis (e.g., LeDoux, 2003; Zajonc, 1984), nor with the Cognitive Primacy Hypothesis (e.g., Lazarus, 1984; Storbeck et al., 2006), and may then argue for the hypothesis of differential affective or cognitive primacy processing depending on the task constraints.

References

- Barrett, L. F., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1521), 1325–1334.
- Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology*, *41*, 167–218.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Rcpp, L. (2015). Package ‘lme4’. *Convergence*, *12*(1).
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276–298. <https://doi.org/10.1037/1528-3542.1.3.276>
- Calvo, M. G., & Nummenmaa, L. (2015). Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion*, 1–26. <https://doi.org/10.1080/02699931.2015.1049124>
- Calvo, M. G., Rodríguez-Chinea, S., & Fernández-Martín, A. (2015). Lateralized discrimination of emotional scenes in peripheral vision. *Experimental Brain Research*, *233*(3), 997–1006. <https://doi.org/10.1007/s00221-014-4174-8>
- Carniglia, E., Caputi, M., Manfredi, V., Zambarbieri, D., & Pessa, E. (2006). The influence of emotional picture thematic content on exploratory eye movements. *Journal of Eye Movement Research*, *5*(4), 1–9.
- Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(4), 1228–1258.
- D’Hondt, F., Lassonde, M., Collignon, O., Lepore, F., Honoré, J., & Sequeira, H. (2013). “Emotions Guide Us”: Behavioral and MEG correlates. *Cortex*, *49*(9), 2473–2483. <https://doi.org/10.1016/j.cortex.2012.12.013>
- D’Hondt, F., Szaffarczyk, S., Sequeira, H., & Boucart, M. (2016). Explicit and implicit emotional processing in peripheral vision: A saccadic choice paradigm. *Biological Psychology*, *119*, 91–100. <https://doi.org/10.1016/J.BIOPSYCHO.2016.07.014>
- Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*(8), 559–564. <https://doi.org/10.1111/j.0956-7976.2004.00719.x>
- de Borst, A. W., & de Gelder, B. (2016). Clear signals or mixed messages: inter-individual emotion congruency modulates brain activity underlying affective body perception.

- Social Cognitive and Affective Neuroscience*, 11(8), 1299–1309.
<https://doi.org/10.1093/scan/nsw039>
- Ellsworth, P. C., & Scherer, K. R. (2003). *Appraisal processes in emotion*. Handbook of Affective Sciences, 572–595.
- Fernández-Martín, A., & Calvo, M. G. (2016). Selective orienting to pleasant versus unpleasant visual scenes. *Cognition*, 155, 108–112.
<https://doi.org/10.1016/J.COGNITION.2016.06.010>
- Hietanen, J. K., & Astikainen, P. (2013). N170 response to facial expressions is modulated by the affective congruency between the emotional expression and preceding affective picture. *Biological Psychology*, 92(2), 114–124.
- Ihssen, N., & Keil, A. (2013). Accelerative and decelerative effects of hedonic valence and emotional arousal during visual scene processing. *The Quarterly Journal of Experimental Psychology*, 66(7), 1276–1301.
- Joubert, O. R., Rousselet, G. a., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, 47(26), 3286–3297.
<https://doi.org/10.1016/j.visres.2007.09.013>
- Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relation of spatial, object, and attraction effects in visual attention. *Frontiers in Psychology*, 1(12).
- Kret, M. E., & de Gelder, B. (2010). Social context influences recognition of bodily expressions. *Experimental Brain Research*, 203(1), 169–180.
<https://doi.org/10.1007/s00221-010-2220-8>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1–26.
<https://doi.org/10.18637/jss.v082.i13>
- Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General*, 144(2), 339–365.
<https://doi.org/http://dx.doi.org/10.1037/xge0000040>
- Lai, V. T., Hagoort, P., & Casasanto, D. (2012). Affective Primacy vs. Cognitive Primacy: Dissolving the Debate. *Frontiers in Psychology*, 3, 243.
<https://doi.org/10.3389/fpsyg.2012.00243>

- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. (Technical). Univ. of Florida, Gaines, FL, US.
- Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist*, *39*(2), 124–129. <https://doi.org/10.1037/0003-066X.39.2.124>
- LeDoux, J. (2003). The Emotional Brain, Fear, and the Amygdala. *Cellular and Molecular Neurobiology*, *23*(4/5), 727–738. <https://doi.org/10.1023/A:1025048802629>
- Marchewka, A., Zurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*(2), 596–610. <https://doi.org/10.3758/s13428-013-0379-1>
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. <https://doi.org/10.3758/s13428-011-0168-7>
- Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, *102*(45), 16518–16523. <https://doi.org/10.1073/pnas.0507650102>
- Mermillod, M., Droit-Volet, S., Devaux, D., Schaefer, A., & Vermeulen, N. (2010). Are Coarse Scales Sufficient for Fast Detection of Visual Threat? *Psychological Science*, *21*(10), 1429–1437. <https://doi.org/10.1177/0956797610381503>
- Mueller, C. J., Fritsch, N., Hofmann, M. J., & Kuchinke, L. (2017). Differences in the dynamics of affective and cognitive processing – An ERP study. *Brain Research*, *1655*, 41–47. <https://doi.org/10.1016/J.BRAINRES.2016.11.018>
- Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Visual Cognition*, *25*(4–6), 563–574. <https://doi.org/10.1080/13506285.2017.1301615>
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion*, *6*(2), 257–268. <https://doi.org/10.1037/1528-3542.6.2.257>
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2010). Semantic categorization precedes affective evaluation of visual scenes. *Journal of Experimental Psychology: General*, *139*(2), 222–246. <https://doi.org/10.1037/a0018858>

- Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computational Vision*, 42, 145–175.
- Padmala, S., Sambuco, N., Codispoti, M., & Pessoa, L. (2018). Attentional Capture by Simultaneous Pleasant and Unpleasant Emotional Distractors. *Emotion*, (March). <https://doi.org/10.1037/emo0000401>
- Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, 13(4), 160–166. <https://doi.org/10.1016/j.tics.2009.01.006>
- Pessoa, Luiz, Kastner, S., & Ungerleider, L. G. (2002). Attentional control of the processing of neural and emotional stimuli. *Brain Research. Cognitive Brain Research*, 15(1), 31–45. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12433381>
- Pilarczyk, Joanna; Kunięcki, M. (2014). Emotional content of an image attracts attention more than visually salient features in various signal-to-noise ratio conditions. *Journal of Vision*, 14(2014), 1–19. <https://doi.org/10.1167/14.12.4.doi>
- Reschke, P. J., Knothe, J. M., Lopez, L. D., & Walle, E. A. (2018). Putting “context” in context: The effects of body posture and emotion scene on adult categorizations of disgust facial expressions. *Emotion*, 18(1), 153–158. <https://doi.org/10.1037/emo0000350>
- Righart, R., & de Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience*, 8(3), 264–272. <https://doi.org/10.3758/CABN.8.3.264>
- Sabatinelli, D., Fortune, E. E., Li, Q., Siddiqui, A., Krafft, C., Oliver, W. T. ... Jeffries, J. (2011). Emotional perception: Meta-analyses of face and natural scene processing. *NeuroImage*, 54(3), 2524–2533. <https://doi.org/10.1016/j.neuroimage.2010.10.011>
- Sander, D., Grandjean, D., & Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Networks*, 18(4), 317–352. <https://doi.org/10.1016/J.NEUNET.2005.03.001>
- Storbeck, J., Robinson, M. D., & McCourt, M. E. (2006). Semantic Processing Precedes Affect Retrieval: The Neurological Case for Cognitive Primacy in Visual Processing. *Review of General Psychology*, 10(1), 41–55. <https://doi.org/10.1037/1089-2680.10.1.41>
- Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, 11(10), 697–709. <https://doi.org/10.1038/nrn2889>
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.

- Vuilleumier, P. (2015). Affective and motivational control of vision. *Current Opinion in Neurology*, 28(1), 29–35.
- Vuilleumier, Patrik. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, 9(12), 585–594.
<https://doi.org/10.1016/j.tics.2005.10.011>
- Wickham, H. (2009). *Elegant graphics for data analysis* (ggplot2).
- Xiao, J., Hays, J., Ehinger, K., Oliva, A., & Torralba, A. (2010). SUN Database: Large-scale Scene Recognition from Abbey to Zoo. *In IEEE Conference on Computer Vision and Pattern Recognition*.
- Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist*, 39(2), 117–123.
<https://doi.org/10.1037/0003-066X.39.2.117>

VI.5 Discussion intermédiaire

Cette étude avait pour objectif d'analyser les traitements sémantiques et émotionnels, ainsi que leurs éventuelles interactions, lors de la perception de scènes visuelles complexes présentées brièvement et latéralisées. Pour ce faire, nous avons proposé deux tâches différentes impliquant un même matériel expérimental constitué d'un objet et d'un contexte pour lesquels, pour la première fois à notre connaissance, la cohérence sémantique et la cohérence émotionnelle ont été simultanément et orthogonalement manipulées. Cette étude a permis de dégager plusieurs résultats importants.

Premièrement nous avons mis en évidence des temps de réponse plus courts pour la catégorisation des images de type contexte présentées en condition de cohérence sémantique par rapport à une condition d'incohérence sémantique. Nous répliquons ainsi l'effet de cohérence sémantique lors de la présentation séparée de l'objet et du contexte tel que mis en évidence dans notre étude du chapitre V (article IV).

Ensuite cette étude montre, pour la première fois à notre connaissance, que l'effet de cohérence émotionnelle décrit dans la littérature pour la présentation visuelle centrale d'un objet intégré dans un contexte est mis en évidence lors d'une présentation visuelle rapide et séparée de l'objet et du contexte. Ceci suggère le traitement en parallèle de deux contenus émotionnels distincts qui sont associés et qui influent sur la catégorisation émotionnelle d'une des deux images. En outre, cet effet de cohérence émotionnelle est mis en évidence quel que soit le niveau (objet ou contexte) de l'image à catégoriser. Ainsi la présence d'un contexte cohérent émotionnellement avec l'objet facilite la catégorisation émotionnelle de celui-ci et réciproquement, la présence du contenu émotionnel de l'objet facilite la catégorisation émotionnelle du contexte lorsque ces deux niveaux sont porteurs du même contenu émotionnel par rapport à des contenus différents.

Notre étude suggère également une interaction entre les traitements sémantiques et émotionnels de l'image cible. Les traitements sémantiques et émotionnels se réalisent conjointement, même de manière involontaire, quelle que soit la tâche demandée au participant (tâche de catégorisation sémantique ou émotionnelle). D'une part, la catégorisation sémantique des images à contenu émotionnel négatif présente une exactitude moindre et un temps de réponse allongé comparativement aux images à contenu émotionnel positif et aux images neutres. Le contenu émotionnel négatif semble ainsi être traité de façon non volontaire lors de la tâche de catégorisation sémantique. D'autre part, lors de la tâche de catégorisation du contenu émotionnel, la mise en commun des contenus sémantiques de l'objet et du contexte semble

venir aider la catégorisation émotionnelle des images dont le jugement sur le plan émotionnel semble plus complexe. Ainsi, les images de contextes négatives sont les images qui présentent la plus faible exactitude à la tâche de catégorisation du contenu émotionnel et ce sont les seules pour lesquelles il est statistiquement mis en évidence un effet de la cohérence sémantique. Nous pouvons faire l'hypothèse que lorsque l'observateur est en difficulté pour évaluer le contenu émotionnel d'une image il va prendre en considération l'ensemble des éléments présents dans son champ visuel afin d'améliorer sa compréhension du sens de la scène visuelle pour soutenir son jugement.

Enfin les effets de cohérence sémantique d'une part et émotionnelle d'autre part sont mis en évidence indépendamment, ce qui évoque que les associations objet-contexte se réalisent sur des informations de nature différente pour chaque effet de cohérence. Nos résultats sont ainsi en faveur de traitements sémantiques et émotionnels indépendants qui se dérouleraient en même temps et en constante interaction en fonction des contraintes. D'une part, la détection d'un contenu émotionnel réduit les ressources allouées à la tâche de catégorisation sémantique et d'autre part un traitement sémantique paraît parfois nécessaire pour que l'observateur puisse évaluer délibérément le contenu émotionnel d'une image, au moins lorsque le contenu émotionnel semble difficile à catégoriser. Dans ce dernier cas, nous pouvons supposer que la reconnaissance sémantique de l'image à juger permette la mise en lien avec l'ensemble des représentations reliées à cette image afin de favoriser son traitement émotionnel. Par exemple une image qui représente un cimetière, suivant l'observateur, pourrait ne pas entraîner de ressenti émotionnel, l'accès à son sens et aux concepts associés (e.g., mort, deuil) aidera alors à évaluer l'image comme négative.

Chapitre VII : Etude des traitements sémantiques et émotionnels lors de la perception rapide de scènes visuelles complexes chez l'enfant d'âge scolaire

Ce chapitre présente la troisième étude empirique de ce travail de thèse qui analyse les traitements sémantiques et émotionnels lors de la perception rapide de scènes visuelles dans le développement précoce, c'est-à-dire chez l'enfant scolarisé en primaire, âgé de 8 à 12 ans. Ce chapitre est constitué d'un rappel des principaux apports théoriques, d'une présentation de la méthodologie de l'étude, de son résumé en français, de la reproduction de l'article en préparation et enfin d'une discussion intermédiaire.

Leroy, A., Spotorno, S. & Faure, S. (in preparation). Semantic and emotional visual scene processing in school-aged children.

VII.1 Rappels des apports théoriques principaux

Les travaux réalisés dans le cadre de cette thèse chez l'adulte (articles IV et V) nous ont permis d'affiner la compréhension du déroulement des traitements sémantiques et émotionnels ainsi que celle des effets d'association objet-contexte au sein des scènes visuelles. Nous pouvons ainsi retenir que lors de la perception d'une scène visuelle nous accédons rapidement à un ensemble d'informations approximatives portées par les différents constituants de la scène (contexte et objet) et qui permettent une compréhension du sens général de la scène grâce à une mise en correspondance avec des représentations en mémoire à long terme. Des prédictions sur d'éventuels autres éléments de la scène sont alors générées permettant à l'individu de s'adapter rapidement à l'environnement. Outre l'extraction d'informations de nature sémantique, des informations de nature émotionnelle sont également extraites; ces dernières peuvent être traitées de façon volontaire ou non. Le traitement non volontaire du contenu émotionnel mobilise des ressources attentionnelles et ainsi interfère avec la réalisation de la tâche en cours (et ce de façon plus marquée lors de la détection d'éléments potentiellement menaçants). Dans la tâche de catégorisation sémantique, présentée dans l'article V, cette interférence entraînait une diminution de l'exactitude et une augmentation des temps de réponse pour les images à contenu négatif comparativement aux images neutres ou positives. A la différence de ce traitement non volontaire (*bottom-up*) du contenu émotionnel, son jugement délibéré, c'est-à-dire la catégorisation volontaire (*top-down*) du contenu émotionnel, semble nécessiter plus d'étapes de traitement et être plus coûteux que la catégorisation sémantique. Ces résultats pourraient suggérer que des traitements sémantiques soient une étape nécessaire à la catégorisation émotionnelle, au moins pour les images les plus difficiles à traiter d'un point de vue émotionnel.

Pour résumer, il semble que lors de la catégorisation sémantique d'une image d'araignée, par exemple, le traitement du contenu émotionnel négatif de peur de cette image va influencer la performance de catégorisation de façon rapide et automatique, et sans que l'individu ait forcément conscience d'avoir détecté une menace. Alors que pour juger l'émotion que lui procure cette image, il aura parfois besoin de la reconnaître (donc de la catégoriser sémantiquement) pour ensuite pouvoir la catégoriser émotionnellement en tant que contenu négatif. Nous reviendrons plus en détail sur le déroulement de ces traitements dans la discussion générale de la thèse.

Nous souhaitons maintenant étudier ces mêmes capacités, de catégorisation sémantique et émotionnelle, auprès d'enfants d'âge scolaire afin de documenter certaines dynamiques développementales du traitement de scènes visuelles complexes. De notre seconde revue de la littérature présentée dans la partie théorique de ce manuscrit (article II), nous retenons que la capacité de détection automatique du contenu émotionnel (surtout négatif) est présente très tôt dans le développement de l'enfant, alors que la capacité à catégoriser volontairement une image selon la valence émotionnelle ou l'émotion de base se développe jusqu'à la fin de l'adolescence. Au niveau sémantique, il semble que la compréhension du sens général de la scène (le *gist*) est possible dès 15 mois si le temps d'exploration est assez long pour le permettre ou lorsque la saillance perceptive de certains constituants de la scène attire l'attention de l'enfant sur une zone importante pour la compréhension de la scène (Duh & Wang, 2014). La prise en compte de l'ensemble des informations d'une scène visuelle (contexte et objets) semble présente tôt dans le développement de l'enfant, dès 4 mois (Bornstein et al., 2011). Néanmoins, l'effet de cohérence sémantique a été peu étudié chez l'enfant : les quelques travaux réalisés suggèrent un effet de l'âge et un rôle des expériences dans le développement des soubassements cognitifs de cet effet (Thomas et al., 2017). Cette synthèse (article II) nous a également permis de mettre en avant que très peu d'études se sont intéressées aux mécanismes de catégorisation sémantique et émotionnelle de scènes visuelles complexes chez l'enfant d'âge scolaire. Cependant, l'analyse du développement de ces différentes dimensions de traitement des informations visuelles est fondamentale pour comprendre comment l'enfant perçoit et représente le monde visuel qui l'entoure. La présente étude a ainsi pour but de mieux déterminer les capacités de traitement des informations sémantiques et émotionnelles visuelles chez l'enfant de huit à 12 ans, aspect primordial dans les processus d'apprentissage et pour l'adaptation au monde physique et social. Cette recherche permettra également de renforcer la compréhension des

processus à l'œuvre durant la perception visuelle par l'étude du développement des capacités de catégorisation sémantique et émotionnelle. Enfin une meilleure compréhension des processus à l'œuvre lors de la perception de scènes visuelles dans le développement typique permettra une caractérisation plus fine des dysfonctionnements dans le cas de pathologies.

VII.2 Présentation de l'étude

Méthodologie expérimentale :

L'étude reprend la même méthodologie que celle du chapitre VI (article V) sur les interactions entre les traitements sémantiques et émotionnels chez l'adulte, avec quelques modifications pour l'adapter à un échantillon d'enfants.

Comme pour l'étude chez les adultes, les enfants doivent juger un objet ou un contexte présenté rapidement (100 ms) au sein d'une paire d'images. Chaque image est latéralisée, une dans le champ visuel droit, l'autre dans le champ visuel gauche, hors de la vision fovéale. La moitié des enfants réalise une tâche de catégorisation sémantique de l'image, l'autre moitié catégorise son contenu émotionnel. La différence de méthodologie avec les tâches de l'étude chez les adultes (article V) est l'absence, ici, de manipulation simultanée des cohérences sémantique et émotionnelle. Dans la tâche de catégorisation sémantique de la présente étude, toutes les paires d'images sont présentées en situation de cohérence émotionnelle et donc seule la cohérence sémantique est manipulée. Dans la tâche de jugement émotionnel, toutes les paires d'images sont présentées en situation de cohérence sémantique et donc seule la cohérence émotionnelle est manipulée. Les mouvements oculaires n'ont pas été enregistrés.

Le matériel expérimental est le même que celui utilisé chez les adultes (article V) à l'exception de 13 images qui ont été remplacées sur les 576 utilisées. Comme décrit dans le chapitre VII, section 6.2, chez l'adulte, les images proviennent de deux banques d'images (NAPS, IAPS) validées selon la valence émotionnelle et le niveau d'activation physiologique et, majoritairement pour la condition neutre, de la banque d'images SUN ou d'images libres de droit sur internet (Google Images). Dans ces bases, nous avons d'abord sélectionné des images à faible niveau d'activation physiologique afin d'écarter les images les plus choquantes (e.g. à caractère érotique, de mutilation et/ou provoquant un fort dégoût). Nous avons ensuite fait évaluer chaque image quant à sa valence émotionnelle par 16 adultes et 6 enfants âgés de 7 à 9 ans (différents de ceux qui ont été inclus dans le protocole). Cette étude pilote a montré que la variable « valence » était correctement opérationnalisée et ce de façon très similaire chez les adultes et les enfants ($\rho = 0,83$; $p < .001$). Enfin, afin d'écarter toute image présentant un

caractère choquant, nous avons, dans une deuxième étude pilote, demandé à cinq couples de parents d'enfants âgés de 7 à 9 ans d'évaluer chaque image sur cette dimension. Nous avons fixé comme critère qu'il suffisait qu'un seul juge sur les 10 estime que l'image est potentiellement choquante pour un enfant âgé entre 7 et 9 ans pour supprimer cette image de notre matériel expérimental : seules 13 images sur 576 (soit 2,26%) ont été retirées. L'utilisation de bases d'images validées et employées consensuellement dans les recherches internationales sur les traitements émotionnels auprès de participants de tous âges, combinée à nos études pilotes, garantit la validité du matériel et minimise autant que possible les risques éthiques liés à un potentiel ressenti émotionnel négatif chez certains participants.

Cette étude a été validée par le comité d'éthique pour les recherches non interventionnelles (CERNI) de l'Université Nice Côte d'Azur (numéro d'attribution 2018-2, cf. Annexe 1).

Modalités de recrutement et de passation:

Le recrutement et les passations ont été réalisés sur le temps scolaire, dans des écoles primaires du Département des Alpes Maritimes (ville de Cagnes sur Mer) suite aux accords de l'inspecteur de l'académie, du médecin conseiller technique et de l'inspecteur de l'éducation nationale de la circonscription de Cagnes-sur-Mer, et sur le temps périscolaire en accord avec la mairie de Nice. Le recrutement des enfants s'est fait sur la base du volontariat des familles. Une première lettre d'information générale présentant l'étude (cf. Annexe 2) ainsi qu'un formulaire d'inclusion (cf. Annexe 3) ont été transmis à toutes les familles d'enfants scolarisés du CE1 au CM2. Les familles souhaitant plus d'informations sur l'étude devaient retourner sous enveloppe cachetée le formulaire d'inclusion. Une seconde enveloppe leur était alors transmise comportant la description détaillée de l'étude (cf. Annexe 4), les formulaires de consentement parents (cf. Annexe 5) et un questionnaire « points forts-points faibles » (questionnaire SDQ version parent, cf. Annexe 6) pour évaluer la présence de troubles du comportement, de troubles scolaires ou de troubles psycho-affectifs chez leur enfant. Après avoir pris connaissance des détails de l'étude, si la famille ne souhaitait plus que son enfant participe, elle n'avait rien à préciser. Dans le cas où elle était intéressée par l'étude et souhaitait que son enfant participe, elle devait transmettre les formulaires de consentement signés et le questionnaire SDQ-Parent rempli.

Les enfants pour lesquels les formulaires de consentement « parents » signés nous ont été retournés étaient vus directement au sein de l'école, en accord avec leur enseignant pour ne

pas gêner le déroulement de leur scolarité, ou sur le temps périscolaire. Les enfants recevaient alors à leur tour un document leur présentant l'étude, celle-ci leur était également expliquée à l'oral et ils pouvaient demander des précisions à l'expérimentateur avant de donner leur accord par la signature du formulaire de consentement enfant (cf. Annexe 7). Les consignes des tâches expérimentales ont été modifiées à partir de celles destinées aux adultes (article V) et adaptées pour être comprises par des enfants d'âge scolaire.

Critères d'inclusion et d'exclusion :

Les critères d'inclusion pour l'étude sont :

- Enfants (sexe masculin et féminin) âgés de 8 ans 0 mois à 12 ans 11 mois ;
- Acuité visuelle normale ou corrigée à la normale ;
- Formulaire de consentement éclairé signé par au moins l'un des deux parents ou un titulaire de l'autorité parentale.

Critères de non inclusion :

- Acuité visuelle anormale limitant la vision ou trouble de la perception des couleurs ;
- Antécédents neurologiques ou d'épilepsie ; enfant présentant un trouble du spectre autistique, un TDA/H ou un retard mental (confirmé par le formulaire d'inclusion renseigné par les parents, cf. Annexe 3).

Hypothèses expérimentales

Nous faisons l'hypothèse que certains processus décrits chez l'adulte sont déjà fonctionnels plus tôt dans le développement et peuvent être mis en évidence chez des enfants de 8 à 12 ans. En conséquence nous devrions répliquer les effets de cohérence sémantique et de cohérence émotionnelle entre l'objet et le contexte d'une scène visuelle et l'effet d'interférence du contenu émotionnel négatif sur les processus de catégorisation sémantique.

Cette étude est la première, à notre connaissance, à proposer à des enfants d'âge scolaire une tâche de catégorisation sémantique et une tâche de catégorisation émotionnelle, sur les mêmes images, lors d'une présentation visuelle rapide en champ visuel divisé. Batty et Taylor (2002) avaient proposé une tâche de détection d'un animal lors d'une présentation visuelle centrale rapide (80 ms) et mis en évidence chez les enfants une exactitude de détection, dès l'âge de huit ans, comparable à celle des adultes. Nous souhaitons analyser les capacités de catégorisation des contenus sémantique et émotionnel d'enfants de 8 à 12 ans lors d'une présentation simultanée de deux images hors de la vision fovéale. Les participants ne sachant

pas dans quel champ visuel l'image cible apparaîtra, doivent traiter, au moins initialement, le contenu des deux images. Nous analysons ainsi les capacités de traitement en parallèle de deux images et la capacité à accéder à leurs contenus sémantiques et émotionnels alors que le temps de présentation est très court et ne permet pas d'exploration de la scène et donc de prise d'information plus précise.

VII.3 Résumé de l'étude

L'enfant doit pouvoir catégoriser rapidement, au niveau sémantique et émotionnel, un nouvel élément qui apparaît dans son environnement visuel. Ces capacités sont en effet primordiales pour son adaptation constante à un environnement changeant. Néanmoins, très peu d'études se sont intéressées à leurs développements, malgré leur importance dans le quotidien, notamment scolaire et social. Nous avons ainsi proposé à des enfants âgés de huit à douze ans une tâche de catégorisation sémantique ou de catégorisation du contenu émotionnel d'une image neutre, positive ou négative, présentée parmi une paire, rapidement et hors de la vision fovéale. De plus, les associations sémantiques entre les deux images de la paire (un objet et un contexte) étaient manipulées dans la tâche sémantique et les associations émotionnelles dans la tâche émotionnelle. Nos résultats montrent que les enfants présentent des capacités de catégorisation sémantique et émotionnelle supérieures au niveau hasard, à part pour les images négatives de type contexte. Les performances à la tâche de catégorisation sémantique sont supérieures à celles de catégorisation du contenu émotionnel, mais diminuées par la présence de contenu émotionnel négatif évoquant un traitement implicite, priorisé du contenu émotionnel. Enfin, la manipulation des associations entre l'objet et le contexte n'a pas d'influence au niveau sémantique alors que c'est le cas pour les informations émotionnelles. Au total cette étude va dans le sens d'un renforcement des associations sémantiques qui se réaliserait au cours du développement et des expériences de vie alors que les associations émotionnelles au sein des scènes visuelles seraient opérantes plus tôt dans le développement, probablement en lien avec un traitement plus implicite, automatique du contenu émotionnel.

VII.4 Article VI. Semantic and emotional visual scene processing in school-aged children.

Le texte ci-dessous est la reproduction d'un article en préparation. Une vérification du niveau d'anglais sera réalisée avant soumission.

Leroy, A., Faure, S. & Spotorno, S. (in preparation). Semantic and emotional visual scene processing in school-aged children.

Semantic and emotional categorization of visual scenes in children

Anaïs Leroy^{a,*}, Sylvane Faure^a, Sara Spotorno^b

^a Laboratoire d'Anthropologie et de Psychologie Cliniques, Cognitives et Sociales (LAPCOS), Université Côte d'Azur, France.

^b School of Psychology, University of Keele, United Kingdom

*corresponding author:

Anaïs LEROY

LAPCOS, MSHS Sud Est

Université Côte d'Azur

Pôle Universitaire Saint Jean d'Angély

24 avenue des Diables Bleus

06357 Nice Cedex 4 (France)

Telephone: +33 (0) 4 89 88 14 60

Email: anais.leroy@univ-cotedazur.fr

Abstract

Despite their importance in daily living, being particularly crucial for the child to adapt in a constantly changing environment, categorisation along the semantic and emotional dimensions of visual scenes have been little studied in the school-age years. Two behavioral experiments assessed semantic (experiment 1) and emotional categorization (experiment 2), as well as the mutual influences of context-object semantic (experiment 1) and emotional (experiment 2) contents in 64 eight to twelve-years-old children, using the same paradigm and material. At each trial, participants had to categorize a scene context or an object briefly presented separately in simultaneous images, one in the right visual field and the other in the left. The semantic relatedness between the two images was manipulated in the semantic categorization task (experiment 1), as was their emotional valence consistency in the emotional categorization task (experiment 2). Results showed that children performed the two tasks above chance, except when context-level images carried a negative emotion. These results suggest that rapid parallel visual processing of complex entries is efficient from the age of eight, allowing access to high-level processes and representations. Interestingly, semantic categorization was more proficient than emotion categorization, but appeared limited by the presence of a negative emotional content, suggesting involuntary access to emotions conveyed by the input, which in turn would necessitate an additional processing that uses cognitive resources at the expense of the semantic decision. The study also illustrates, for the first time, that the modulation of emotion categorization by emotional consistency occurs for the two image levels (i.e; object and context), while it failed to replicate the semantic consistency advantage that characterises semantic categorization in adults. This pattern of findings underpins the view of a slowly progressive strengthening of semantic associations on the basis of accumulated knowledge in the course of development and through life experiences, while by contrast the processing of emotional relationships within a scene would mature earlier, probably beginning with an early or innate drive to perceive and react to emotions, in a more automatic, rapid and spontaneous way.

Key words: children; emotions; semantics; categories; object-context;

Highlights

- Primary school-aged children categorize efficiently and both on the semantic and emotional dimensions images of common objects and contexts.
- Semantic categorization is better performed than emotional categorization.
- For the first time, modulation of the emotional categorisation of the context of the scene by the emotional content of the object, i.e. an emotion consistency advantage, is evidenced, whereas the robust semantic consistency advantage found in adult research is not evidenced.
- The study enriches literature on differences in the early development of semantic and emotion visual processing as well as emotion-cognition interactions more generally.

Introduction

In day-to-day life, children are constantly surrounded by a multitude of visual stimuli that they must be able to select and process quickly in order to extract patterns of information, make predictions and respond and adapt optimally to the changing constraints of their physical and social environment. This speed of processing is made possible by prioritizing certain information and categorizing it. Prioritization depends both on involuntary processes, when the characteristics of the stimuli are prominent and/or relevant enough to attract exogenous attention (bottom-up process), and on voluntary processes (endogenous allocation of attention) directed by the observer's expectations or goals (Corbetta & Shulman, 2002). These two modes are necessary for adaptation and survival, interact continuously during visual perception (Corbetta, Patel, & Shulman, 2008) and must be well balanced. Thus, when the child is engaged in an activity, he or she focuses on the task-relevant stimuli (top-down process). If a prominent stimulus appears in his or her visual field, such as a snake, although task-irrelevant, it will be processed involuntarily and this may even suspend the current activity if necessary. The saliency of the stimulus may depend in particular on perceptual (e.g., color, brightness, contrast, etc.) or emotional characteristics (for a review see: Zadra & Clore, 2011). Several studies carried out in adults (Öhman, Flykt, & Esteves, 2001; for reviews see: Pourtois, Schettino, & Vuilleumier, 2012; Vuilleumier, 2015) and in children (LoBue, 2009; LoBue & DeLoache, 2010; Vaish, Grossmann, & Woodward, 2008) have described a prioritization of stimuli with emotional content versus neutral stimuli (i.e., without emotional content).

The selected elements will then be categorized. Categorizing allows viewers to process large amounts of information, while minimizing cognitive costs. Indeed, to avoid that the multitude of information that the child perceives, as well as that available to him in long-term memory, overloads his cognitive capacities, the child will categorize, i.e. group together different elements sharing a certain specificity in the same category (Harnad, 2017; Mareschal & Quinn, 2001). Thus, whatever their shapes, colours and location (on a desk, on a bedstand or in the street), the child will be able to group a desk lamp, a bedside lamp and a street light in the "lamp" category and thus infer that each of them can be used to illuminate. Similarly, when the child sees someone crying, he or she may infer that the person is sad, and adapt his or her behaviour. The categories thus formed make it possible to process information more quickly, transfer knowledge of a specimen from one category to another and communicate more easily. Categorization can be done according to perceptual (e.g., shape, color), semantic (e.g.,

taxonomic or thematic characteristics) or emotional (e.g., this person is happy or sad) attributes (Brosch, Pourtois, & Sander, 2010; Harnad, 1987; Rosch, 1978).

Although these prioritization and categorization processes have been studied extensively in adults, the study of their development during childhood has been neglected. In particular, very few studies have focused on categorization of semantic and emotional information during rapid visual perception in elementary school-aged children, a process that is central to learning (O'Sullivan, 1996) and to the development of social skills (see for a comprehensive view, Happé & Frith, 2014).

Semantic categorization

Many adult research studies have highlighted the ability to semantically categorize, within a few hundred milliseconds (ms), images of isolated objects as well as of complex scenes (Rousselet, Joubert, & Fabre-Thorpe, 2005; Thorpe, Fize, & Marlot, 1996; for a review see, Fabre-Thorpe, 2011) that characterize everyday visual world. These scenes are composed of multiple objects, organized within a background (the scene's context) following semantic and spatial rules, which define the plausibility of object occurrence in the scene and of the scene's layout, respectively. Object-context associations are at the core of the concept of "scene" and a key, omnipresent aspect in our visual experience. This is the case both on the semantic-categorical and on the emotional dimensions. However, it is still unclear how object-context associations influence visual information processing, namely in which conditions and at what stages they operate. In children research, the categorization process has been mainly studied during the first two years of life. This literature describes an evolution from a categorization based essentially on perceptual characteristics to the increasingly importance of semantic characteristics and a progressive refinement of categories (for reviews see: Poulin-Dubois & Pauen, 2017; Quinn & Johnson, 2000). Although categorization processes develop from early childhood throughout adolescence and into early adulthood (Bova et al., 2007), and despite the importance of categorization in building other cognitive functions such as memory and reasoning (O'Sullivan, 1996), only two studies to our knowledge have examined semantic categorization skills in a rapid visual presentation in school-aged children and adolescents. In the Batty and Taylor's (2002) study, participants aged seven to 15 years were asked to detect the presence of an animal in an image presented centrally and very briefly (80ms). Their results suggested similar categorization capabilities in terms of accuracy (although associated with longer response times) to those of adults. The study by Vanmarcke and Wagemans (2015),

which also proposed an animal detection task but at an even faster visual presentation (33 ms) to adolescents aged 11 to 16 years, again shows similar abilities to those obtained by adults. These studies therefore have suggested that from the school period (from seven years of age) children would exhibit semantic categorization abilities, at least in terms of accuracy, that are essentially identical to those of adults. However, tasks consisted in detecting an object (presence or not of an animal) presented centrally and did not require the categorization of an entire visual scene. By contrast in his daily life, the child must take into account multiple stimuli simultaneously present in different regions of his or her visual field and be able to process in parallel and integrate several distinct levels of information. On this topic, it should be noted that in Vanmarcke and Wagemans's (2015) study, the detection of the presence of an animal was better when the image represented the target included in a context compared to a condition without background.

The importance of the association between local level information (objects) and global level information (context) in the rapid categorization process is widely recognized in the adult literature. Numerous studies (Davenport & Potter, 2004; Joubert, Fize, Rousselet, & Fabre-Thorpe, 2008; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Munneke, Brentari, & Peelen, 2013) have indeed shown that semantic categorisation performance was better when both levels (i.e, context and object) of a visual scene were presented in a condition of semantic consistency (i.e., according to an association frequently encountered in everyday life as a toaster in the kitchen) compared to semantic inconsistency (i.e., a toaster in a bathroom). At the developmental level, some studies suggest, particularly through the recording of eye movements or evoked potentials, that from the early years of life human beings are able to detect the presence of a semantic inconsistency (Bornstein, Mash, & Arterberry, 2011; Helo, Azaiez, & Rämä, 2017; Helo, van Ommen, Pannasch, Danteny-Dordoigne, & Rämä, 2017). As far as we know, these effects of semantic consistency between the different levels of a visual scene (context and object) have not been studied at school age. Recently, Thomas, Nardini, and Mareschal (2017) used a Stroop-like paradigm in children aged six to nine years and suggested that the speed advantage of semantic consistency between an animal image and a sound develops with age and is reinforced with life experiences and as learning progress.

Emotional categorization

The child thus builds and groups information from his or her environment into perceptual or semantic categories but also into emotional categories, requiring recognition of

emotional content (Brosch et al., 2010). The categorization of emotional content is gradually refined until late adolescence to allow a more precise differentiation of the type of emotion felt, expressed by others or inferred from a situation (Conte, Brenna, Ricciardelli, & Turati, 2018; McManis, Bradley, Berg, Cuthbert, & Lang, 2001; Vicari, Reilly, Pasqualetti, Vizzotto, & Caltagirone, 2000; Widen, 2013; Widen & Russell, 2008).

The association between the emotional content of the different levels of a visual scene (i.e., context and objects) influences their categorization: performance in adults is facilitated by emotional consistency. Specifically, when the context and objects present in a visual scene carry the same emotional content (i.e., emotional consistency situation), their emotional categorization is facilitated in comparison to a situation of emotional inconsistency where each level would convey different content (Kret & de Gelder, 2010; Nelson & Mondloch, 2017; Righart & de Gelder, 2008). The ability to detect an emotional inconsistency between the emotions carried by face and by body attitudes (e.g., a happy face on a body in an attitude of fear) has been described very early in development, from the first year of life (Heck, Chroust, White, Jubran, & Bhatt, 2018; Rajhans, Jessen, Missana, & Grossmann, 2016). The facilitating effect of emotional consistency, for the categorization of the emotional expression of a face when it is presented with a context or body in a bodily attitude carrying the same emotional content, has been shown from the age of five (Mondloch, 2012; Mondloch, Horner, & Mian, 2013; Theurel et al., 2016). For example, children categorize the facial emotion "fear" more accurately when the two levels convey the same fearful content, e.g. a face expressing fear presented in a fearful context such as a storm at sea versus the same face presented in a joyful context such as a heavenly beach. However, to our knowledge, no study has studied the reciprocal, i.e. the effect of the emotional content of an object on the emotional categorization of a context. In the previous example, would the emotional categorization of the context - a storm at sea as fearful emotion - be facilitated when the face included in the image expresses fear compared to when it expresses joy? In addition, similarly to studies on semantic categorization in children, the stimuli of the studies described above were presented in central vision and/or displayed on the screen for a period of time long enough to allow visual exploration and thus the viewer to fixate various components of the scene. These methodologies make it difficult to understand what information carried by visual scenes is perceived and processed instantly, before any more detailed voluntary exploration and analysis. However, the fast coarse-grain initial processing of information that appears in any region of the visual field is essential to interpret, and adapt instantly to, a changing environment.

This study

The purpose of this study is to analyze how school-age children process semantic and emotional information carried by local (object) and global (contextual information) level elements that quickly appear outside the foveal vision, before a more precise analysis can begin. We intended to study which elements will be processed prioritally, in isolation or taking into account another component of the scene to improve the categorization process (consistency effects). These initial processings are essential, albeit among others, as they make it possible to evaluate the relevance of the stimulus for the.

At each trial, participants aged eight to twelve years old had to categorize a scene context or an object briefly presented separately in simultaneous images (one in the left hemifield, one in the right), along the semantic-categorical dimension (Experiment 1) or along the emotional valence (Experiment 2). To get closer to perception in real-world settings, we used stimuli made up of images that reproduced environments (context images) and prominent objects that form the visual scenes encountered and processed in everyday life. This is in sharp contrast with the stimuli mainly used by child research studies to date: isolated objects for semantic categorisation, or facial expressions for emotion categorization, that also lack contextual information. In addition to allow measuring categorization abilities, these tasks, by differently orienting the attentional focus, may provide access to the indirect and therefore non-voluntary processing of emotional content (during the semantic task that directs attention to semantic content) and its direct and therefore voluntary processing (during the emotion task). Lastly, we examined whether children process and associate information from the various components of the visual scene (i.e., object and context): if so, semantic consistency and/or emotional consistency advantage should be highlighted. As already mentioned, to our knowledge, no study has analyzed the effect of semantic consistency within a visual scene during this developmental period. However, we aimed analyzing whether the child combines contextual semantic information and local semantic information when a stimulus rapidly appears in his or her environment, to activate predictions and improve categorization, as demonstrated in adulthood (Trapp & Bar, 2015). Regarding emotions, child research studies have highlighted a facilitating effect of contextual information on the emotional categorization of the local level but none has, to our knowledge, studied the influence of the emotional content of the object on the emotional categorization of the context. We hypothesised that the emotional consistency advantage is characterized by a mutual interaction between the emotional content from the object and that from the context. We therefore expected that emotional consistency should

facilitate emotional categorization and occur for categorizing the emotional content of the object as well as that of the context.

General Method

Participants

Sixty four children aged eight to twelve years old, with mean age (SD): 9.34 (0.77), 16 different in each experiment (Experiment 1: 16 females, Experiment 2: 19 females), voluntarily participated. Children were recruited in primary schools in the region of Nice (France). *All reported normal or corrected-to-normal visual acuity, normal color vision and no history of neurological disorders.* Exclusion criteria were abnormal visual acuity or colour perception disorder, neurological or epilepsy history, and a diagnosis of autism spectrum disorder or ADHD (two neurodevelopmental disorders where emotional problems are frequently noted, see Sheppes, Suri & Gross, 2015).

This study was approved by the ethics committee for non-interventional research at the Université Côte d'azur (2018-2) and carried out in accordance with the declaration of Helsinki. The informed consent of at least one of the two parents, as well as the child, was required for the child to participate in the study. All participants gave informed written consent before participating.

Stimulus material and design

Three hundred and ninety-six full-color photographs of animate and inanimate objects, plus 24 as practice, and 396 full-color photographs of indoor and outdoor scene's contexts, plus 24 as practice, were used. "Object" images were defined by a unique object (e.g., book, daim). "Context" images were mainly defined by global, spatial information (e.g., office, forest), even though they also contained local information (such as in a set of houses in a street). People, animals or objects that were present in the initial context images were removed using Adobe Photoshop CS 9.0 software (Adobe, San Jose, CA). We also created a white noise mask in Matlab. The computer background was medium grey (RGB: 127, 127, 127).

Half of the images (context and object) were emotionally neutral and the other half carried positive (25%) or negative (25%) emotional valence. The images were selected from the International Affective Picture System (IAPS, Lang, Bradley, & Cuthbert, 2008), the Nencki Affective Picture System (Marchewka, Zurawski, Jednoróg, & Grabowska, 2014), the SUN database (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010) or Google Images. As part of a

pilot study, we submitted the images to independent judges (five couples parents of children aged eight to nine years) for evaluation in order to exclude all those believed shocking in the event of displaying to a child. Another pilot study was conducted with independent judges, 16 adults and 6 children aged 7 to 9 years, who were not participating in the current study, to obtain a French-language emotional categorization of the experimental images. The results showed that the Emotional Content variable was correctly operationalized and in a very similar way in adults and children (the association between scoring from the two samples was strong: $\rho = 0.83$; $p < .001$).




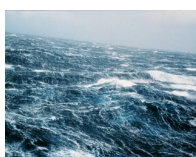

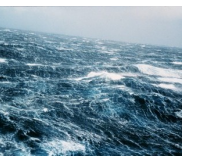


Images were presented in pairs consisting of a context and an object. Neutral object and context images were always presented in a situation of semantic and emotional consistency (e.g., a neutral image was always presented with another neutral image), thus avoiding participants to become habituated to the condition of inconsistency (semantic or emotional) and to the presence of emotional content. First, the original pairs of emotional images were formed so that the two images (context and object) were semantically consistent (i.e., a forest and a snake) and emotionally consistent (i.e., the two images were either of neutral or positive or negative valence). Second, to create a semantically inconsistent (i.e., a forest and a shark) or an emotionally inconsistent (i.e., a positive image associated with a negative image, such as a paradise beach and a shark) pairs, the constituent images of the original pairs were swapped according to the experimental versions (see Figure 1).

Each experimental image was presented only once. In the semantic categorization task (Exp.1), all image pairs were emotionally consistent, only semantic consistency was manipulated. In the emotion categorization task (Exp.2), all pairs of images were semantically consistent, only emotional consistency was manipulated. For this experimental methodology - analysing consistency effects by using two separate images, - we rely on the results of previous adult studies in our team, which have highlighted that the effects of semantic and emotional consistency between two images (an object and a context) are evidenced even for a separate presentation of the two images, each in opposite lateral visual fields (Leroy, Faure and Spotorno, under revision; in preparation).

The experimental factors were Image Level (Context vs. Object), Semantic (Exp.1) or Emotional (Exp.2) Consistency of the image pair (Consistent vs. Inconsistent) and Emotional Content (Neutral vs. Positive vs. Negative) of the target image (image to categorize). All of the experimental factors were counterbalanced within and between participants according to four experimental versions for each task.

Each experiment lasted approximately 25 minutes.

A

a1. Semantic consistency				
a2. Semantic Inconsistency				

B

b1. Emotional Consistency				
b2. Emotional Inconsistency				

Figure 1: Examples of semantically consistent and inconsistent pairs for experiment 1 (A, panels at the top) and emotionally consistent and inconsistent pairs for experiment 2 (B, panels at the bottom). The conditions of semantic (a2) or emotional (b2) inconsistency were obtained by swapping images from two semantically (a1) or emotionally (b1) consistent pairs.

After completing the consent form, each child was randomly assigned one of the four groups: two were attributed the 1st main experimental task and the two others the 2nd main experimental task, as follows. Experimental task 1 (experiment 1) evaluated semantic categorization capabilities through two different instructions, one relating to the context level and the other relating to the object level. Experiment 1a concerned the semantic categorization of the image representing the context within the pair of images: instruction were to decide whether it was a natural or manufactured environment; experiment 1b concerned the semantic categorization of the image representing the object level within the pair of images: instructions were to decide whether it was an animal, a person or an inanimate object. Experimental task 2 (experiment 2) evaluated emotional categorization giving participants the instruction to decide

whether the target image carried a neutral, positive or negative emotional content. Experiment 2a concerned the emotional categorization of the context and experiment 2b the emotional categorization of the object.

The participant sat comfortably in a quiet room in front of the monitor, with his/her head supported by a chin-rest (chin and forehead), so that the distance between his/her eyes and the computer screen (size: $32.5^\circ \times 24.5^\circ$, resolution: 1024×768 pixels, refresh rate: 85 Hz) was 57 cm. At the beginning of the experiment, 12 training trials are carried out with images that were not used for the experimental phase. Each trial (see Figure 2) begins with the display of a central fixation cross for 500 milliseconds (ms), then two experimental images ($10^\circ \times 8^\circ$, 315×252 pixels), a "context" and an "object", were presented simultaneously for 100 ms, one in the right visual field and the other in the left (internal edge of the image at 2.5° from the fixation point), on the horizontal midline of the screen, followed by the visual noise masks for 80 ms and then by the cue, a black arrow ($2.5^\circ \times 1^\circ$, 79×32 pixels) pointing toward the screen's half (right or left) where the image to categorize (of a given level : context or object) had been presented. The participant responded with a forced choice by pressing a key (L, S or space) designated by a sticker on an azerty keyboard. Response assignments were counterbalanced between participants in order to avoid potential spatial stimulus/response compatibility effects. Once the response has been given, a new trial can begin; a total of 288 trials are thus carried out, separated into 4 experimental blocks (72 trials per block). The total duration of the experiment is approximately 45 minutes including three 5-minute breaks between each block.

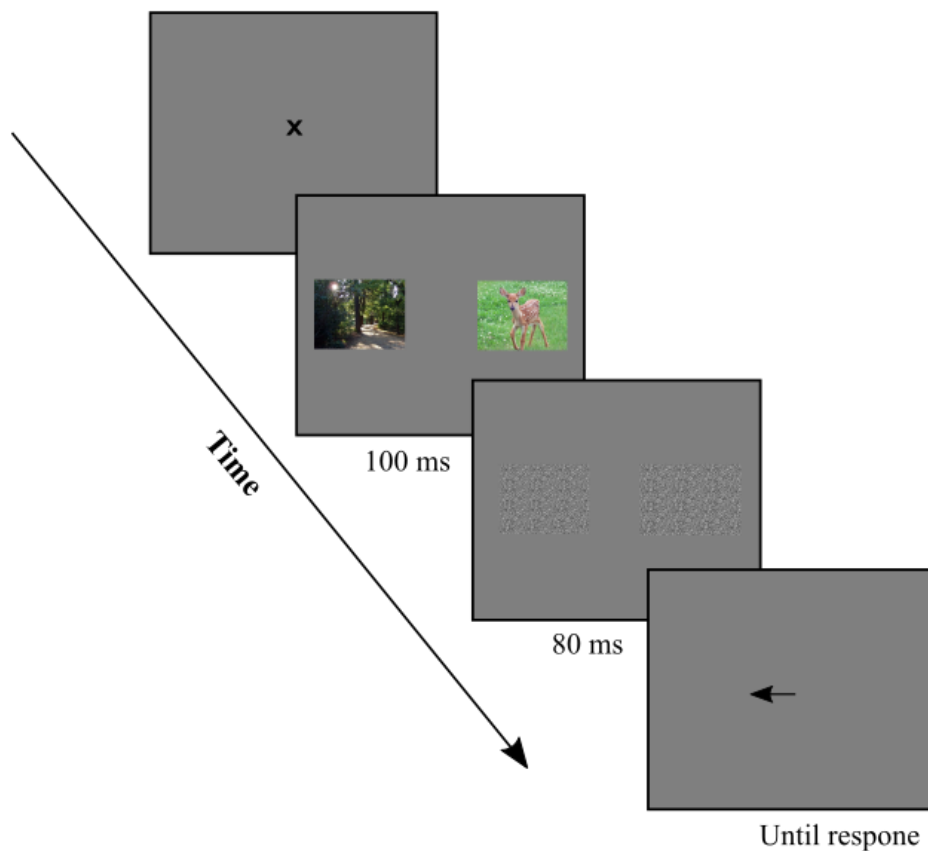


Figure 2: General paradigm and trial procedures. Two images were presented simultaneously to the right and left of a central fixation point followed by two corresponding masks and then an arrow indicating to the participant where the target image (context or object) had been displayed (here, the context had to be categorized). The participant was given different forced-choice response alternatives depending on the task assigned to him or her (semantic categorization or emotional categorization) and had to give a manual response (see text for further details).

Data analysis

Statistical analyses were mainly performed by generalized linear mixed models (GLMMs) on accuracy (proportion of correct response) and by linear mixed models (LMMs) on response times of correct response: RTs). These models allow both subject and item analyses to be performed (Baayen, Vasishth, Kliegl, & Bates, 2017; González, De Boeck, & Tuerlinckx, 2014). We ran several models with fixed effects describing the variables of interest as predictors: Image Level (Context vs. Object), Emotional Content (Positive vs. Negative),

Consistency (Consistent vs. Inconsistent), whether Semantic (Exp. 1) or Emotional (Exp. 2), of the image pair (see relevant sections of the Results that follow) and their interactions. For each experiment, participants and items were specified as random factors.

We first analyzed semantically consistent images. We then excluded trials in which the images were presented in a situation of inconsistency (semantic for experiment 1, emotional for experiment 2) in order to be able to compare data from neutral images (always presented in a situation of semantic and emotional consistency, see in the material section) with those from emotional images (positive and negative). This model thus has as a predictor the Emotional Content (Neutral vs. Positive vs. Negative), the Image Level (Context vs. Object) and their interactions. The Emotional Content variable being in three modes, we contrast the data from neutral to positive images and from neutral to negative images.

Second, neutral images were excluded from the analyses in which the consistency variable (semantic, experiment 1 or emotional, experiment 2) was manipulated. The predictors of our main analyses are the Image Level (Context vs. Object), Emotional Content (Positive vs. Negative), Consistency (Consistent vs. Inconsistent) Semantic (Exp. 1) or Emotional (Exp. 2) of the image pair and their interactions.

For the analysis of significant interactions, follow-up models are performed to examine the differences between the conditions involved. Before the analysis of response times, outliers, i.e. responses given before 300ms and after 5000ms, are excluded (167 trials corresponding to 2.8% of the total trials for experiment 1 and 171 trials or 4.6% of the total trials for experiment 2). All data are analyzed by R software (The R Foundation for Statistical Computing, Version 3.3.2), GLMMs and LMMs using the `lmer()` function of the `lme4` package (Bates et al., 2015) and graphics are created using the `ggplot2` package (Wickham, 2009). For each model we report the coefficients of the predictors (β -values), the *SE* values, the *z*-values for GLMMs and *t*-values for LMMs and the associated *p*-value. We will only analyze significant effects ($p < .05$) and not trends ($.10 > p \geq .05$) for which we will report only *z*(GLMMs) or *t*(LMMs) and the *p*-associated value.

Experiment 1: Semantic categorization

We study the ability to semantically categorize an image presented within a pair. We also analyzed whether semantic categorization performance was modulated by the level of the image (context or object), by the emotional content (neutral vs. positive vs. negative) of the image to

be judged and by the manipulation of semantic relationship with the other image of the pair (by presenting them in condition of semantic consistency vs. inconsistency).

Results:

Analyses focused first on semantically coherent images in order to provide a baseline for main analyses.

Accuracy. Analyses showed an effect of the Image Level, $\beta=1.25$, $SE=.23$, $z=5.44$, $p<.001$ (Fig.3a), with higher accuracy for object ($M=0.92$, $SE=0.005$) over context images ($M=0.81$, $SE=0.007$). The effect of Emotion Content, by the comparison of neutral images with negatives was nearly significant, $z=-1.96$, $p=.050$, which suggests an advantage for neutral images ($M=0.87$, $SE=0.005$) over negatives ($M=0.84$, $SE=0.01$). There were no other significant main or interaction effects (all $z\leq 1.21$, all $p\geq .226$).

Response times. Analyses showed an interaction between the Emotional Content, for the comparison of neutral and negative images, and the Image Level, $\beta=-.08$, $SE=.01$, $t=-5.30$, $p<.001$ (Fig.3b). Analysis of this interaction showed that for neutral images RTs of, context ($M=1616$ ms, $SE=18$) and object ($M=1609$ ms, $SE=16$) did not differ significantly, $\beta=-.003$, $SE=.01$, $t=-0.21$, $p=.827$, while images of negative contexts had longer RTs ($M=1832$ ms, $SE=42$) than images of negative objects ($M=1525$ ms, $SE=33$), $\beta=-.04$, $SE=.01$, $t=-3.06$, $p=.004$. RTs for negative contexts were significantly longer than RTs for neutral contexts and RTs for negative objects were faster than RTs for neutral objects (all $t\geq 2.38$, all $p\leq .017$). The analysis also showed a trend for there to be an effect of Emotional Content, for the comparison between neutral and negative images ($t=1.96$, $p=.051$) and a trend towards an interaction between Emotional Content, when comparing neutral to positive images, and Image Level ($t=1.95$, $p=.052$). There were no other significant main effects or interactions (all $t\leq 1.64$, all $p\geq .112$).

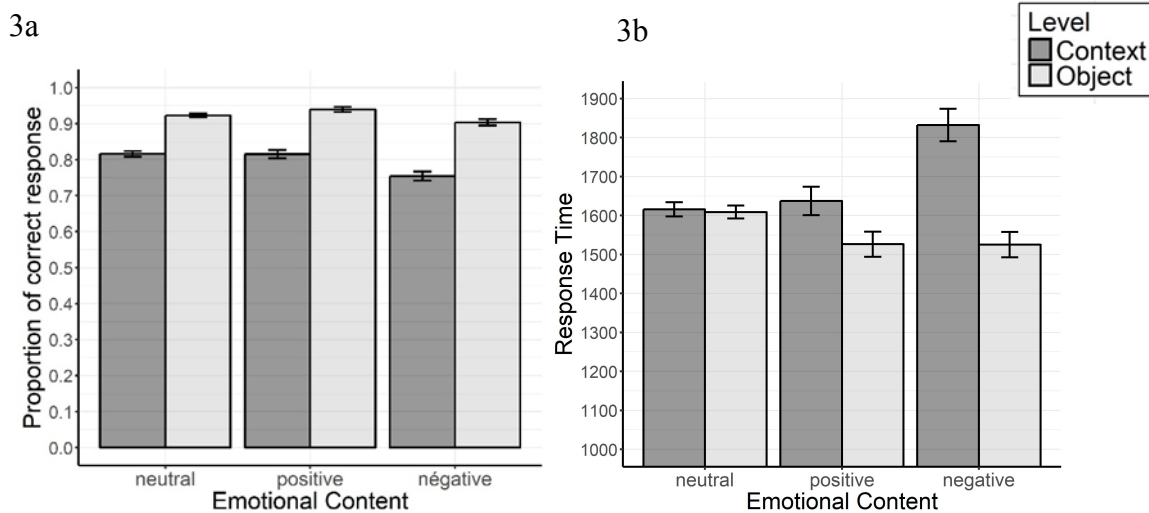


Figure 3: Proportion of correct responses (3a) and response time (3b) for the semantic categorization task according to the Level (context or object) and the Emotional Content of the target image.

We then carried out main analyses.

Accuracy. Analyses showed a significant effect of Emotional Content of the image to be judged, $\beta = -.41$, $SE = .15$, $z = -2.85$, $p = .004$, with an advantage for images with positive emotional content ($M = 0.88$, $SE = 0.007$) compared to negative ones ($M = 0.83$, $SE = 0.008$). There was also a significant effect of Level of the image, $\beta = 1.28$, $SE = .21$, $z = 6.08$, $p < .001$, with a better performance of objects ($M = 0.92$, $SE = 0.006$) than of contexts ($M = 0.78$, $SE = 0.009$). There were no other significant main effects or interactions (all $z \leq 1.52$, all $p \geq .128$).

Response times. Analyses showed significant main effects of Emotional Content, $\beta = .03$, $SE = .008$, $t = 3.20$, $p = .003$, and of Image Level, $\beta = -.06$, $SE = .02$, $t = -2.43$, $p = .021$, as well as a significant interaction between these two variables, $\beta = -.04$, $SE = .02$, $t = -2.73$, $p = .01$. The analysis of the interaction showed that there was no significant difference between the RTs for context ($M = 1650$ ms, $SE = 27$) and object ($M = 1515$ ms, $SE = 23$) images when these images were positive ($t = 1.47$, $p = .151$), while RTs for context ($M = 1824$ ms, $SE = 30$) were significantly longer than for object ($M = 1542$ ms, $SE = 23$) when these images had negative content, $\beta = .08$, $SE = .02$, $t = 3.15$, $p = .003$. No other main or interaction effects were observed (all $t \leq 1.09$, all $p \geq .279$).

Discussion:

This study replicates the results of the rare ones (Batty & Taylor, 2002; Vanmarcke & Wagemans, 2015) that examined the semantic categorization capabilities of school-age children for rapid visual presentation, and allows generalizing findings, for the first time to our knowledge, to a lateralized presentation, outside the foveal and simultaneous vision of two images. Studies so far have highlighted the efficiency, from primary school age, of semantic categorization of an object when the child's attention is focused on this object. We show here that the visual processing of objects and contexts can be done in parallel without the need for focused attention. In addition, these parallel processes are not limited to low-level visual processing but allow access to high-level semantic representations. Our results are thus in line with those described in adults (Rousselet, Fabre-Thorpe, & Thorpe, 2002; Rousselet, Thorpe, & Fabre-Thorpe, 2004) and suggest that these capacities for parallel processing of visual scenes are effective from the age of 8. The results thus highlight the already very efficient semantic categorization capabilities in primary school children (with an average of 85% of images correctly categorized), despite the difficulty of the task (rapid and lateralised presentation of two simultaneous images). The performances highlighted in this study are indeed quite identical in terms of response accuracy (but with longer response times) to adults' ones in a similar study (accuracy $M= 86.5\%$; response time $M= 912$ ms ; Leroy, Faure, & Spotorno, in preparation). Object images (foreground, local level) are categorized more accurately and quickly than context images (background images, representing a more diffuse set of information). This advantage of objects can be explained by a local level processing bias that still predominates over the global level in children aged seven to twelve years (Harrison & Stiles, 2009; Poirel, Mellet, Houdé, & Pineau, 2008). It could also be due to the experimental task itself, the differentiation between, on one hand, an object, an animal or a person being perhaps easier for a primary-school child than the differentiation between, on the other hand, natural and manufactured concepts that are perhaps more abstract, thus requiring a higher level of thinking. Whatever it may be, categorization is very good for both types of images: 92% for objects and 78% for contexts.

We also show that the accuracy of semantic categorization of negative images (regardless of their level: context or object) is lower than that of positive images. The robust effect of interference from indirect and involuntary processing of negative emotional content (Augusti, Torheim, & Melinder, 2014; Cohen-Gilbert & Thomas, 2013; Kramer, Lagattuta, & Sayfan, 2015; Rahman & Wiebe, 2019) is highlighted here with a rapid and lateralised

presentation and in a similar way to findings in adults (Leroy, Faure, & Spotorno, under preparation). These results therefore suggest that negative emotional content, although not relevant to the task, is detected and that its processing requires some of the attention resources allocated to the semantic categorization task. The child therefore process negative emotional information that appears in his or her visual field even if his or her attention is focused on an activity for which the emotional content is not relevant. This reinforces the hypothesis of prioritised, or at least reinforced, processing of negative emotional stimuli over neutral ones. Finally, we wanted to examine whether the semantic content of the other image in the pair would influence the semantic categorization of the target image. No effect of semantic consistency was found, either on speed or accuracy, whereas this effect has been frequently highlighted in adult research studies (Davenport & Potter, 2004; Joubert et al., 2008; Rémy, Vayssière, Pins, Boucart, & Fabre-Thorpe, 2014). This result can be discussed within various frameworks. First, according to some authors, the semantic consistency effect develops as experiences and the repetition of co-occurrences between contextual information and the presence of certain objects cumulate, and as learning progresses (Thomas et al., 2017). Second, it may also be the case that the simultaneous presentation of two images outside the foveal vision and very briefly does not allow school-age children to process accurately and then integrate the information from the two sources. Indeed, this particular stimulation implies that a lower quantity and quality of visual information is perceived compared to a central visual presentation, as in most studies. However, an important limit of this classical-central vision approach is that it is far from natural visual environments: when navigating real-world settings, individuals face often crowded surroundings with many objects and rapidly changing background. Finally, the divided lateral visual field presentation implies that each image was initially addressed to opposite cerebral hemispheres, the contralateral hemisphere to the visual field of stimulation (Hellige, 2001). Sharing information, initially received by each hemisphere, requires structural substrates and functions that are not mature in primary school children. Indeed, several studies have highlighted an increase in inter-hemispheric transfer capacities and thus an improvement in functional connectivity between hemispheres associated with a decrease in information transfer time between hemispheres with age and this very intensively during the period of primary school, with a peak in maturation between ten and twelve years (Banich & Brown, 2000; Chicoine, Proteau, & Lassonde, 2000; Fagard, Hardy-Léger, Kervella, & Marks, 2001; Keshavan et al., 2002).

Experiment 2 : Emotional categorization

This second experiment was designed to study the ability to categorize the emotional content of an image when presented with another one conveying the same emotional content (condition of emotional consistency) or a different one (condition of emotional inconsistency). Experiment 2 uses the same experimental design (see Fig.2) and stimuli as in experiment 1, thus avoiding any bias in the interpretation of results that would be secondary to perceptual differences between the images used. A notable exception is that here all images appeared in a situation of semantic consistency, and emotional consistency only is manipulated. We study the emotional categorization capacities, and the influence of the image level (context or object) as well as the influence of the associated image (emotional consistency effect) on the categorization performance.

Results:

At first we only analyze emotionally consistent images.

Accuracy. Analyses revealed a significant effect of Emotional Content for the comparison of neutral images with positive ones, $\beta=.31$, $SE=.11$, $z=2.96$, $p=.003$, with an advantage for images with positive emotional content ($M=0.62$, $SE=0.01$) on neutral ones ($M=0.55$, $SE=0.01$), as well as a significant effect of Emotional Content for the comparison of neutral images with negative ones, $\beta=-.96$, $SE=.11$, $z=-9.02$, $p<.001$, with a lower performance for negative images ($M=0.36$, $SE=0.01$) (Fig.4a). Analyses also showed a significant effect of Image Level, $\beta=.74$, $SE=.27$, $z=2.73$, $p=.006$, with an advantage for object type images ($M=0.63$, $SE=0.01$) over context images ($M=0.44$, $SE=0.01$). This effect is modulated by an interaction with Emotional Content for the comparison between neutral and positive images, $\beta=-.72$, $SE=.18$, $z=-3.97$, $p<.001$, while there were no interaction with Emotional Content for the comparison between neutral and negative images ($z=1.38$, $p=.169$). The performance of context images ($M=0.59$, $SE=0.02$) and positive objects ($M=0.65$, $SE=0.02$) does not differ, $z<1$, $p=.740$, while context-type images have a lower categorization accuracy ($M=0.45$, $SE=0.01$) than object-type images ($M=0.66$, $SE=0.01$) for neutral images, $\beta=.33$, $SE=.06$, $z=5.28$, $p<.001$.

Response times. Analyses highlighted a significant main effect of the Emotional Content for the comparison of neutral and positive images, $\beta=-.07$, $SE=.008$, $t=-8.28$, $p<.001$, this effect interact with Image Level, $\beta=.04$, $SE=.02$, $t=2.55$, $p=.011$ (Fig.4b). The interaction reflected the greater categorization advantage of positive ($M= 1499$ ms, $SE=26$) over neutral stimuli ($M=$

1797 ms, $SE=18$) whatever the level of the image (all $t \geq 4.06$, all $p \leq .001$) but largely for contexts images. We did not show other main or interaction effects (all $z \leq 1.28$, all $p \geq .201$).

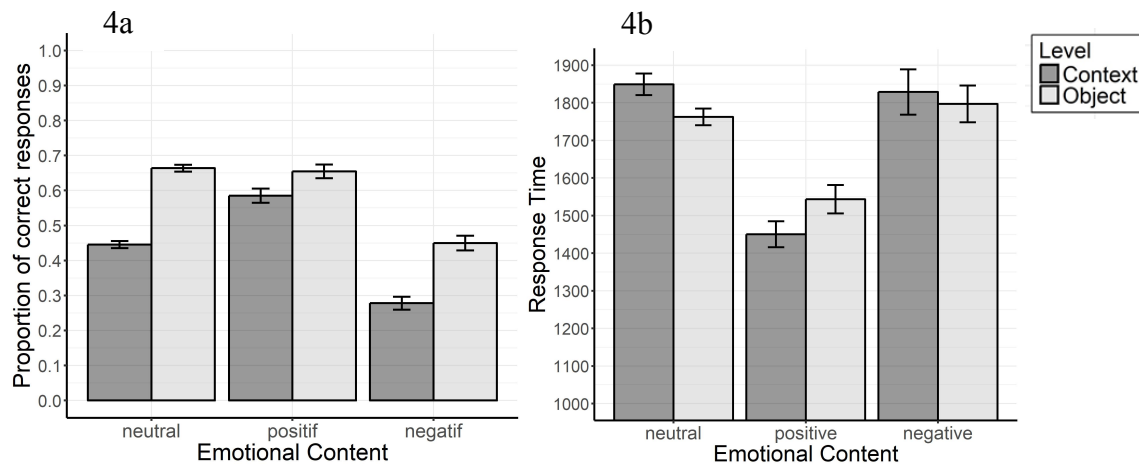


Figure 4: Proportion of correct answers (4a) and response times (4b) for performance in the emotional categorization task according to the Level (context or object) and Emotional Content (positive, negative or neutral) of the target image.

We then carried out main analyses

Accuracy. Analyses, showed a significant effect of the Emotional Content, $\beta=-1.28$, $SE=.26$, $z=-4.92$, $p<.001$, with an advantage for images with positive emotional content ($M=0.60$, $SE=0.01$) compared to negative images ($M=0.36$, $SE=0.01$). As the accuracy of categorization was low for negative images, we compared the data obtained according to the level of the object (context or object) and the emotional content (positive or negative) of the target image with the chance level (which corresponds to 0.33 since 3 forced choice responses are possible). The categorization of positive images was better than chance level for both context ($t(15)=4.97$, $p<.001$) and object ($t(15)=6.42$, $p<.001$) images. The categorization of negative images was not significantly different from chance level for contexts ($t(15)=-1.83$, $p=.957$) but was significantly superior to chance for objects ($t(15)=2.53$, $p=.011$). Assessing the emotional valence appears therefore more difficult for negative context-level images ($M=0.26$, $SE=0.01$) than negative object-level images ($M=0.46$, $SE=0.01$) and more difficult for negative than positive images, whether contexts ($M=0.55$, $SE=0.01$) or objects ($M=0.64$, $SE=0.01$).

Analyses also showed a significant effect of Emotional Consistency, $\beta=-.16$, $SE=.08$, $z=-2.04$, $p=.041$, with performance in the condition of emotional consistency ($M=0.49$, $SE=0.01$) being superior to that of emotional inconsistency ($M=0.46$, $SE=0.01$) and an Image

Level main effect, $\beta=.72$, $SE=.24$, $z=2.94$, $p=.003$, contexts images were categorized with a lower accuracy ($M=0.40$, $SE=0.01$) than objects ones ($M=0.55$, $SE=0.01$). No interaction effects were significant (all $z \leq 1.57$, all $p \geq .117$).

Response times. Analyses revealed a significant main effect of Emotional Content, $\beta=.08$, $SE=.009$, $t=8.87$, $p<.001$, as well as a significant interaction with Image Level, $\beta=-.06$, $SE=-.02$, $t=-3.40$, $p<.001$. There was also a trend toward an interaction between Emotional Content and Emotional Consistency $\beta=-.03$, $SE=.02$, $t=-1.70$, $p=.099$. No other main or interaction effects were found (all $t \leq 1$, all $p \geq .474$).

The analysis of the interaction between Emotional Content and Image Level specified that the RTs for positive images were shorter than for negative images regardless of the image level, but more markedly so for context-type images.

Discussion:

In this second experiment, children had to categorize the emotional content (positive, negative or neutral) of an image presented quickly within a pair. Performance, except for negative context images, was better than chance level, suggesting that school-aged children have the ability to categorize the emotional content of images (objects and contexts) that appears very briefly and laterally. This study suggests that the amount of sensory information gathered during such a glance and outside foveal vision is sufficient to evaluate the emotional content of an image and classify it according to the emotional valence dimension (i.e., positive or negative).

There was an advantage for positive images over negative ones, the first being, moreover, categorized nearly as accurately ($M= 60\%$) as in adults (i.e. $M= 68\%$, as shown by a similar study by Leroy, Faure and Spotorno, in preparation), but with slower responses (mean RTs $M= 1534\text{ms}$ and $M= 1096\text{ms}$, respectively) while the negative images categorization was lower ($M=36\%$ in children vs. $M= 52\%$ in adults). In addition, images of contexts having a negative emotional content appeared to be the most challenging to assess, with a categorization performance which did not differ from chance and a processing time considerably longer than for positive contexts. The lower categorizing performance for negative compared to positive images may be explained by the higher cognitive control required in the presence of negative emotional content, which thus requires more attention resources than positive or neutral images (Pessoa, 2009; Schel & Crone, 2013). For some authors, when a stimulus is perceived, a rapid processing of emotional valence is achieved: in the case of positive valence

detection, synonymous with safety, there would be no need for more specific processing, while the detection of negative valence, and thus potential danger, would require additional, more precise and detailed processing (Ihssen & Keil, 2013). This would explain a faster and less demanding processing for positive stimuli compared to negative ones. Another explanation of this superiority of positive stimuli over negative ones is based on the difference in the number of possible basic emotions included in the categories "positive valence" versus "negative valence". The positive condition includes only joy while the negative condition comprises several basic emotions: i.e., fear, anger, sadness and disgust. Even when the instructions are to categorize emotional content at the superordinate level (i.e., positive, negative or neutral), as in the current study, some findings suggest that basic emotional processing would nevertheless be carried out involuntarily, and the decision would require a longer processing time and a greater amount of perceptual information than for taking a decision at a superordinate level (Aguado, Martínez-García, Solís-Olce, Dieguez-Risco, & Hinojosa, 2018).

This experiment replicates the emotional consistency effect described in the literature on early development (Mondloch, 2012; Mondloch et al., 2013; Nelson & Mondloch, 2017; Theurel et al., 2016) and generalizes it to a rapid and lateralised presentation. This result therefore suggests that the child is able to simultaneously process information from two separate images presented in opposite visual fields (and therefore initially addressed to different cerebral hemispheres), extract the respective emotional content and integrate it. The effect of emotional consistency is obtained without interaction with the level of the image to be categorized, which suggests a facilitating effect of contextual emotional information on the emotional categorization of the object but also a facilitating effect of the emotional content of the object on the emotional categorization of the context when these two levels carry the same emotional value compared to a situation of emotional inconsistency.

General discussion

The general aim of this study was to analyze the abilities for semantic and emotional categorization of visual scenes in primary school children (aged eight to twelve). To do this, an object image (level of local, precise information) and a context image (contextual information, global level, diffuse) were presented simultaneously, quickly and separately to children who had to perform either a semantic categorization task (with manipulation of the semantic consistency between the object and the context) or an emotional categorization task (with manipulation of the emotional consistency). This study is the first, to our knowledge, to use the

same images, which were moreover complex images of real scenes, in order to analyse how semantic consistency and emotional consistency modulate scene understanding in children attending primary school.

The first notable result of this study is that, with the exception of the emotional categorization of negative context-level images, eight to twelve years old children have the ability to categorize, both semantically and emotionally, images presented simultaneously and outside the foveal vision (i.e., tachistoscopic presentation) clearly better than chance, despite a very rapid presentation. This study thus replicates the results obtained for central vision (Batty & Taylor, 2002; Vanmarcke et al., 2016) and extends them to the simultaneous processing of two lateralized images. Our study shows that the parallel visual processing of an object image and a context image, presented for only 100 ms, is efficient from at least eight years old. In addition, the study shows that these parallel processes allow access to high-level representations and thus to the recognition of the semantic or emotional category of the image. These findings provide evidence that not only primary school-aged children can characterize scenes along several essential dimensions, but also that these skills are already proficient as far as the perceptual scene configuration is not present in the visual field but only mentally constructed by the young viewer (Lowe et al., 2018). Furthermore, our study showed that the presence of negative emotional content, although irrelevant to the current task, was detected.

This study also suggests different capacities for processing semantic and emotional dimensions of a visual scene. First, children appear more proficient at semantic categorization than for emotional categorization. This suggests that categorizing emotional content requires additional processing steps, and/or more perceptual information than semantic content. Second, the emotional consistency effect, which implies a pooling of information from the target image and associated image, was highlighted for the emotional categorization task while the semantic consistency effect was not found in the semantic categorization task. Finally, results show an effect of emotional content on the performance of the semantic categorization task, which suggests indirect, involuntary and rapid processing of the emotional component. This therefore put forward the issue of different levels of processing of emotional content depending on whether the processing is carried out through bottom-up, rapid and non-voluntary processes, or through top-down, voluntary processes.

Processing semantic information

This study highlights an ability of school-aged children to semantically categorize the object and context, which is essentially identical to that of adults, but with a longer processing time. Thus, the information gathered during the rapid and lateralised presentation of images is sufficient for the child to access a long-term memory representation and to determine the semantic category to which the target belongs. These capacities are highlighted for object images, therefore at the local level, made up of more precise information, as well as for contextual information, therefore at a more global and diffuse level. The accuracy of categorization of the object and context is high (92% and 78% respectively) but we note a better categorization of the local level compared to the global level. These results are consistent with those obtained with experimental material consisting of geometric shapes or letters and that describe a predominance of processing from the local level over the global level at this age (Harrison & Stiles, 2009; Poirel et al., 2008).

The substantial effect of semantic consistency, frequently described in the adult population, is not observed in this study with children aged eight to twelve years. The suggested lack of sharing of information from the local and global levels could be explained by the use of the divided visual field paradigm, which requires interhemispheric transfer for sharing information presented in different visual fields and therefore addressed to different hemispheres. The corpus callosum is not mature at this age (Keshavan et al., 2002) and the functional connectivity of the corpus callosum, along with the efficiency of interhemispheric transfer, continues to develop until late adolescence (Banich & Brown, 2000). However, the evidence of the effect of emotional consistency in experiment 2 suggests functional inter-hemispheric transfer capacities, at least for emotional information. Another explanation therefore seems more likely, namely that the object-context association effect may emerge later as experiences and learning grows (Thomas et al., 2017). The influence of semantic consistency would thus depend on the strength of the associations, which is due to the frequent co-occurrences between the different components of a visual scene. It is therefore very likely that in experimental tasks where the child has time to detail a visual scene, as well as in his or her daily life, the perception of an single element of the scene makes it possible to activate semantic representations associated with this element and thus guide visual exploration and promote the recognition of the other components of the scene. However, the temporal (rapid presentation) and spatial (lateralized presentation) constraints of our experimental device limit the amount of perceptual information available in a speeded task. To adapt to these demands, it is likely that

children have processed the target image in a more narrowed way than adults, and relied more on its perceptual content. In fact, the processing of perceptual information (i.e., bottom-up processings) during visual exploration has been highlighted as predominant until adolescence, compared to predominant guidance by cognitive processes (i.e., top-down) in adults (Helo, Pannasch, Sirri, & Rämä, 2014; Walker, Bucker, Anderson, Schreij, & Theeuwes, 2017). Indeed, according to some authors, mature visual perception is not only achieved through bottom up (stimulus driven) processing of visual content gradually along the ventral route to access object identification until decision-making, but downward cognitive influences also modulate visual information processing very early on. These top-down influences would activate memory representations via the orbitofrontal cortex and the parahippocampal cortex (Bar, 2004; Chaumon, Kveraga, Barrett, & Bar, 2014; Kauffmann, Ramanoël, & Peyrin, 2014; Trapp & Bar, 2015). However, the orbitofrontal cortex would only be activated for known stimuli (Chaumon et al., 2014), a finding thus highlighting the importance of past experiences and learning (Bar, 2007). We therefore propose that from early development, semantic visual information would be processed in a mainly bottom-up manner, this processing being possible for two stimuli in parallel, as this study has shown. The processing of visual perceptual information allows access to its semantic representation, then as semantic associations are strengthened in long-term memory, the visual perception of a stimulus will allow not only the activation of its representation, and thus its identification, but also the activation of its associative network. The activation of this network will then make it possible, via top-down processes, to generate predictions, guide visual exploration and speed up the recognition of elements present in the visual scene when they are consistent with this semantic network (Trapp & Bar, 2015).

Processing of emotional information

In this study we also examined the ability to categorize the emotional content conveyed by both objects (animal, human or inanimate) and contexts (natural or manufactured). It should be remembered that we used images that are representative of everyday life as stimuli: these are complex images for which the evaluation of emotional content is not limited to the categorization of facial emotions. Despite the complexity of the images and the constraints of their presentation, our results show that children from the age of eight have the ability to identify the emotional content of an image presented outside their foveal vision and very quickly, with the notable exception of images of the context-level that conveyed a negative emotion.

Images conveying positive emotional content, whether at the object or at the context level, were better categorized (with greater accuracy and more quickly) than negative images. Our results are in line with studies of categorization of emotional content of faces in children that showed superior performance in categorizing positive faces compared to negative faces (Vesker, Bahn, Degé, Kauschke, & Schwarzer, 2018). The ability to categorize positive faces would be identical in five-year-olds and adults, while the ability to categorize other emotions would develop until early adolescence (Gao & Maurer, 2010). Our results show that, beyond the understanding of facial expressions, this advantage of positive content also applies to the categorization of everyday objects and contexts. The facilitation process of positive versus negative emotional content thus seems to be a robust effect, highlighted with different types of stimuli. It has been suggested that there is a processing bias towards positive stimuli in young children that changes in the course of development to a negative bias in young adults (Vesker et al., 2018). It was suggested that this superiority in categorizing positive stimuli over negative could come from a hierarchy of emotional information processing (Ihssen & Keil, 2013). Namely, Ihssen & Keil (2013) contend that the detection of positive emotional content would be synonymous with safety, so the decision to categorize emotion is made quickly (it requires few attention resources and is rapid) while the detection of negative emotional content, warning about potential threats and dangers for the individual's survival, would require more precise and detailed processings to specify the nature of this danger. The processing of negative information would accordingly be slower and more expensive in terms of attention resources than that of positive information.

The apparently poor categorizing skills for negative contexts (analyses revealed not-better-than-chance performance) could be explained in part by this need for additional processing: the detection of a hostile environment would require that this surrounding be characterized more precisely than a neutral or pleasant ones. Above the emotional valence, there is also a preferential local to global processing in childhood, which can, at least in part, explain why negative contexts are not recognized better than chance, whereas this is the case for negative objects.

If the voluntary judgment of negative images thus seems more difficult than that of positive images, interestingly, negative emotional content can be accessed quickly and involuntarily. Indeed, there was an interference effect of the processing of negative emotional content on the semantic categorization. This detrimental influence of the presence of negative emotional content not relevant to the task has recently been parallelised with an increase in the

latency of the N2 wave during a task in evoked potentials, a wave that suggests an increase in the cognitive control needed to resolve a conflict (Rahman & Wiebe, 2019). Negative emotional content therefore seems to capture and consume attentional resources, at the expense of the current task. Thus, negative emotional content is detected early and processed involuntarily, while its voluntary processing (e.g. in order to categorize its emotional content) seems to require several additional processing steps. Our study therefore suggests different characteristics and dynamics of emotional information processing depending on the task at hand. Indeed, many studies have highlighted an involuntary processing of emotional content, mainly negative, very early in development (Jessen, Altvater-Mackensen, & Grossmann, 2016; for a review see, Vaish et al., 2008) while the ability to categorize an image according to emotional valence or basic emotion (i.e. volitional processing) develops until the end of adolescence (Widen, 2013). The ability to categorize the emotional content of an image is thus dependent, for some researchers, on the maturation of certain high-level and essential cognitive skills such as executive functions (Martins, Osório, Veríssimo, & Martins, 2016), non-verbal capacities like analytical reasoning (von Salisch, Haenel, & Freund, 2013), as well as visuo-spatial processing and language proficiency (Vicari et al., 2000).

Finally, this study replicates the effect of emotional consistency according to which the presence of a context carrying the same emotional valence as an object facilitates the emotional categorization of that object (Mondloch, 2012; Mondloch et al., 2013; Nelson & Mondloch, 2017; Theurel et al., 2016). In addition, and to our knowledge for the first time, we highlight not only a facilitating effect of contextual emotional information on the categorization of the emotional content of the object but also its reciprocal, namely an influence of the emotional content of the object on the emotional categorization of the context. Moreover, this effect of emotional consistency is achieved with a glimpse as short as 100 milliseconds at two simultaneous images. It thus appears that a child is able to integrate emotional information that is conveyed simultaneously by different components of his or her visual field and may use it to refine and strengthen his or her understanding of the environment.

Conclusion

Taken as a whole, the results of this study suggest that from the age of eight years onwards, children can process an object and a context in parallel and access their semantic or emotional representations. Our study also suggests that emotional categorization is more challenging than semantic categorization for school-aged children, while involuntary

processing of negative emotional content seems to occur at the expense of the semantic categorization. These results call for further studies to help distinguishing the indirect, involuntary processing of emotions, a capacity present early in life and probably related to survival, from the voluntary processing necessary to assess the emotion carried by an image, a capacity that will only mature at the end of adolescence. Together with recent findings discussed above, our results provide evidence that information that can be used to make fundamental characterizations of natural scenes required for the perception of the visual environment is available early in the course of cognitive development, albeit not mature at the onset of adolescence.

Références

- Aguado, L., Martínez-García, N., Solís-Olce, A., Dieguez-Risco, T., & Hinojosa, J. A. (2018). Effects of affective and emotional congruency on facial expression processing under different task demands. *Acta Psychologica, 187*, 66–76.
- Augusti, E. M., Torheim, H. K., & Melinder, A. (2014). The effect of emotional facial expressions on children's working memory: Associations with age and behavior. *Child Neuropsychology, 20*(1), 86–105.
- Baayen, R. H., Vasishth, S., Kliegl, R., & Bates, D. (2017). The cave of shadows: Addressing the human factor with generalized additivemixed models. *Journal of Memory and Language, 94*, 206–234.
- Banich, M. T., & Brown, W. S. (2000). A Life-Span Perspective on Interaction Between the Cerebral Hemispheres. *Developmental Neuropsychology, 18*(1), 1–10. https://doi.org/10.1207/S15326942DN1801_1
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Rcpp, L. (2015). Package 'lme4'. *Convergence, 12*(1).
- Batty, M., & Taylor, M. J. (2002). Visual categorization during childhood: An ERP study. *Psychophysiology, 39*(4), S0048577202010764. <https://doi.org/10.1017/S0048577202010764>
- Bornstein, M. H., Mash, C., & Arterberry, M. E. (2011). Perception of object-context relations: eye-movement analyses in infants and adults. *Developmental Psychology*. <https://doi.org/10.1037/a0021059>
- Bourne, V. (2006). *The divided visual field paradigm: Methodological considerations*. (Psychology).
- Bova, S. M., Fazzi, E., Giovenzana, A., Montomoli, C., Signorini, S. G., Zoppello, M., & Lanzi, G. (2007). The development of visual object recognition in school-age children. *Developmental Neuropsychology, 31*(1), 79–102.
- Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition and Emotion, 24*(3), 377–400.
- Chicoine, A.-J., Proteau, L., & Lassonde, M. (2000). Absence of Interhemispheric Transfer of Unilateral Visuomotor Learning in Young Children and Individuals With Agenesis of the Corpus Callosum. *Developmental Neuropsychology, 18*(1), 73–94. https://doi.org/10.1207/S15326942DN1801_5
- Cohen-Gilbert, J. E., & Thomas, K. M. (2013). Inhibitory control during emotional distraction across adolescence and early adulthood. *Child Development, 84*(6), 1954–1966. <https://doi.org/10.1111/cdev.12085>
- Conte, S., Brenna, V., Ricciardelli, P., & Turati, C. (2018). The nature and emotional valence of a prime influences the processing of emotional faces in adults and children. *International Journal of Behavioral Development, 42*(6), 554–562. <https://doi.org/10.1177/0165025418761815>
- Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science, 15*(8), 559–564. <https://doi.org/10.1111/j.0956-7976.2004.00719.x>
- Fabre-Thorpe, M. (2011). The characteristics and limits of rapid visual categorization. *Frontiers in Psychology, 2*(October), 243. <https://doi.org/10.3389/fpsyg.2011.00243>
- Fagard, J., Hardy-Léger, I., Kervella, C., & Marks, A. (2001). Changes in Interhemispheric Transfer Rate and the Development of Bimanual Coordination during Childhood. *Journal of Experimental Child Psychology, 80*(1), 1–22. <https://doi.org/10.1006/jecp.2000.2623>
- Gao, X., & Maurer, D. (2010). A happy story: Developmental changes in children's sensitivity to facial expressions of varying intensities. *Journal of Experimental Child Psychology, 107*(2), 67–86. <https://doi.org/10.1016/j.jecp.2010.05.003>

- González, B., De Boeck, P., & Tuerlinckx, F. (2014). Linear mixed modelling for data from a double mixed factorial design with covariates: A case-study on semantic categorization response times. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *63*, 289–302.
- Happé, F. and Frith, U. (2014), Annual Research Review: Towards a developmental neuroscience of atypical social cognition. *J Child Psychol Psychiatr*, *55*: 553-577. doi:[10.1111/jcpp.12162](https://doi.org/10.1111/jcpp.12162)
- Harnad, S. (1987). *Categorical perception: The groundwork of cognition* (Cambridge).
- Harnad, S. (2017). *To cognize is to categorize: Cognition is categorization*. In *Handbook of categorization in cognitive science* (Elsevier, pp. 21–54).
- Harrison, T. B., & Stiles, J. (2009). Hierarchical forms processing in adults and children. *Journal of Experimental Child Psychology*, *103*(2), 222–240. <https://doi.org/10.1016/J.JECP.2008.09.004>
- Heck, A., Chroust, A., White, H., Jubran, R., & Bhatt, R. S. (2018). Development of body emotion perception in infancy: From discrimination to recognition. *Infant Behavior and Development*, *50*, 42–51. <https://doi.org/10.1016/J.INFBEH.2017.10.007>
- Helo, A., Azaiez, N., & Rämä, P. (2017). Word Processing in Scene Context: An Event-Related Potential Study in Young Children. *Developmental Neuropsychology*, *42*(7–8), 482–494. <https://doi.org/10.1080/87565641.2017.1396604>
- Helo, Andrea, Pannasch, S., Sirri, L., & Rämä, P. (2014). The maturation of eye movement behavior: Scene viewing characteristics in children and adults. *Vision Research*, *103*, 83–91. <https://doi.org/10.1016/J.VISRES.2014.08.006>
- Helo, Andrea, van Ommen, S., Pannasch, S., Danteny-Dordoigne, L., & Rämä, P. (2017). Influence of semantic consistency and perceptual features on visual attention during scene viewing in toddlers. *Infant Behavior and Development*, *49*, 248–266. <https://doi.org/10.1016/J.INFBEH.2017.09.008>
- Ihssen, N., & Keil, A. (2013). Accelerative and decelerative effects of hedonic valence and emotional arousal during visual scene processing. *The Quarterly Journal of Experimental Psychology*, *66*(7), 1276–1301.
- Jessen, S., Altvater-Mackensen, N., & Grossmann, T. (2016). Pupillary responses reveal infants' discrimination of facial emotions independent of conscious perception. *Cognition*, *150*, 163–169. <https://doi.org/10.1016/J.COGNITION.2016.02.010>
- Joubert, O. R., Fize, D., Rousselet, G. a, & Fabre-Thorpe, M. (2008). Early interference of context congruence on object processing in rapid visual categorization of natural scenes. *Journal of Vision*, *8*(13), 11.1-18. <https://doi.org/10.1167/8.13.11>
- Joubert, O. R., Rousselet, G., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, *47*(26), 3286–3297. <https://doi.org/10.1016/j.visres.2007.09.013>
- Keshavan, M. S., Diwadkar, V. A., DeBellis, M., Dick, E., Kotwal, R., Rosenberg, D. R., ... Pettegrew, J. W. (2002). Development of the corpus callosum in childhood, adolescence and early adulthood. *Life Sciences*, *70*(16), 1909–1922. [https://doi.org/10.1016/S0024-3205\(02\)01492-3](https://doi.org/10.1016/S0024-3205(02)01492-3)
- Kramer, H. J., Lagattuta, K. H., & Sayfan, L. (2015). Why is happy–sad more difficult? Focal emotional information impairs inhibitory control in children and adults. *Emotion*, *15*(1), 61–72. <https://doi.org/10.1037/emo0000023>
- Kret, M. E., & de Gelder, B. (2010). Social context influences recognition of bodily expressions. *Experimental Brain Research*, *203*(1), 169–180. <https://doi.org/10.1007/s00221-010-2220-8>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): affective ratings of pictures and instruction manual*. Tech Rep A-8. University of Florida, Gainesville.

- LoBue, V. (2009). More than just another face in the crowd: superior detection of threatening facial expressions in children and adults. *Developmental Science*, *12*(2), 305–313. <https://doi.org/10.1111/j.1467-7687.2008.00767.x>
- LoBue, V., & DeLoache, J. S. (2010). Superior detection of threat-relevant stimuli in infancy. *Developmental Science*, *13*(1), 221–228. <https://doi.org/10.1111/j.1467-7687.2009.00872.x>
- LoBue, V., Rakison, D. H., & DeLoache, J. S. (2010). Threat Perception Across the Life Span. *Current Directions in Psychological Science*, *19*(6), 375–379. <https://doi.org/10.1177/0963721410388801>
- Marchewka, A., Zurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*(2), 596–610. <https://doi.org/10.3758/s13428-013-0379-1>
- Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences*, *5*(10), 443–450. [https://doi.org/10.1016/S1364-6613\(00\)01752-6](https://doi.org/10.1016/S1364-6613(00)01752-6)
- McManis, M., Bradley, M., Berg, W., Cuthbert, B., & Lang, P. (2001). Emotional reactions in children: Verbal, physiological, and behavioral responses to affective pictures. *Psychophysiology*, *38*(2), 222–231.
- Mondloch, C. J. (2012). Sad or fearful? The influence of body posture on adults' and children's perception of facial displays of emotion. *Journal of Experimental Child Psychology*, *111*(2), 180–196. <https://doi.org/10.1016/J.JECP.2011.08.003>
- Mondloch, C. J., Horner, M., & Mian, J. (2013). Wide eyes and drooping arms: Adult-like congruency effects emerge early in the development of sensitivity to emotional faces and body postures. *Journal of Experimental Child Psychology*, *114*(2), 203–216. <https://doi.org/10.1016/j.jecp.2012.06.003>
- Munneke, J., Brentari, V., & Peelen, M. V. (2013). The influence of scene context on object recognition is independent of attentional focus. *Frontiers in Psychology*, *4*(AUG), 1–10. <https://doi.org/10.3389/fpsyg.2013.00552>
- Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Visual Cognition*, *25*(4–6), 563–574. <https://doi.org/10.1080/13506285.2017.1301615>
- O'Sullivan, J. T. (1996). Children's metamemory about the influence of conceptual relations on recall. *Journal of Experimental Child Psychology*, *62*, 1–29.
- Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466–478. <https://doi.org/10.1037/0096-3445.130.3.466>
- Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, *13*(4), 160–166. <https://doi.org/10.1016/j.tics.2009.01.006>
- Poirel, N., Mellet, E., Houdé, O., & Pineau, A. (2008). First came the trees, then the forest: Developmental changes during childhood in the processing of visual local-global patterns according to the meaningfulness of the stimuli. *Developmental Psychology*, *44*(1), 245–253. <https://doi.org/10.1037/0012-1649.44.1.245>
- Poulin-Dubois, D., & Pauen, S. (2017). *The Development of Object Categories: What, When, and How?* In Handbook of Categorization in Cognitive Science (Elsevier, pp. 653–671).
- Pourtois, G., Schettino, A., & Vuilleumier, P. (2012). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology*, *92*(3), 492–512.
- Quinn, P. C., & Johnson, M. H. (2000). Global-before-basic object categorization in connectionist networks and 2-month-old infants. *Infancy*, *1*(1), 31–46.

- Rahman, A., & Wiebe, S. (2019). Valence matters: An electrophysiological study on how emotions influence cognitive performance in children. *Developmental Psychobiology*, *61*, 290–303. <https://doi.org/10.1002/dev.21813>
- Rajhans, P., Jessen, S., Missana, M., & Grossmann, T. (2016). Putting the face in context: Body expressions impact facial emotion processing in human infants. *Developmental Cognitive Neuroscience*, *19*, 115–121. <https://doi.org/10.1016/j.dcn.2016.01.004>
- Rémy, F., Vayssière, N., Pins, D., Boucart, M., & Fabre-Thorpe, M. (2014). Incongruent object/context relationships in visual scenes: Where are they processed in the brain? *Brain and Cognition*, *84*(1), 34–43. <https://doi.org/10.1016/J.BANDC.2013.10.008>
- Righart, R., & de Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(3), 264–272. <https://doi.org/10.3758/CABN.8.3.264>
- Rosch, E. (1978). *Principles of categorization*. In *Cognition and Categorization* (In Rosch E). Potomac, MD, Erlbaum.
- Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get to the gist of real-world natural scenes. *Visual Cognition*, *12*(6), 852–877. <https://doi.org/10.1080/13506280444000553>
- Schel, M. A., & Crone, E. A. (2013). Development of response inhibition in the context of relevant versus irrelevant emotions. *Frontiers in Psychology*, *4*, 383. <https://doi.org/10.3389/fpsyg.2013.00383>
- Sheppes, G., Suri, G., & Gross, J. J. (2015). Emotion Regulation and Psychopathology. *Annual Review of Clinical Psychology*, *11*(1), 379–405. <https://doi.org/10.1146/annurev-clinpsy-032814-112739>
- Theurel, A., Witt, A., Malsert, J., Lejeune, F., Fiorentini, C., Barisnikov, K., & Gentaz, E. (2016). The integration of visual context information in facial emotion recognition in 5- to 15-year-olds. *Journal of Experimental Child Psychology*, *150*, 252–271. <https://doi.org/10.1016/J.JECP.2016.06.004>
- Thomas, R. L., Nardini, M., & Mareschal, D. (2017). The impact of semantically congruent and incongruent visual information on auditory object recognition across development. *Journal of Experimental Child Psychology*, *162*, 72–88. <https://doi.org/10.1016/J.JECP.2017.04.020>
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.
- Trapp, S., & Bar, M. (2015). Prediction, context, and competition in visual recognition. *Annals of the New York Academy of Sciences*, *1339*(1), 190–198.
- Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. <https://doi.org/10.1037/0033-2909.134.3.383>
- Vanmarcke, S., & Wagemans, J. (2015). The influence of age and gender on ultra-rapid categorization. *Visual Cognition*, *23*(7), 894–916. <https://doi.org/10.1080/13506285.2015.1091801>
- Vesker, M., Bahn, D., Degé, F., Kauschke, C., & Schwarzer, G. (2018). Developmental changes in the categorical processing of positive and negative facial expressions. *PLOS ONE*, *13*(8), e0201521. <https://doi.org/10.1371/journal.pone.0201521>
- Vicari, S., Reilly, J. S., Pasqualetti, P., Vizzotto, A., & Caltagirone, C. (2000). Recognition of facial expressions of emotions in school-age children: the intersection of perceptual and semantic categories. *Acta Paediatrica*, *89*(7), 836–845. <https://doi.org/10.1111/j.1651-2227.2000.tb00392.x>
- Vuilleumier, P. (2015). Affective and motivational control of vision. *Current Opinion in Neurology*, *28*(1), 29–35.

- Walker, F., Bucker, B., Anderson, N., Schreij, D., & Theeuwes, J. (2017). Looking at paintings in the Vincent Van Gogh Museum: Eye movement patterns of children and adults. *PLoS ONE*, *12*(6), e0178912. <https://doi.org/10.1371/journal.pone.0178912>
- Wickham, H. (2009). *Elegant graphics for data analysis* (ggplot2).
- Widen, S. C. (2013). Children's Interpretation of Facial Expressions: The Long Path from Valence-Based to Specific Discrete Categories. *Emotion Review*, *5*(1), 72–77. <https://doi.org/10.1177/1754073912451492>
- Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development*, *23*(2), 291–312. <https://doi.org/10.1016/J.COGDEV.2008.01.002>
- Xiao, J., Hays, J., Ehinger, K., Oliva, A., & Torralba, A. (2010). SUN Database: Large-scale Scene Recognition from Abbey to Zoo. In *IEEE Conference on Computer Vision and Pattern Recognition*.

VII.5 Discussion intermédiaire

Cette étude a analysé les capacités d'enfants de 8 à 12 ans à catégoriser sémantiquement et émotionnellement des scènes visuelles. Un objet et un contexte ont été présentés rapidement et séparément à des enfants qui devaient réaliser soit une tâche de catégorisation sémantique (avec manipulation de la cohérence sémantique entre l'objet et le contexte) soit une tâche de catégorisation émotionnelle (avec manipulation de la cohérence émotionnelle). L'ensemble des résultats suggère que dès 8 ans l'enfant peut traiter en parallèle un objet et un contexte et accéder à leurs représentations sémantique ou émotionnelle.

Les résultats mettent en évidence chez les enfants des capacités de catégorisation sémantique, en termes d'exactitude de réponse, comparables à celles des adultes (mais avec des temps de réponse plus longs). Ces résultats rejoignent ceux obtenus lors d'une présentation visuelle centrale (Batty & Taylor, 2002) et montrent, de façon originale, que le processus de catégorisation sémantique est efficace même lorsque l'enfant doit, au moins initialement, diviser son attention sur deux images présentées simultanément dans des hémichamps latéraux différents. L'effet de cohérence sémantique, mis en évidence dans les études auprès d'adultes, n'est pas observé ici auprès d'enfants de 8 à 12 ans. Il est possible que la mise en relation des contenus objet-contexte se développe au fur et à mesure des expériences et des apprentissages, comme déjà suggéré dans d'autres travaux (Thomas et al., 2017).

Notre étude suggère également que la catégorisation émotionnelle est plus difficile que la catégorisation sémantique pour des enfants d'âge scolaire. En outre, un traitement non volontaire du contenu émotionnel négatif au détriment de la tâche de catégorisation sémantique est mis en évidence. Ces résultats distinguent ainsi le traitement indirect, non volontaire du contenu émotionnel, capacité présente tôt dans le développement, du traitement volontaire nécessaire pour évaluer l'émotion portée par une image, capacité qui ne sera mature qu'à la fin de l'adolescence, comme le suggère notre seconde revue de littérature (article II, chapitre II).

Enfin cette étude a mis en évidence la présence d'un effet facilitateur de la cohérence émotionnelle : lorsque l'objet et le contexte d'une scène visuelle sont porteurs du même contenu émotionnel, la catégorisation émotionnelle des deux niveaux est améliorée par rapport à une situation d'incohérence émotionnelle. La mise en commun des informations émotionnelles des niveaux local et global semble ainsi s'opérer plus tôt dans le développement que pour les informations sémantiques et serait donc probablement moins dépendante des apprentissages.

Chapitre VIII: Application au Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H)

Ce chapitre présente la quatrième et dernière étude empirique de ce travail de thèse. Cette étude analyse les capacités de traitements sémantiques et émotionnels lors de la perception rapide de scènes visuelles auprès d'enfants de sept à 12 ans qui présentent un TDA/H. La totalité des inclusions n'étant pas finalisée lors de la rédaction de ce manuscrit, nous présentons ici les résultats préliminaires de l'étude, nous n'incluons donc pas d'article et ce chapitre se structurera comme un chapitre de thèse « classique ». Il est constitué d'un rappel des apports théoriques, de la présentation de l'étude, des analyses préliminaires et d'une discussion intermédiaire.

VIII.1 Rappel des apports théoriques principaux

Les travaux réalisés dans les trois études expérimentales des chapitres V, VI et VII (articles IV, V et VI respectivement) ont permis d'améliorer la compréhension des processus à l'œuvre lors de la perception visuelle rapide de scènes complexes dans le fonctionnement cognitif normal et de spécifier certains aspects développementaux. Au-delà de ces enjeux fondamentaux, l'intérêt de ces travaux est également à visée clinique.

Dans le cadre de cette dernière étude de la thèse, nous souhaitons analyser les particularités de traitement des informations visuelles sémantiques et émotionnelles dans le TDA/H. En effet, un déficit du traitement des émotions a souvent été décrit dans cette population. Les troubles comportementaux et les conflits avec les pairs, fréquemment notés dans le TDA/H, seraient, selon certains auteurs, en partie dus à de faibles capacités de reconnaissance et de gestion des émotions. Il a ainsi été mis en évidence que la compréhension des émotions est perturbée, aussi bien s'agissant de les identifier sur des visages expressifs que de les inférer à partir de scènes visuelles (voir article III, chapitre III). Les recherches ont également mis en évidence des profils de marqueurs oculométriques particuliers chez les enfants TDA/H lors de l'exploration de visages émotionnels. Par exemple, le temps d'exploration visuelle de la bouche est plus long que chez les enfants neurotypiques ou ceux présentant un trouble du spectre autistique (Muszkat et al., 2015). On note également un temps d'exploration des images émotionnelles négatives anormalement long et leur catégorisation selon le contenu émotionnel est moins exacte que celles des images neutres ou au contenu émotionnel positif (Pelc, Kornreich, Foisy, & Dan, 2006). Cependant, les résultats des études sur le traitement des émotions dans le TDA/H sont parfois divergents et, comme nous l'avons

souligné dans notre troisième revue de littérature (article III, chapitre III), la majorité des études utilisent des visages émotionnels comme stimuli et très peu ont étudié les capacités de traitement sémantique des informations visuelles auprès de cette population. Or, pour valider l'évaluation de l'influence des émotions sur l'exécution des tâches, et mieux comprendre d'où proviennent les difficultés de traitement des émotions dans le TDA/H, il semble nécessaire de pouvoir dissocier ce qui relève d'anomalies des traitements émotionnels et des traitements sémantiques. Pour cela il convient de s'assurer que les enfants et les adolescents qui présentent un TDA/H ont une compréhension préservée de la signification générale de la scène visuelle.

VIII.2 Présentation de l'étude

Objectifs

L'objectif principal de cette étude est d'analyser les capacités de traitement des informations émotionnelles véhiculées par des scènes visuelles chez les enfants TDA/H. Les résultats obtenus nous permettront d'appréhender si les difficultés des enfants TDA/H se situent à une étape de traitement plutôt non volontaire, automatique des processus de traitement des informations émotionnelles ou plutôt à une étape de traitement volontaire, contrôlée. Une influence du contenu émotionnel des images sur leur catégorisation sémantique indiquerait un traitement automatique, non volontaire du contenu émotionnel. La comparaison de l'intensité de l'interférence émotionnelle entre enfants TDA/H et au développement typique nous permettra de définir si le traitement automatique du contenu émotionnel diffère dans le TDAH, et si cela est le cas, dans quelle mesure (traitement renforcé ou diminué). La tâche de catégorisation émotionnelle nous permettra d'évaluer les capacités de traitement volontaire du contenu émotionnel des enfants TDA/H lors d'une présentation visuelle rapide. Nous pourrons ainsi analyser les traitements précoces de l'information émotionnelle avant la mise en place d'explorations volontaires et guidées.

Cette étude permettra également de mieux comprendre quelles sont les informations sélectionnées et traitées en priorité par les enfants qui présentent un TDA/H. Pour ce faire nous testerons si les performances de catégorisation sont modulées par le contenu émotionnel, le niveau de l'image à analyser (contexte ou objet), ainsi que par l'association avec un autre constituant de la scène visuelle (effets de cohérence sémantique ou émotionnelle).

Comme dans l'étude présentée précédemment (article VI, chapitre VII) et qui a été réalisée auprès d'enfants au développement typique, une paire d'images, un contexte et un objet, est présentée rapidement à chaque essai de la tâche, une image dans le champ visuel

droit et l'autre dans le champ visuel gauche. Chaque image peut avoir un contenu émotionnel positif, négatif ou neutre. Les données obtenues par le groupe d'enfants qui présentent un TDA/H seront comparées aux données issues de l'étude antérieure réalisée avec la même méthodologie auprès d'enfants au développement typique (cf. article VI, chapitre VII).

Calcul de l'échantillon

Nous présentons ici les résultats préliminaires de l'étude puisque nous avons à ce jour inclus 40 participants TDA/H sur les 56 prévus. En effet le calcul de l'échantillon réalisé, sur le logiciel G*Power 3.1², nous a permis de calculer la valeur critique de $F=3.09$ et ainsi la taille totale de l'échantillon nécessaire : celle-ci est de 52 participants (50% TDA/H, 50% neurotypiques) par tâche (tâche de catégorisation sémantique et tâche de jugement émotionnel) que nous élevons à 56 pour le contrebalancement des facteurs expérimentaux, soit 112 au total. Nous comparerons donc les données des 56 enfants TDA/H qui participeront à cette étude aux données de 56 enfants au développement typique (i.e., groupe contrôle) qui ont été sélectionnés parmi l'ensemble des participants neurotypiques (article VI) de façon à obtenir le meilleur appariement possible selon l'âge, le sexe et le niveau socioéducatif.

Le calcul de l'échantillon a pris en compte :

- le type d'analyse statistique : analyses de la variance (ANOVA) à mesures répétées avec le contenu émotionnel en intrasujet et le groupe de participants en intersujets ;
- le nombre de groupes (deux : TDA/H et contrôles) ;
- le nombre des variables intrasujet, qui est de trois (pour les trois modalités de la variable contenu émotionnel : neutre vs. positif vs. négatif) ;
- la spécification du choix de l'option de la taille d'effet tel que recommandé classiquement dans les études en sciences humaines (Cohen, 1988) ;
- la taille de l'effet (*Effect size*) : estimation des effets attendus $\geq .15$. Cette estimation a été faite en prenant en compte la littérature dans le domaine³ ainsi que l'étude déjà menée chez des participants adultes sains par notre équipe (article V). Ces références amènent à prédire des effets forts ou moyens pour les effets principaux et pour les interactions ; en considérant le caractère innovant, chez les enfants, de notre étude et la variabilité individuelle entre les participants, nous avons retenu comme mesure de

² <http://www.gpower.hhu.de/en.html>

³ si non spécifiée, la taille des effets dans chaque étude considérée a été calculée en utilisant <https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD5.php>

prudence l'adoption du seuil indiquant la valeur minimale d'un effet de taille moyenne ($f^2 = .15$);

- une puissance statistique (1-Beta) de .80 (valeur standard en psychologie);
- un seuil de probabilité alpha de l'erreur de type I (*Alpha error probability*) de 0.05 (valeur standard en psychologie).

Méthodologie expérimentale

Cette étude reprend la même méthodologie expérimentale que celle décrite dans l'article VI (chapitre VII) auprès d'enfants au développement typique (mêmes tâches expérimentales et mêmes questionnaires). Ainsi, la tâche de jugement émotionnel des scènes, qui oriente l'attention des enfants sur les contenus émotionnels, permettra d'analyser le traitement volontaire des émotions et la tâche de catégorisation sémantique des scènes permettra d'analyser le traitement non volontaire des émotions. Nous testerons sur cette base si et comment ces traitements volontaires et non volontaires des informations émotionnelles sont modulés par le niveau de la scène, c'est-à-dire son niveau plutôt global (contexte) ou plutôt local (objet).

Chaque enfant se voit attribuer aléatoirement une des tâches : tâche de catégorisation sémantique du contexte, tâche de catégorisation sémantique de l'objet, tâche de catégorisation émotionnelle du contexte ou enfin tâche de catégorisation émotionnelle de l'objet. Les stimuli et la méthodologie expérimentale (temps de présentation, position des images sur l'écran) sont exactement les mêmes que ceux présentés dans le chapitre VII (article VI) auprès d'enfants neurotypiques.

Analyse des données

La comparaison des performances des enfants TDA/H aux performances des enfants contrôles se base sur deux mesures : l'exactitude de réponse et les temps de réponse des réponses correctes. L'analyse de ces deux mesures complémentaires est importante : l'exactitude nous permettra de quantifier et ainsi de différencier les capacités de traitement des enfants TDA/H de celles des enfants contrôles (efficacité), et les temps de réponse d'aborder les dynamiques des processus cognitifs et des stratégies de réalisation de la tâche (efficacité). La prise en compte de ces deux mesures nous permettra également de mettre en évidence un effet éventuel lié à l'impulsivité des enfants TDA/H, qui impliquerait une exactitude faible associée à des temps de réponse anormalement courts.

Les analyses statistiques des données sont réalisées principalement par des modèles mixtes linéaires généralisés (GLMMs) sur l'exactitude et par des modèles mixtes linéaires (LMMs) sur les temps de réponse des réponses correctes transformés en logarithmes. Les modèles ont en facteurs aléatoires les participants et les items.

Notre objectif principal est d'analyser l'influence du contenu émotionnel des scènes sur les performances en fonction du groupe de participants (groupe « TDA/H » vs. Groupe « contrôle » constitué d'enfants au développement typique). Les analyses principales qui portent uniquement sur les images cohérentes sémantiquement (expérience 1) ou émotionnellement (expérience 2), auront ainsi en prédicteurs le Groupe de participants (TDA/H vs. Contrôle), le Contenu Emotionnel de l'image-cible (Neutre vs. Positive et Neutre vs. Négative) et leurs interactions.

Pour les analyses complémentaires liées à nos objectifs secondaires d'étude de l'influence du niveau de l'image et des associations sémantiques ou émotionnelles au sein des scènes sur les performances de catégorisation, les modèles auront comme prédicteurs le Niveau de l'image-cible (Contexte vs. Objet), la Cohérence (Cohérent vs. Incohérent) Sémantique (tâche de catégorisation sémantique) ou Emotionnelle (tâche de catégorisation émotionnelle) de la paire d'images et leurs interactions.

Les essais avec un temps de réponse aberrant (supérieur à 5s ou inférieur à 300ms) ont été exclus (2,9% des essais pour la tâche sémantique et 4,3% pour la tâche émotionnelle). Tous les modèles ont d'abord été exécutés avec une structure aléatoire complète. Lorsque le modèle maximal ne convergait pas, nous avons simplifié le modèle en supprimant les corrélations entre les pentes aléatoires et les intercepts, puis, si nécessaire, en réduisant la complexité du modèle. Les analyses ont été réalisées en utilisant la fonction `lmer` du progiciel `lme4` (Bates et al., 2015) dans l'environnement de programmation R (The R Fondation for Statistical Computing, version 3.3.3, 2016). Pour chaque modèle, nous présentons les coefficients des variables prédicteurs (β valeurs), les valeurs *SE*, les valeurs *z* (GLMM) ou les valeurs *t* (LMM), et les valeurs *p* associées. Les valeurs *p* ont été générées par la bibliothèque `lmerTest` (Kuznetsova, Brockhoff, & Christensen, 2017). Pour les effets montrant une tendance à un effet significatif ($.05 < p < 0.1$), nous avons rapporté des valeurs *z*- (pour les GLMM) ou *t*- (pour les LMM) et *p*- mais pas les valeurs β - et *SE*. Les graphiques ont été créés à l'aide du package `ggplot2` (Wickham, 2009). Des modèles de suivi ont été réalisés dans le cas d'une interaction significative pour examiner les différences entre les conditions impliquées.

Modalités de recrutement et de passation

Cette étude est réalisée en partenariat avec le Centre de Référence des Troubles des Apprentissages (CERTA) des hôpitaux pédiatriques de Nice CHU-Lenval. Elle bénéficie d'un soutien logistique et financier obtenu en réponse à un appel à projet interne du CHU de Nice (référence : RECH/AP/DD/LM/18477) et a reçu un avis éthique favorable du Comité de Protection des Personnes (CPP) Sud Méditerranée I (référence 19 19, Annexe 8). Le numéro d'identifiant d'enregistrement clinique unique de l'étude (sur *ClinicalTrials.gov*) est le 18-HPNCL-07 (cf. Annexe 9).

Les patients TDA/H sont recrutés au CERTA. Il s'agit de patients diagnostiqués, et dont le suivi est réalisé chaque année dans le centre, l'endroit leur est donc familier. Les médecins du CERTA proposent au patient de participer à cette recherche lors d'une consultation classique de suivi. Les familles intéressées par l'étude reçoivent à cette occasion la description détaillée de l'étude et les formulaires de consentement « parents » qui doivent être signés par au moins un des deux représentants légaux de l'enfant (cf. Annexe 10). Chaque famille reçoit également un formulaire d'inclusion (cf. Annexe 11) et un formulaire « points forts-points faibles » concernant l'enfant (cf. Annexe 6). Avant de débiter la passation, l'étude est expliquée à l'enfant qui doit donner son accord et signer à son tour un formulaire de consentement (cf. Annexe 12).

La participation à l'étude suppose que les patients TDA/H se rendent au CERTA, ils reçoivent en dédommagement des frais de transport un forfait de 25 euros.

Critères d'inclusion et d'exclusion

Les critères d'inclusion de l'étude sont :

- Enfants (sexe masculin et féminin) âgés de 7 à 12 ans ;
- Acuité visuelle normale ou corrigée à la normale ;
- Diagnostic de TDA/H posé par un pédiatre, neuro-pédiatre, pédo-psychiatre ou médecin, selon les critères du DSM-5 ;
- Absence de traitement médicamenteux (méthylphénidate) durant les 24 heures précédant la passation de l'expérience ;
- Affiliation à la sécurité sociale ;
- Formulaire de consentement éclairé signé par un représentant légal de l'enfant, et formulaire de consentement éclairé signé par l'enfant.

Les critères de non-inclusion pour l'étude sont :

- Acuité visuelle anormale qui même corrigée limite la vision ;
- Antécédents neurologiques ou d'épilepsie ; enfant présentant un trouble psychotique, un retard mental ou un trouble spécifique du langage oral (confirmé par le formulaire d'inclusion renseigné par les parents, cf. Annexe 2).

VIII.3 Analyses préliminaires

Nous présentons ici les résultats issus de l'analyse des données de 40 enfants présentant un TDA/H (10 par expérience, âge moyen 10.2 ans, déviation standard 1.27, 10 filles) et de 64 enfants contrôles (16 par expérience, âge moyen 9.34 ans, déviation standard 0.77, 35 filles). L'ensemble des participants contrôles est utilisé dans ces analyses préliminaires pour augmenter la puissance statistique et pour avoir une comparaison avec un échantillon le plus représentatif que possible de la population générale. Une fois les inclusions terminées nous ferons un appariement en fonction de l'âge, du sexe et du niveau socioéducatif. Ces résultats seront donc à confirmer : l'étude se clôturera lorsque nous aurons atteints 56 enfants TDA/H soit quatorze par tâche expérimentale, ce qui correspond à un participant supplémentaire par version expérimentale dans chaque tâche expérimentale.

VIII.3.1 Tâche de catégorisation sémantique

Résultats préliminaires

Cette première tâche nous permet d'étudier la capacité de catégorisation sémantique d'une image présentée parmi une paire rapidement et hors de la vision fovéale. Nous étudions ainsi le traitement non volontaire du contenu émotionnel par l'analyse de l'influence de la présence de contenu émotionnel sur les performances de catégorisation sémantique.

Au niveau de l'exactitude, l'analyse montre un effet significatif du Contenu Emotionnel de l'image cible pour la comparaison des images neutres aux images négatives, $\beta = -.32$, $SE = .11$, $z = -2.87$, $p = .004$, les images neutres présentent une exactitude de catégorisation supérieure ($M = 0.85$, $SE = 0.004$) aux images à contenu négatif ($M = 0.82$, $SE = 0.004$). Nous obtenons également un effet significatif du Groupe, $\beta = -.44$, $SE = .22$, $z = -1.97$, $p = .049$, les enfants du groupe contrôle présentent des performances supérieures ($M = 0.87$, $SE = 0.004$) à celles du groupe TDA/H ($M = 0.82$, $SE = 0.006$). Nous ne mettons pas en évidence d'autres effets principaux ou d'interaction (tous $z \leq 1$, tous $p \geq .344$).

Au niveau des temps de réponse, nous obtenons également un effet principal significatif du Groupe, $\beta = -.07$, $SE = .03$, $t = -2.78$, $p = .008$, les enfants du groupe contrôle présentent des temps de réponse plus longs ($M = 1615$ ms $SE = 10.16$) que ceux du groupe TDA/H ($M = 1426$ ms, $SE = 12.94$). Les résultats montrent également une tendance à ce qu'il y ait un effet du Contenu Emotionnel pour la comparaison des images neutres aux images négatives, $t = 1.87$, $p = .066$. Nous n'obtenons aucun autre effet principal ni d'interaction (tous $t \leq 1.36$, tous $p \geq .178$).

Pour nos objectifs secondaires, nous étudions si les performances de catégorisation sémantique sont modulées par le niveau de l'image (contexte ou objet) cible et l'association sémantique avec l'autre image de la paire (présentation en situation de cohérence ou d'incohérence sémantique).

Au niveau de l'exactitude, nous mettons en évidence un effet significatif du Niveau de l'image à juger, $\beta = 1.18$, $SE = .14$, $z = 8.32$, $p < .001$, avec un avantage pour la catégorisation des images de type objet ($M = 0.91$, $SE = 0.005$) par rapport aux images de type contexte ($M = 0.76$, $SE = 0.007$). Nous montrons également un effet significatif du Groupe, $\beta = -.47$, $SE = .14$, $z = -3.25$, $p = .001$, le groupe contrôle présentant des performances supérieures ($M = 0.85$, $SE = 0.005$) à celles du groupe TDA/H ($M = 0.80$, $SE = 0.007$). Nous n'obtenons aucun autre effet principal ni d'interaction (tous $z \leq 1.23$, tous $p \geq .22$).

Au niveau des temps de réponse, nous obtenons également un effet principal significatif du Groupe, $\beta = -.07$, $SE = .03$, $t = -2.76$, $p = .008$, les enfants du groupe contrôle ($M = 1621$ ms, $SE = 12.93$) présentent des temps de réponse allongés par rapport à ceux du groupe TDA/H ($M = 1436$ ms, $SE = 16.17$). Cependant une interaction avec le niveau de l'image, $\beta = .12$, $SE = .05$, $t = 2.25$, $p = .029$, met en évidence que cette différence n'est significative que pour les images de type contexte ($M = 1734$ ms, $SE = 20.31$ pour le groupe contrôle ; $M = 1366$ ms, $SE = 25.07$ pour le groupe TDA/H), $\beta = -.06$, $SE = .02$, $t = -3.53$, $p < .001$. Les temps de réponses pour les images de type objet ne diffèrent pas entre les deux groupes ($t < 1$, $p = .728$). Nous n'obtenons pas d'autres effets principaux ni d'interaction (tous $t \leq 1$, tous $p \geq .530$).

Discussion préliminaire

Les enfants TDA/H présentent des performances de catégorisation sémantique, en termes d'exactitude, plus faibles que celles du groupe contrôle et associées à des temps de réponse plus courts. Ceci peut indiquer la présence d'impulsivité (plus de réponses inexactes données rapidement). Néanmoins, les performances des enfants TDA/H restent de haut niveau

(81% contre 86% pour le groupe contrôle) ce qui suggère l'efficacité des capacités de catégorisation sémantique même pour une présentation visuelle rapide et latéralisée et donc avec des entrées perceptives limitées. Ce résultat est supporté par d'autres études, en oculométrie, qui ont utilisé des scènes visuelles comme matériel expérimental et ont mis en avant que la première saccade d'enfants TDA/H se dirigeait vers les informations pertinentes de la scène, ce qui va dans le sens d'une analyse initiale du contenu global de la scène efficiente (Karatekin & Asarnow, 1999; Serrano, Owens, & Hallowell, 2015; Türkan, Amado, Ercan, & Perçinel, 2016). En outre, dans notre étude les temps de réponse pour les images de type objet sont semblables entre les deux groupes. Cependant, les enfants au développement typique présentent des temps de réponse allongés, comparativement au groupe TDA/H, pour les images de type contexte. Dans le groupe TDA/H, de façon similaire au groupe contrôle, l'étude met en évidence une exactitude de catégorisation des objets supérieure à celle des contextes, mais sans différence significative entre les temps de réponse. Les enfants du groupe TDA/H ne semblent pas adapter leur temps de réponse aux contraintes de la tâche, cela pouvant être lié à l'aversion au délai (Sonuga-Barke, 2002). Néanmoins, il est important de prendre en considération la variabilité, fréquemment reportée dans la littérature, des temps de réponse dans le TDA/H (pour une revue voir Kofler et al., 2013) qui suggère de rester précautionneux dans l'analyse des temps de réponse dans cette population. De plus, il a été suggéré que dans des conditions qui incitent à accélérer (comme dans notre étude avec des présentations très rapides, de 100 ms) les personnes TDA/H présentent une diminution des temps de réponse plus marquée que les contrôles (Andreou et al., 2007).

En ce qui concerne notre objectif principal, qui est l'analyse des capacités de traitement non volontaire, donc probablement automatique, du contenu émotionnel chez les enfants TDA/H, nous mettons en évidence un effet du contenu émotionnel sur les performances à la tâche de catégorisation sémantique sans interaction avec le groupe expérimental, donc de manière semblable pour les enfants au développement typique et ceux avec TDA/H. La diminution des performances (baisse de l'exactitude et augmentation des temps de réponse) pour la catégorisation des images négatives, comparativement aux images neutres et aux images positives, laisse inférer que la présence d'un contenu émotionnel négatif interfère avec la tâche en cours. Les enfants qui présentent un TDA/H semblent ainsi traiter non volontairement le contenu émotionnel d'une image qui leur est présenté rapidement et hors de la vision fovéale, de façon semblable aux enfants neurotypiques. Ceci suggère un traitement automatique du contenu émotionnel, préservé chez les enfants avec TDA/H. Nos résultats rejoignent ainsi les

théories qui postulent que les particularités de traitements des émotions dans le TDA/H ne proviendraient pas d'un déficit émotionnel spécifique mais découleraient des troubles primaires du TDA/H, c'est-à-dire déficit attentionnel et d'inhibition (Barkley, 1997). Par exemple Shin et collaborateurs (2008) ont mis en évidence que les difficultés de reconnaissance des émotions dans le TDA/H étaient corrélées au déficit attentionnel (évalué par un test informatisé de performance continue). Villemonteix *et al.* (2017) ont analysé l'effet d'interférence d'un contenu émotionnel non pertinent sur la réalisation d'une tâche et ont mis en évidence un défaut de contrôle inhibiteur général dans le TDA/H et non spécifiquement lié à un contenu émotionnel. Ainsi il semble que les capacités de traitement précoce des émotions ne diffèrent pas entre enfants avec TDA/H et au développement typique mais le défaut de contrôle inhibiteur, qui altère les capacités d'attention soutenue, ne permet pas aux enfants avec TDA/H d'affiner leur prise d'information.

Nos objectifs secondaires étaient d'analyser si le niveau de l'image (contexte ou objet) et l'association sémantique avec l'autre image de la paire pouvaient influencer les capacités de catégorisation sémantique d'images présentées rapidement. Nous avons retrouvé un effet du niveau de l'image, comme discuté plus haut, mais nous n'avons pas observé d'effet de la cohérence sémantique, c'est-à-dire que les enfants des deux groupes semblent avoir catégorisé l'image cible de façon isolée sans que le contenu sémantique de l'image associée n'influe sur cette catégorisation. Cette absence d'effet de la cohérence sémantique peut s'expliquer par la difficulté à cet âge de traiter plusieurs informations simultanément, surtout dans une tâche complexe et qui limite le contenu visuel disponible. Une autre explication serait que l'effet de cohérence sémantique se développe avec la répétition des co-occurrences rencontrées dans les scènes visuelles et ainsi avec les apprentissages et les expériences de vie (Thomas et al., 2017). Ces associations ne seraient pas encore assez stabilisées chez des enfants d'âge scolaire pour être activées, notamment lors d'une présentation si rapide des images. Nous reviendrons plus en détail sur ces hypothèses dans la discussion générale de la thèse.

VIII.3.2 Tâche de catégorisation émotionnelle

Résultats préliminaires

Dans cette seconde tâche nous avons étudié la capacité de catégorisation du contenu émotionnel d'une image et ainsi le traitement explicite du contenu émotionnel.

Au niveau de l'exactitude, nous obtenons un effet significatif du Contenu Emotionnel de l'image à juger pour la comparaison des images neutres aux images positives, $\beta = .54$, $SE = .24$,

$z=2.28, p=.022$, ainsi que pour la comparaison des images neutres aux images négatives, $\beta=-.83, SE=.22, z=-3.79, p<0.001$, avec un avantage pour les images à contenu émotionnel positif ($M=0.61, SE=0.008$) par comparaison aux images neutres ($M=0.52, SE=0.006$) et aux images à contenu négatif ($M=0.36, SE=0.008$). Nous n'obtenons pas d'effet significatif du Groupe, ni d'effets d'interaction (tous $z \leq 1.57$, tous $p \geq 0.117$).

Au niveau des temps de réponse, les analyses montrent également un effet principal significatif du Contenu Emotionnel de l'image à juger, pour la comparaison des images neutres aux images positives, $\beta=-.06, SE=.01, t=-4.29, p<0.001$, et pour la comparaison des images neutres aux images négatives, $\beta=.036, SE=.01, t=2.70, p=0.01$. Les images à contenu émotionnel positif sont catégorisées plus rapidement ($M=1432$ ms, $SE=20.51$) que les images neutres ($M=1699$ ms, $SE=14.11$) qui sont catégorisées plus rapidement que les images à contenu négatif ($M=1768$ ms, $SE=30.11$). Nous obtenons une tendance à ce qu'il y ait un effet du Groupe, $t=-1.95, p=.057$. Nous n'obtenons pas d'effets d'interaction (tous $t \leq 1.39$, tous $p \geq 0.170$).

Pour nos objectifs secondaires, nous avons étudié si les performances de catégorisation émotionnelle étaient modulées par le niveau de l'image (contexte ou objet) à juger et par le contenu émotionnel de l'autre image de la paire (présentation en situation de cohérence ou d'incohérence émotionnelle). Comme pour la tâche sémantique, les analyses relatives aux objectifs secondaires n'ont été réalisées que sur les images émotionnelles (positives et négatives).

Au niveau de l'exactitude, un effet significatif principal du Niveau de l'image à juger est mis en évidence, $\beta=.58, SE=.17, z=3.42, p<0.001$, les images de contextes ($M=0.42, SE=0.008$) sont moins bien catégorisées que les images d'objets ($M=0.55, SE=0.008$). L'analyse des données montre également un effet de la Cohérence Emotionnelle, $\beta=-.16, SE=.07, z=-2.10, p=.036$, les images présentées en condition de cohérence émotionnelle ($M=0.50, SE=0.008$) sont catégorisées avec une exactitude supérieure aux images présentées en situation d'incohérence émotionnelle ($M=0.47, SE=0.008$). Nous obtenons également une tendance pour un effet d'interaction entre ces deux facteurs, $z=1.83, p=.067$. Aucun autre effet principal ni d'interaction (tous $z < 1$, tous $p \geq 0.468$) n'est observé.

Au niveau des temps de réponse, nous observons une tendance à ce qu'il y ait un effet du Groupe, $\beta=-.05, SE=.03, z=-1.85, p=.071$. Aucun effet significatif principal ni d'interaction n'est obtenu, (tous $t \leq 1.57$, tous $p \geq 0.126$).

Discussion préliminaire

Les résultats suggèrent que les enfants TDA/H ont des capacités de traitement volontaire du contenu émotionnel préservées lors d'une présentation visuelle rapide : en effet, les performances à la tâche de catégorisation du contenu émotionnel sont équivalentes entre les deux groupes.

Ces résultats suggèrent, ainsi, que la prise d'information initiale, « en un coup d'œil », lors de l'apparition d'une scène visuelle est efficace chez les enfants TDA/H et leur permet d'accéder à la même quantité d'information émotionnelle et à la même évaluation que les enfants au développement typique. Ces résultats sont soutenus par certains travaux qui mettent en évidence, lors de l'exploration de scènes visuelles, des capacités de traitement initial des informations visuelles fonctionnelles chez les enfants TDA/H et un déficit plus tardif lié à l'exploration et la prise d'information plus détaillée (Serrano et al., 2015; Türkan et al., 2016). Ces travaux ont ainsi décrit que lors de l'apparition d'une scène visuelle l'enfant TDA/H fixe les mêmes zones que l'enfant au développement typique mais que l'exploration visuelle diffère ensuite : les enfants TDA/H réalisent moins de fixations et/ou le temps total d'exploration de l'image est plus court. Notre étude ouvre ainsi des questions intéressantes que la recherche future pourra développer. Ainsi, en se référant au modèle séquentiel de la perception visuelle, il semblerait donc que le traitement automatique du contenu émotionnel, guidé par le stimulus (et donc l'attention exogène), soit fonctionnel dans le TDA/H. Par contre le stade ultérieur d'exploration plus fine et de mise en correspondance de différentes informations au sein de la scène visuelle soit moins performant, ce qui serait dû au déficit d'attention soutenue constitutif de ce trouble et à un défaut de désengagement. Il semble que le même contenu émotionnel soit extrait dans les deux groupes mais les enfants au développement typique vont ensuite préciser, enrichir leur compréhension de l'image et ainsi l'identification des émotions en augmentant les prises d'informations contrairement aux enfants TDA/H.

Ceci pourrait expliquer en partie certaines divergences dans les travaux s'intéressant aux capacités de catégorisation du contenu émotionnel chez les enfants TDA/H. L'évaluation de catégories émotionnelles d'un niveau superordonné (i.e., positif versus négatif) serait plus accessible pour les enfants TDA/H que l'accès à un niveau basique, plus précis (e.g., joie, peur, colère, dégoût, tristesse), pour lequel la différenciation entre deux émotions nécessite la comparaison d'informations plus détaillées et/ou la mise en lien avec d'autres informations de l'image (Widen, 2013; Widen & Russell, 2008).

Pour les deux groupes, comparativement aux images neutres, la catégorisation émotionnelle des images positives est facilitée (meilleure exactitude et temps de réponse plus court) par rapport à celle des images négatives. Différentes propositions peuvent être avancées pour expliquer les différences de performances de catégorisation du contenu émotionnel des images positives par rapport aux négatives. Tout d'abord, une reconnaissance et une orientation préférentielle vers les images positives ont été décrites au cours du développement (Vesker et al., 2018). Il a également été suggéré que la présence d'un contenu émotionnel négatif mobilise plus de ressources attentionnelles que la présence d'un contenu positif ou neutre (Pessoa, 2009; Schel & Crone, 2013). Cela pourrait en partie s'expliquer par le fait que la détection d'une valence négative, et donc potentiellement de danger, nécessite un traitement plus précis et détaillé afin de s'assurer qu'il ne s'agit pas d'une menace (Ihssen & Keil, 2013). Enfin, il a également été avancé que le traitement des informations négatives nécessite des traitements sous-corticaux supplémentaires avant d'accéder aux traitements corticaux, allongeant ainsi le temps total de traitement de l'information (Adolphs, 2002).

Enfin, nous obtenons un effet du niveau de l'image, les images de type objet présentent une exactitude de catégorisation supérieure à celle des images de type contexte. Inférer le contenu émotionnel d'informations contextuelles peut en effet constituer une activité moins aisée et moins habituelle pour des enfants d'âge scolaire que de juger le caractère plaisant ou déplaisant d'une personne ou d'un animal. Nous obtenons également un effet de cohérence émotionnelle uniquement pour les contextes. Ainsi, il semble que, lorsque l'attribution d'une émotion est plus complexe, tel que pour juger l'émotion véhiculée par un lieu, l'enfant va également traiter et associer d'autres informations de son champ visuel pour aider la réalisation de la tâche. Il a en effet été suggéré un impact plus important des connaissances antérieures, donc des représentations activées par les associations sémantiques, lorsque le stimulus est plus ambigu ou plus difficile à reconnaître (Summerfield & de Lange, 2014).

VIII.4 Discussion intermédiaire

Cette étude a évalué les capacités d'enfants TDA/H à catégoriser sémantiquement et émotionnellement des images qui apparaissent rapidement et de manière latéralisée dans les champs visuels droits et gauches. Avant de discuter les résultats, il est important de rappeler que ceux-ci sont préliminaires et seront confirmés, ou infirmés, suite aux prochaines inclusions.

Ces premiers résultats mettent en avant des performances plus faibles du groupe TDA/H comparativement au groupe contrôle lors de la tâche de catégorisation sémantique. Ce déficit

ne s'est pas répercuté sur les capacités de catégorisation émotionnelle ni sur la présence de l'effet de cohérence émotionnelle, qui ne diffèrent pas entre les deux groupes, renforçant ainsi la notion d'une indépendance des traitements émotionnels par rapport aux sémantiques. Les performances plus faibles en tâche de catégorisation sémantique chez les enfants TDA/H peuvent s'expliquer par la présence d'impulsivité des réponses (exactitude plus faible associée à des temps de réponses plus courts). Plusieurs études néanmoins suggèrent qu'un déficit d'encodage des informations serait à l'origine de plus faibles capacités de mémoire à long terme dans le TDA/H (pour une revue, Skodzik, Holling, & Pedersen, 2017), ce qui pourrait également expliquer les plus faibles capacités d'accès aux représentations mis en évidence dans notre étude.

Cette étude suggère également que les capacités de traitement, volontaire et non volontaire, des informations émotionnelles sont préservées chez les enfants TDA/H lorsque la prise d'information est rapide. Ainsi lorsqu'une image apparaît dans son champ visuel, l'enfant TDA/H est capable de la traiter afin de lui attribuer une valence émotionnelle, de manière comparable au développement typique. Il semble ainsi que les difficultés de traitements des émotions, fréquemment décrites dans la littérature qui s'est intéressée aux émotions faciales et aux scènes émotionnelles (voir article III, chapitre III), ne proviennent pas de l'altération de l'extraction et du traitement initial du contenu émotionnel, mais d'étapes ultérieures du traitement des informations visuelles. Ces étapes, dépendantes des capacités d'attention soutenue et spatiale (organisation de l'exploration visuelle), seraient plus sensibles aux déficits constitutifs du trouble. Ces résultats de la littérature sont importants pour la clinique du TDA/H et la mise en place de protocoles de remédiation du déficit de reconnaissance des émotions. L'efficacité préservée de l'extraction et de l'encodage initial des informations visuelles dans le TDA/H est un aspect encourageant. Les protocoles de remédiations portant sur l'amélioration des capacités d'attention soutenue et d'organisation de l'exploration visuelle sont donc pertinents car, en permettant d'augmenter le temps et l'efficacité de l'exploration visuelle, ils favoriseront probablement l'identification des émotions, tout en améliorant les capacités attentionnelles nécessaires aux apprentissages scolaires. Il sera, de plus, important d'y associer un travail plus spécifique sur l'exploration d'images émotionnelles et sur la mise en commun de différentes sources d'informations permettant une compréhension plus précise du contenu émotionnel.

Enfin, les enfants TDA/H présentent des capacités équivalentes à celles des enfants au développement typique quant à la prise en compte d'autres images présentes dans le champ

visuel. Ainsi, dans la tâche de catégorisation sémantique, aucun effet de cohérence sémantique n'a été mis en avant pour aucun des deux groupes, ce qui suggère un traitement ciblé de l'image à juger. Pour la tâche de catégorisation émotionnelle, les enfants des deux groupes ont la capacité de traiter simultanément le contenu émotionnel de deux images. Cet effet, dit effet de cohérence émotionnel, est identique entre les deux groupes, ce qui met en avant la capacité d'enfants qui présentent un TDA/H à analyser simultanément deux sources d'informations différentes, chacune dans un champ visuel différent et ainsi adressée initialement à un hémisphère différent. Il a été suggéré des capacités de transfert inter-hémisphérique moindres dans le TDA/H par rapport au développement normal (Hale, Loo, Zaidel, Hanada, Maciona & Smalley, 2009). Nos résultats, cependant, suggèrent que, au moins pour les informations émotionnelles, l'intégration des informations d'un hémisphère à un autre semble se réaliser de façon équivalente par des enfants TDA/H et des enfants au développement typique.

DISCUSSION GENERALE

Dans ce dernier chapitre, nous rappellerons notre démarche expérimentale et synthétiserons les principaux résultats de la partie expérimentale de ce document. Nous présenterons ensuite une vision plus intégrative de l'ensemble des résultats, en détaillant les implications théoriques en regard de la littérature existante.

1. Rappel de la démarche et synthèse des résultats

Le premier article de ce manuscrit (article I, chapitre I) est une revue de littérature dans laquelle nous avons proposé une synthèse critique des connaissances actuelles sur les traitements sémantiques et émotionnels des scènes visuelles. Cette synthèse nous a permis de mettre en avant un processus dynamique et interactif dans lequel les contenus perceptifs, sémantiques et émotionnels extraits des scènes visuelles interagissent. Cette synthèse nous a également permis de mettre en évidence l'importance des liens sémantique et émotionnel entre les objets et le contexte des scènes visuelles pour faciliter respectivement leurs processus de catégorisation sémantique et émotionnelle. Cependant, les mécanismes sous-jacents à ces effets d'associations sont encore sources de débats théoriques.

Notre seconde revue de la littérature (article II, chapitre II) nous a permis de synthétiser les résultats des travaux principaux concernant le développement des processus de catégorisation de la naissance à l'adolescence. Notre analyse de ces travaux nous permet de retenir une prédominance des traitements perceptifs (bottom-up) lors de tâches de catégorisation ou d'exploration visuelles chez les enfants et ce, jusqu'à la fin de l'adolescence. Les catégories sémantiques et émotionnelles s'affinent progressivement, ce qui semble lié aux expériences de vie et à l'accroissement de l'influence des traitements de plus haut niveau cognitif (top-down). Nous avons également mis en évidence une quasi-absence de travaux s'intéressant aux liens entre les objets et le contexte des scènes visuelles et ainsi à la façon dont l'enfant intègre l'ensemble des informations présentes dans son champ visuel. Notre troisième article (article III, chapitre III) est une revue systématique des études empiriques portant sur le traitement des scènes visuelles dans le TDA/H. De nombreuses études se sont intéressées au traitement volontaire et non volontaire des informations émotionnelles dans le TDA/H en utilisant des visages comme matériel expérimental. La compréhension des expressions faciales est en effet d'importance primordiale pour les relations sociales, cependant au quotidien l'individu doit traiter des contenus émotionnels provenant de sources de nature différente. Nous avons pu montrer dans cette revue que, bien que peu étudié, le traitement du contenu émotionnel des scènes visuelles semble également impacté dans ce trouble (i.e., performances de

catégorisation émotionnelle plus faibles et influence des images négatives plus importante chez les TDA/H que chez les enfants au développement typique). Les particularités de traitement des émotions chez l'enfant et l'adolescent avec TDA/H ne semblent donc pas spécifiquement liées à des difficultés de reconnaissance des émotions faciales. Cette synthèse suggère plutôt une altération générale des traitements visuels émotionnels dans le TDA/H, c'est-à-dire même lorsque le contenu émotionnel est porté par des informations contextuelles et par d'autres stimuli que des visages (e.g., posture corporelle).

A partir de l'ensemble des éléments mis en évidence dans nos revues de littérature, nous avons défini la problématique de la thèse (cf. chapitre IV). L'objectif général est de participer à la compréhension des processus à l'œuvre lors de la perception visuelle rapide de scènes complexes, dans le fonctionnement typique à l'âge adulte et au cours du développement durant l'âge de scolarisation en primaire (8 à 12 ans), ainsi que dans le cadre du TDA/H. Nous nous sommes ainsi intéressé à l'existence d'interactions entre les traitements sémantiques et émotionnels lors de la présentation rapide, hors de la vision fovéale (i.e., vision centrale qui permet l'analyse fine des détails), d'une scène du quotidien, ainsi qu'à la façon dont les constituants de cette scène (i.e., les objets et les informations contextuelles) participent à ces éventuelles interactions. Pour étudier notre problématique générale, nous avons défini quatre objectifs. Il s'agissait de préciser les mécanismes sous-jacents à l'influence des relations sémantiques (objectif 1) et émotionnelles (objectif 2) entre l'objet et le contexte des scènes visuelles, d'étudier les interactions entre les traitements volontaires et non volontaires des informations sémantiques et émotionnelles (objectif 3) et d'analyser les capacités précoces de traitement des informations émotionnelles dans le TDA/H (objectif 4). Quatre études expérimentales ont été réalisées afin d'aborder ces questions (cf. tableau 2).

	Tâche sémantique				Tâche émotionnelle			
	Image cible		Image distractrice		Image cible		Image distractrice	
Etude	Dimension pertinente : sémantique	Dimension non pertinente : émotionnelle	Dimension pertinente : sémantique	Dimension non pertinente : émotionnelle	Dimension pertinente : émotionnelle	Dimension non pertinente : sémantique	Dimension pertinente : émotionnelle	Dimension non pertinente : sémantique
1 : associations sémantiques chez l'adulte	Contexte (C) > Objet (O) (exactitude)		Effet de cohérence sémantique (CS)					
2 : traitements sémantiques et émotionnels chez l'adulte	O > C (exactitude) Sauf pour images négatives où O=C	Images négatives : performances diminuées (i.e., exactitude diminuée et temps de réponse allongé)	CS pour les contextes (temps de réponse)	Pas d'effets significatifs	Exactitude > et temps de réponse < des images positives par rapport aux neutres et des neutres par rapport aux négatives	Performances O > C (exactitude > et temps de réponse plus courts) Sauf pour images positives où C=O (en exactitude et en temps de réponse)	Effet de cohérence émotionnelle (CE) (exactitude)	CS pour les contextes négatifs (temps de réponse)
3 : traitements sémantiques et émotionnels chez l'enfant au développement typique	O > C (exactitude) C négatifs > O (temps de réponse)	Images négatives : performances diminuées	Pas d'effets significatifs		Exactitude > et temps de réponse < des images positives par rapport aux neutres et aux négatives. Exactitude images neutres > négatives	O > C (exactitude) Sauf pour images positives où C=O (en exactitude)	CE (exactitude)	
4 : traitements sémantiques et émotionnels chez l'enfant présentant un TDA/H	O > C (exactitude) TDA/H : exactitude < contrôles, et pour C temps de réponse < contrôles	Images négatives : performances diminuées	Pas d'effets significatifs		Images positives > neutres > négatives (exactitude)	O > C (exactitude)	CE (exactitude)	

Tableau 1 : Synthèse des principaux résultats des quatre études empiriques. Pour chaque tâche expérimentale, les résultats en fonction des dimensions sémantique et émotionnelle de l'image cible et de l'image distractrice présentée simultanément sont décrits. Les résultats encadrés en bleu correspondent au premier objectif de la thèse, ceux en rouge au second et ceux en vert au troisième. Le quatrième objectif n'a pas été encadré par soucis de lisibilité, il correspond à la ligne de l'article 4. Les cases grisées indiquent les dimensions non évaluées.

Notre première étude empirique (article IV, chapitre V) a analysé l'effet facilitateur de la cohérence sémantique, qui décrit des performances de catégorisation supérieures lorsque l'objet et le contexte d'une scène visuelle sont présentés en condition de cohérence sémantique (e.g., un cochon et une ferme) par rapport à une condition d'incohérence sémantique (e.g., un cochon et une banquise). Dans cette étude nous avons manipulé la cohérence sémantique entre un contexte et un objet présentés rapidement, séparément et simultanément dans une tâche de catégorisation sémantique où les réponses étaient non contraintes par un choix limité. Nos résultats enrichissent ainsi la compréhension des mécanismes sous-jacents aux influences des associations sémantiques objet-contexte (objectif 1) en suggérant que l'avantage de la cohérence sémantique pour la catégorisation du contexte et de l'objet se produit au stade cognitif de l'appariement d'un percept avec les connaissances antérieures stockées en mémoire à long terme. Cette étude a ainsi mis en évidence qu'une représentation unifiée de la scène visuelle est formée même lorsque les constituants de cette scène (les objets et le contexte) sont présentés séparément.

Dans notre seconde étude empirique (article V, chapitre VI), nous avons proposé une tâche de catégorisation sémantique et une tâche de catégorisation émotionnelle d'un objet et d'un contexte présentés rapidement, séparément et simultanément, chacun dans un champ visuel différent. Chaque image cible de niveau contexte ou objet était présentée avec une image de l'autre niveau, suivant les conditions expérimentales, en condition de cohérence ou d'incohérence sémantique et de cohérence ou d'incohérence émotionnelle. Les liens sémantiques et émotionnels entre l'objet et le contexte étaient manipulés de façon orthogonale dans les deux tâches.

Dans la tâche de catégorisation sémantique, nous avons montré un effet de cohérence sémantique sur les temps de réponse des images de contextes, avec une présentation encore plus courte que dans l'étude 1 et des réponses contraintes par un choix manuel forcé entre des catégories prédéfinies. Ce résultat suggère à nouveau la robustesse de l'effet de cohérence sémantique, qui ne dépend ni du matériel expérimental (les images étant différentes entre la première et la seconde étude empirique) ni de la modalité de réponse (objectif 1). Dans la même tâche, nous avons également mis en évidence un traitement non volontaire du contenu émotionnel de l'image cible (performances plus faibles pour les images négatives que pour les images neutres et positives). Le traitement non volontaire du contenu émotionnel influence donc les traitements sémantiques et mobilise des ressources attentionnelles (objectif 3).

Au regard de notre second objectif, les résultats à la tâche de catégorisation émotionnelle ont mis en avant un effet du contenu émotionnel du contexte sur la catégorisation du contenu

émotionnel de l'objet (voir aussi Righart & de Gelder, 2008) mais également, de manière originale, un effet du contenu émotionnel de l'objet sur la catégorisation émotionnelle du contexte. Les traitements des contenus émotionnels de l'objet et du contexte interagissent donc, y compris lorsque ces constituants potentiels de la scène sont présentés séparément. Nous montrons également dans cette tâche un effet de cohérence sémantique entre l'objet et le contexte sur la catégorisation émotionnelle des contextes négatifs. Ce résultat suggère que la catégorisation émotionnelle d'une image nécessite, au moins pour les images probablement les plus complexes à évaluer et/ou les plus ambiguës, une étape préalable de traitement sémantique de l'image (objectif 3), ce qui pourrait corroborer les théories de la primauté cognitive (Lazarus, 1984). Enfin, les traitements des liens sémantiques et émotionnels entre l'objet et contexte ne s'influencent pas. Ce résultat suggère que les effets de cohérence sémantique et de cohérence émotionnelle dépendraient de processus distincts et se réaliseraient sur des contenus de nature différente.

Notre troisième étude empirique, réalisée auprès d'enfants d'âge scolaire (8 à 12 ans) au développement typique (article VI, chapitre VII), reprend dans ses principes généraux la méthodologie que l'étude 2. Cependant, dans la tâche de catégorisation sémantique, seule la cohérence sémantique est manipulée (tous les essais sont cohérents émotionnellement) et, dans la tâche de catégorisation émotionnelle, seule la cohérence émotionnelle est manipulée (tous les essais sont cohérents sémantiquement). Cette étude, à l'instar de l'étude 2, suggère un traitement non volontaire du contenu émotionnel négatif (objectif 3). Nous ne mettons pas en évidence d'effet de cohérence sémantique (objectif 1) dans la tâche de catégorisation sémantique, mais un effet de cohérence émotionnelle (objectif 2) est montré dans la tâche de catégorisation émotionnelle. Ceci suggère des trajectoires différentes dans le développement des associations sémantiques et émotionnelles entre objets et contextes. L'effet de cohérence sémantique semble lié aux apprentissages, alors que l'effet de cohérence émotionnelle est présent plus tôt dans le développement. Nous reviendrons plus en détails sur ces aspects dans la partie « implications théoriques » de cette discussion.

Enfin, dans notre dernière étude (étude 4 en cours, chapitre VIII) nous avons comparé les performances d'enfants présentant un TDA/H aux performances d'enfants au développement typique (i.e., groupe contrôle), selon la même méthodologie que celle de l'étude 3. Nos résultats montrent un traitement non volontaire du contenu émotionnel à la tâche sémantique et des performances de catégorisation du contenu émotionnel de scènes visuelles (i.e., traitement volontaire) à la tâche émotionnelle comparables entre les deux groupes. Les enfants qui présentent un TDA/H semblent donc extraire et traiter le contenu émotionnel d'une

scène visuelle présentée pour un temps très court de façon comparable à ce que font des enfants au développement typique. Cela suggère que les difficultés d'identification des émotions, fréquemment relevées dans le TDA/H (pour des revues: Martel, 2009; Shaw, Stringaris, Nigg, & Leibenluft, 2014; van Stralen, 2016), ne proviendraient pas d'un défaut dans la prise initiale d'information (objectif 4).

Nous allons maintenant situer ces résultats en regard de la littérature existante.

2. Implications théoriques

2.1 Traitements sémantiques des scènes visuelles

Différents modèles théoriques ont été proposés dans la littérature pour décrire les relations sémantiques entre objet et contexte au sein des scènes visuelles. Les études réalisées dans le cadre de cette thèse nous permettent de nous positionner face à ce débat théorique. Nos résultats sont ainsi en faveur des modèles de sélection de l'objet, appelés aussi modèles d'amorçage (Bar, 2004; Bar & Aminoff, 2003 ; Palmer, 1975 ; Ullman, 1996), qui situent les associations sémantiques objet-contexte au stade d'une mise en correspondance entre un percept et une représentation en mémoire à long terme.

Ces modèles de sélection de l'objet décrivent que le traitement initial rapide d'informations grossières véhiculées par les basses fréquences spatiales (BFS) des constituants de la scène permet d'activer des représentations et d'élaborer des prédictions sur les éléments associés à cette scène. Ces prédictions vont permettre de diminuer la quantité d'information perceptive nécessaire à la reconnaissance du stimulus si celui-ci correspond aux hypothèses. Dans le cas inverse, c'est-à-dire en présence d'un stimulus incohérent sémantiquement, les prédictions ne correspondront pas à l'image à reconnaître et vont ralentir la catégorisation ou entraîner des erreurs de catégorisation. Ainsi, à partir de la même représentation approximative de l'objet, une image, par exemple de champignon, sera catégorisée plus facilement qu'une image de parapluie si elle est présentée avec un contexte de forêt. Et réciproquement pour la catégorisation de la scène visuelle. La perception visuelle ne se réaliserait ainsi pas uniquement dans un sens ascendant de traitement du contenu visuel progressivement le long de la voie ventrale pour accéder à l'identification de l'objet jusqu'à la prise de décision dans le cortex pré-frontal, des influences cognitives descendantes modèleraient également très tôt le traitement de l'information visuelle (cf. figure 2). Ces influences descendantes, *top-down*, seraient possibles grâce aux voies magnocellulaires qui transmettent rapidement des informations grossières et qui permettraient d'activer des représentations mnésiques via notamment le cortex orbito

frontal (COF) et le cortex parahippocampique (e.g., Bar, 2004; Chaumon, Kveraga, Barrett, & Bar, 2014; Kauffmann, Ramanoël, & Peyrin, 2014; Trapp & Bar, 2015). Il a été proposé que le COF aurait une fonction primordiale dans la facilitation cognitive de la perception, en permettant d'intégrer les informations perceptives et les représentations mnésiques afin de générer des prédictions (Barrett & Bar, 2009; Trapp & Bar, 2015). Cependant, le COF ne serait activé que pour des stimuli connus de l'observateur (Chaumon et al., 2014), ce qui met en évidence l'importance des expériences passées et des apprentissages nécessaires à l'élaboration d'analogies et de prédictions (e.g., Bar, 2007). Grâce à ses expériences sensorimotrices avec les objets de son quotidien et par la répétition de co-occurrences entre certains objets et contextes, l'enfant va progressivement développer ses concepts sémantiques (Kalénine, 2009). Ainsi, les capacités de catégorisation sémantique se développent au fur et à mesure des expériences de vie.

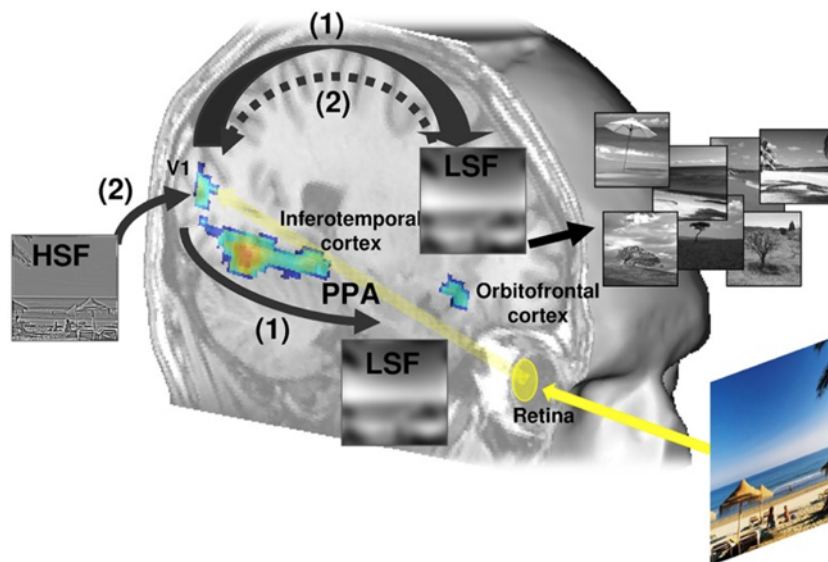


Figure 2 : Représentation schématique d'un modèle prédictif de facilitation cognitive de la perception visuelle.

(1) Les informations constituées de basses fréquences spatiales (BFS- LSF pour l'anglais « low spatial frequencies » et dans le schéma) rejoignent rapidement des aires de haut niveau cognitif de la voie visuelle dorsale pour permettre une première impression grossière de la scène (dont des interprétations possibles de la catégorie de la scène via notamment le COF), avant que la propagation réalisée en parallèle de ces mêmes BFS ne soit complétée le long de la voie ventrale (cortex inféotemporal), ce qui modulera au final la reconnaissance de la scène. (2) Cette analyse initiale grossière active les prédictions sur les scènes candidates, qui sont similaires dans leur apparence BFS, et qui seront ensuite « rétro-injectées » par des boucles retour dans des aires de plus bas niveau incluant le cortex visuel primaire (V1) pour guider l'analyse plus lente des hautes fréquences spatiales (HFS- HSF pour l'anglais « high spatial frequencies » et dans le schéma) le long de la voie visuelle ventrale en permettant la sélection des détails nécessaires pour la reconnaissance de la scène. Il est à préciser que les mêmes processus sont à l'œuvre pour la reconnaissance des objets.

Illustration tirée de Kaufmann et al. (2014).

Au niveau développemental, il a été décrit que l'enfant et l'adolescent montrent des performances de catégorisation sémantique d'une image présentée rapidement comparables à celles des adultes en terme d'exactitude, mais avec des temps de réponse allongés (Batty & Taylor, 2002; Vanmarcke & Wagemans, 2015). Notre troisième étude ajoute à cette littérature que cette capacité de catégorisation sémantique est mise en évidence y compris lorsque deux images sont présentées simultanément et ce, dès 8 ans. Nous montrons ainsi que l'enfant a la capacité de traiter en parallèle des images différentes, à minima jusqu'à ce que l'image pertinente à la tâche soit sélectionnée, et d'accéder à une représentation sémantique, donc de haut niveau, de cette image. Nous avons montré des performances de catégorisation sémantique comparables en termes d'exactitude entre enfants et adultes mais avec des temps de réponse allongés. Nous pouvons proposer que cette différence d'efficacité des traitements sémantiques entre enfants et adultes s'explique en partie par l'absence de mise en relation des contenus sémantiques entre les différents constituants de la scène visuelle. Les réseaux associatifs sémantiques sont moins développés chez les enfants, comparativement aux adultes, ce qui est probablement lié à la maturation des structures cérébrales impliquées dans l'encodage et la récupération telles que l'hippocampe, le cortex préfrontal et les régions visuelles associatives (Casey, Giedd, & Thomas, 2000; Rosen et al., 2018; Schlichting, Guarino, Schapiro, Turk-Browne, & Preston, 2017). Nous faisons donc l'hypothèse d'un traitement visuel majoritairement ascendant chez les enfants. Ainsi l'entrée perceptive serait traitée le long de la voie ventrale jusqu'à son identification mais le processus de catégorisation semble peu bénéficier de prédictions qui permettraient de gagner en vitesse de reconnaissance. Dans ce sens certains travaux ont suggéré que le déroulement des processus de catégorisation sémantique chez l'enfant présenterait des différences du point de vue des structures cérébrales activées par rapport à l'adulte. Notamment, les différences de patterns de recrutement des réseaux neurocognitifs sont en faveur d'une perception moins dirigée par des processus de haut niveau chez l'enfant (Ofen & Shing, 2013). En outre il a été montré que, lors de la perception de scènes visuelles, les régions préfrontales sont moins activées et les régions visuelles associatives présentent moins de spécificité d'activation (Casey et al., 2000; Rosen et al., 2018; Schlichting et al., 2017). D'autres auteurs ont mis en avant qu'au cours du développement le nombre d'aires cérébrales activées lors de la réalisation d'une tâche de catégorisation sémantique diminue, ce qui permettrait des réponses plus rapides et moins coûteuses (Batty & Taylor, 2002). L'ensemble de ces résultats suggère donc une trajectoire dans l'édification du processus de catégorisation sémantique dont le développement semble dépendre de la maturation cérébrale et des expériences de vie.

2.2 Traitements émotionnels des scènes visuelles

2.2.1 Distinction entre traitements volontaire et non volontaire des scènes émotionnelles

L'ensemble des résultats de ce travail de thèse suggère que des processus différents opèrent lors d'un traitement non volontaire (dirigé par le stimulus, bottom-up) et lors d'un traitement volontaire (dirigé par l'observateur, top-down) des informations émotionnelles. D'une part, une diminution des performances à la tâche de catégorisation sémantique des images à contenu émotionnel négatif a été mise en évidence chez l'adulte et l'enfant, ce qui suggère une détection priorisée non volontaire du contenu émotionnel. D'autre part, l'exactitude de la catégorisation du contenu émotionnel, qui nécessite un traitement volontaire des informations émotionnelles, est plus faible que pour la tâche de catégorisation sémantique aussi bien chez l'adulte que chez l'enfant. De plus, la catégorisation du contenu émotionnel est associée à des temps de réponse allongés chez l'adulte par rapport à la catégorisation du contenu sémantique. Ces résultats suggèrent une distinction entre une détection non volontaire, probablement automatique du contenu émotionnel et un traitement explicite plus complexe de ce contenu. Il est à noter que ces niveaux distincts de traitements volontaires ou non des informations émotionnelles font échos aux travaux sur le développement du traitement des informations émotionnelles. Ainsi la seconde revue de littérature (article II) de cette thèse a permis de distinguer des traitements non volontaires du contenu émotionnel présents dès la naissance, avec la présence d'un biais négatif quasiment identique à celui de l'âge adulte dès la période préscolaire (3-4 ans), alors que les capacités de catégorisation volontaire des émotions se développent jusqu'à la fin de l'adolescence. L'ensemble de ces résultats différencie donc un traitement non volontaire du contenu émotionnel qui semble inné, automatique donc probablement peu coûteux en ressources attentionnelles, d'un traitement explicite, volontaire qui se développe avec les expériences de vie et semble dépendre de processus plus contrôlés donc probablement plus coûteux en ressources attentionnelles (Schneider, & Shiffrin, 1977). Nos résultats peuvent être interprétés dans le cadre des modèles dits « double-système » (Gawronski & Bodenhausen, 2006; March, Gaertner, & Olson, 2018) qui décrivent des effets comportementaux, en potentiels évoqués (PE) et des activations en IRMf différents selon que le contenu émotionnel soit traité de façon volontaire ou non (Scheuerecker et al., 2007). Des études en PE ont ainsi décrit des réponses précoces (notamment de 100 à 300 ms post onset-stimulus) et des réponses plus tardives (entre 300 et 700 ms, Late Positive Potential) à des stimulations émotionnelles (pour une revue voir: Olofsson, Nordin, Sequeira, & Polich, 2008). Il a également été mis en avant des activations cérébrales différentes suivant le type de tâche

proposée aux participants, i.e., selon qu'elle sollicite un traitement volontaire ou non du contenu émotionnel. L'amygdale et l'aire frontale inférieure droite/insula sont activées lors de la présence d'un contenu émotionnel quel que soit le type de tâche, ce qui suggère des traitements implicites d'évaluation du contenu émotionnel. Par contre, les aires latérales du COF et le cortex cingulaire antérieur ne seraient activés que lors de l'évaluation volontaire du contenu émotionnel (Cunningham, Raye, & Johnson, 2004) qui nécessite une activité réflexive et lors d'ambiguïté ou pour des émotions complexes (telles que la jalousie ou le remords). Ainsi il semble qu'une évaluation non volontaire du traitement émotionnel se réalise rapidement de façon directe (feedforward), suivie par une évaluation volontaire plus lente car qui nécessite des processus plus contrôlés, passant par des structures cérébrales telles que le COF associées notamment au traitement des représentations sémantiques, et/ou des boucles de retour (feedback) pour affiner la prise d'informations.

Un premier traitement rapide, approximatif, permettrait donc de détecter la présence d'un contenu émotionnel afin de prioriser son traitement si cela est pertinent. L'évaluation volontaire du contenu émotionnel nécessiterait parfois un accès à une représentation plus fine de l'image visuelle. Cependant, bien que la rapidité de traitement du contenu émotionnel soit largement démontrée, le déroulement de ce traitement et les voies recrutées pour permettre cette rapidité de traitement sont encore débattus. Certains auteurs suggèrent que le traitement du contenu émotionnel se réaliserait via une voie sous-corticale, qui transmettrait des informations grossières en BFS, notamment à l'amygdale, avant que le traitement visuel cortical ne soit finalisé (e.g., Tamietto & de Gelder, 2010). Comme l'amygdale présente des connexions réciproques avec de nombreuses régions corticales dont les réseaux perceptifs (e.g., cortex visuel primaire, cortex temporal inférieur) et les régions préfrontales (e.g., cortex orbito-frontal, cortex cingulaire antérieur), des influences en provenance de l'amygdale peuvent renforcer le traitement perceptif initial et la représentation des stimuli émotionnels. D'autres auteurs ont proposé des modèles alternatifs en spécifiant que la rapidité de traitement des informations émotionnelles pourrait se réaliser via des traitements cortico-corticaux. Le modèle à double niveau (Pourtois et al., 2012) s'inscrit dans cette hypothèse : il décrit un traitement initial grossier des informations visuelles émotionnelles via la voie visuelle dorsale pour atteindre des régions corticales de haut niveau, telles que le COF, qui modulerait ensuite, par des activations descendantes, le fonctionnement des aires visuelles primaires et déclencherait des processus d'évaluation émotionnelle sous-tendus par l'amygdale. Bar et collaborateurs ont proposé une extension de leur modèle de la perception visuelle au traitement des objets et des scènes à contenu émotionnel (Barrett & Bar, 2009). Ce modèle constructiviste intègre une approche

incarnée de la reconnaissance visuelle. En effet, selon l'hypothèse de ce modèle, l'état affectif de la personne influencerait le processus de catégorisation. Le ressenti corporel serait une source supplémentaire pour faciliter la reconnaissance de l'objet ou de la scène. Enfin, Pessoa et Adolphs (2010) ont proposé un modèle « multi-vagues » qui se base sur la connectivité tout au long de la voie visuelle ventrale. Selon ces auteurs, l'ensemble de ces connexions permet des séries d'activations à l'origine de la rapidité de traitement du contenu émotionnel, via notamment des connexions entre l'amygdale, le COF, l'insula et le cortex cingulaire antérieur. Le recrutement de ces régions permet de renforcer le traitement des stimuli émotionnels. En outre, les informations émotionnelles seraient initialement traitées à partir des BFS, ce qui permettrait un traitement rapide.

Bien que ces différents modèles soient encore largement débattus, il semble néanmoins admis que le traitement initial du contenu émotionnel est permis par la transmission rapide d'information grossières provenant des BFS via les voies magnocellulaires (Beffara et al., 2015; Mermillod et al., 2017).

2.2.2 Distinction entre le traitement du contenu émotionnel positif et du contenu émotionnel négatif des scènes visuelles

Des différences entre traitements volontaire et non volontaire selon le contenu émotionnel (positif ou négatif) des images sont mises en évidence dans nos travaux. Nous avons ainsi montré une interférence du traitement non volontaire du contenu émotionnel sur les performances de catégorisation sémantique, même si uniquement pour le contenu émotionnel négatif. En outre, chez les adultes, l'effet d'interférence du contenu émotionnel négatif est plus marqué pour les objets que pour les contextes. Ce résultat peut s'expliquer par un impact plus fort des images représentant des objets vivants, donc potentiellement plus menaçants, que des objets non vivants (Mermillod et al., 2010) et donc probablement que des contextes. Les images d'objets négatifs doivent donc être traitées prioritairement et leur interférence sur la tâche en cours est plus marquée que celle des images de contextes négatifs. Concernant les traitements volontaires (i.e., catégorisation du contenu émotionnel), des travaux récents en modélisation ont montré l'influence des processus descendants pour aider à la catégorisation émotionnelle en situation d'ambiguïté, à part pour les stimuli positifs, qui bénéficient d'un avantage de catégorisation (Mermillod et al., 2019). Cet avantage de catégorisation du contenu émotionnel positif a été largement démontré dans la littérature, notamment avec des visages (e.g. Leppänen & Hietanen, 2004). Il a également été décrit que les émotions positives élargissaient le focus attentionnel et pouvaient améliorer les performances cognitives (Fredrickson, 2001). Nos

travaux, de manière originale, ajoutent à cette littérature que cet avantage est également mis en évidence chez l'enfant et chez l'adulte, quel que soit le niveau de l'image à juger (objet ou contexte). De plus, cet avantage est montré même pour un temps de présentation très court et lorsque deux images sont présentées simultanément, ce qui implique un traitement en parallèle des images, au moins jusqu'au moment où l'image cible sera identifiée. Dans toutes nos études, chez l'adulte et l'enfant, les performances de catégorisation du contenu émotionnel des images de contextes négatives et neutres sont plus faibles que celles des objets négatifs et neutres, respectivement, alors que les performances de catégorisation des contextes et objets positifs sont équivalentes. Nos travaux suggèrent donc que le contenu émotionnel positif est extrait rapidement et de manière semblable quel que soit le niveau de l'image (local ou global).

Pour résumer, il semblerait qu'une détection initiale de la présence de contenu émotionnel s'opère lors de l'apparition d'une image par l'amygdale ; si ce contenu est négatif, qu'il soit pertinent ou non pour l'activité en cours, le traitement de l'image sera renforcé en partie par des feed-back de l'amygdale sur les régions corticales perceptives et, cela, de façon plus marquée pour les images d'objets animés potentiellement plus menaçants pour l'individu. Ce traitement nécessite des ressources attentionnelles supplémentaires ce qui peut expliquer l'effet d'interférence de la présence d'un contenu émotionnel non pertinent à la tâche en cours. La détection d'un contenu positif, synonyme de sécurité et bénéfique pour l'individu, semble permettre une catégorisation rapide quel que soit le niveau de l'image. La catégorisation émotionnelle des images neutres et négatives est plus lente et moins exacte, surtout pour les contextes, et semble nécessiter, en moins en partie, l'accès à une représentation sémantique.

2.2.3 Compétition attentionnelle entre deux images émotionnelles

Les résultats de nos études permettent également de préciser certains aspects du traitement automatique non volontaire du contenu émotionnel. Notre seconde étude, en présentant de façon originale deux images émotionnelles simultanément et en demandant la catégorisation sémantique d'une seule de ces images, a permis d'étudier l'influence du contenu émotionnel, non pertinent à la tâche, non seulement de l'image cible mais également de l'image distractive associée. La littérature a maintes fois montré la capture attentionnelle d'une image émotionnelle par rapport à une image neutre, mais la compétition attentionnelle entre deux contenus émotionnels non pertinents à la tâche (portés par l'image distractive et l'image cible) n'a pas, à notre connaissance, été étudiée. Nous avons ainsi pu montrer que seul le traitement non volontaire du contenu émotionnel de l'image cible, image qui bénéficie de l'attention de l'observateur, interférait avec la catégorisation sémantique de cette image. Il semble donc que

le contenu émotionnel de l'image distractive ne soit pas traité, ou que celui-ci n'interfère pas avec les traitements en cours. Pour aller plus loin sur cette question et déterminer si le contenu émotionnel de l'image distractive a été traité, nous avons, dans des travaux annexes à cette thèse, analysé la reconnaissance à long terme des images. Nous avons étudié l'influence du contenu émotionnel (positif, négatif ou neutre) de l'image vue par le participant lorsqu'il réalisait la tâche de catégorisation sémantique, sur son rappel 15 jours plus tard. Il en ressort que les images négatives sont mieux rappelées que les images positives, même pour les images qui étaient distractives dans la tâche de catégorisation sémantique lors de l'encodage (cf. Annexe 13 : communication affichée des résultats de cette étude à la Société Française de Psychologie, 2017). Ces résultats suggèrent que le contenu émotionnel de l'image distractive est traité, la présence d'un contenu négatif renforce la mémorisation de l'image. Néanmoins, ce traitement n'interfère pas avec les processus de catégorisation sémantique de l'image cible. Par contre, lorsque l'attention du participant est orientée sur la dimension émotionnelle de l'image cible (tâche de catégorisation émotionnelle), nous mettons en évidence, chez l'adulte et chez l'enfant, un effet de la cohérence émotionnelle qui montre que le contenu émotionnel de l'image distractive a été traité et influence la catégorisation émotionnelle de l'image cible et ce, quel que soit le niveau de l'image cible et le niveau de l'image distractive (i.e., objet ou contexte). Ces résultats nous permettent d'enrichir la littérature sur l'effet de cohérence émotionnelle en précisant certains de ses mécanismes. La mise en évidence d'une influence réciproque du contenu émotionnel du contexte sur la catégorisation émotionnelle de l'objet et du contenu émotionnel de l'objet sur la catégorisation émotionnelle du contexte est importante, puisque cela suggère que le contenu émotionnel est extrait d'informations locales et globales (constituées de fréquences spatiales différentes), et ce même pour une présentation très brève. De plus cela nous permet de spécifier que l'étape de mise en relation des contenus émotionnels de l'objet et du contexte se situe à un stade précoce de la chaîne de traitement des informations émotionnelles. En effet, bien que les images de contextes négatifs sont moins bien catégorisées émotionnellement que les autres images émotionnelles, leur contenu émotionnel peut néanmoins faciliter la catégorisation émotionnelle des objets si ceux-ci sont porteurs du même contenu émotionnel.

Pour résumer, lorsque deux images émotionnelles apparaissent dans le champ visuel, l'observateur va traiter les deux contenus émotionnels, mais la mise en relation de ces informations émotionnelles dépendra de ses attentes ou de ses buts, et donc de l'orientation de l'attention endogène. Ces résultats peuvent également être interprétés selon les modèles cognitifs d'évaluation de la pertinence et plus particulièrement le modèle des processus

composants (*Component Process Model*; Scherer, 2001). En effet nos résultats ont mis en évidence des temps de réponse plus longs pour la catégorisation sémantique et émotionnelle des images négatives comparativement à ceux des images positives (i.e., pour le traitement non volontaire et volontaire du contenu émotionnel), ce qui pourrait aller à l'encontre d'un besoin d'action rapide face à un évènement menaçant. Il est ainsi probable que nos images ne présentent pas un niveau de pertinence fort pour l'individu. Leur contenu émotionnel est détecté, et traité tel que nous l'avons décrit ci-dessus, mais le niveau de pertinence n'est pas assez fort pour permettre le déclenchement d'un comportement de préparation à l'action (Mazzietti & Koenig, 2014). Il est ainsi probable que, lorsque l'individu oriente son attention sur le contenu émotionnel (i.e., tâche de catégorisation émotionnelle) la pertinence du contenu émotionnel des images augmente, ce qui expliquerait la mise en évidence de l'effet de cohérence émotionnelle uniquement lorsque l'individu oriente son attention sur le contenu émotionnel.

2.3 Associations sémantiques et émotionnelles entre l'objet et le contexte au sein des scènes visuelles

L'ensemble de nos travaux suggère des traitements indépendants des contenus sémantiques et émotionnels qui interagissent notamment en fonction des attentes ou buts de l'observateur.

En accord avec nos hypothèses, l'ensemble de nos résultats suggère que la question de la primauté des, et de la dépendance entre, traitements sémantiques et traitements émotionnels n'a pas de réponse simple, et que la priorisation de traitement d'une dimension, ainsi que les interactions entre les traitements sémantiques et émotionnels, dépendent en partie de l'orientation attentionnelle de l'observateur. L'avantage de la cohérence sémantique, spécifiquement sur la catégorisation émotionnelle des contextes négatifs, laisse penser que le passage par un traitement sémantique pour la catégorisation émotionnelle d'une image ne serait pas une étape obligatoire, mais se réaliserait pour les images plus complexes à évaluer (Mueller et al., 2017). En outre, nos travaux ne montrent pas d'influence entre les associations sémantiques et émotionnelles entre l'objet et le contexte ce qui suggère une indépendance entre les effets de cohérence sémantique et émotionnelle. Enfin, la mise en évidence d'un effet de cohérence émotionnelle au cours du développement (enfants de 8 à 12 ans) alors que l'effet de cohérence sémantique n'est pas montré, suggèrent que le processus de catégorisation émotionnelle se réalise indépendamment de la catégorisation sémantique, et ce, de façon

précoce dans le développement. Les résultats de l'étude 4 vont également dans ce sens. Dans cette étude, les enfants avec un TDA/H présentent des performances de catégorisation sémantique inférieures à celles des enfants au développement typique, mais des performances de catégorisation du contenu émotionnel semblables. De plus, l'effet de cohérence émotionnelle entre objet et contexte ne diffère pas entre les deux groupes. Si la catégorisation émotionnelle était secondaire à la catégorisation sémantique, les performances de catégorisation émotionnelle auraient dû être diminuées dans le groupe TDA/H par rapport au groupe contrôle de manière analogue à celles de catégorisation sémantique. Néanmoins, comme évoqué précédemment, lorsque l'évaluation émotionnelle de l'image s'avère complexe, le traitement sémantique de l'image réalisé en parallèle de son traitement émotionnel permet d'activer la représentation de l'image pour soutenir la catégorisation émotionnelle.

Ces résultats vont dans le sens de certains travaux en potentiels évoqués qui montrent des temps d'intégration des informations globales et locales différents selon qu'il s'agisse de traitements émotionnels ou sémantiques. Ainsi, la détection d'une incohérence entre l'émotion exprimée par un visage émotionnel et le contenu émotionnel du contexte dans lequel il est inclus a comme signature corticale une variation de l'amplitude de la N170 alors que la détection d'une incohérence sémantique suscite une modulation plus tardive de la N400 (Hietanen & Astikainen, 2013). Ces résultats suggèrent un traitement plus précoce, à un stade perceptif, de la cohérence émotionnelle comparativement à un stade de plus haut niveau pour la prise en compte de la cohérence sémantique (Hietanen & Astikainen, 2013; Righart & de Gelder, 2006). Aldunate, López et Bosman (2018) proposent ainsi que les effets de cohérence émotionnelle ne sont pas le reflet d'une intégration sémantique, mais qu'ils sont plus précoces et liés à un recrutement de ressources attentionnelles qui permettent de prioriser et de faciliter les traitements perceptifs des stimuli émotionnels. Il semblerait donc que l'effet de cohérence émotionnelle, au-delà de se distinguer de l'effet de cohérence sémantique, se déroule, au moins en partie, plus précocement. Cet effet paraît ainsi dépendre du stade initial de traitement automatique des émotions plutôt que du stade représentationnel (Aldunate, López & Bosman, 2018).

2.4 Synthèse

Pour synthétiser l'ensemble des résultats obtenus dans ces travaux de thèse et dans la littérature, nous regroupons les étapes de traitements suggérées par les résultats de nos études empiriques avec les données de la littérature afin de préciser et compléter les hypothèses sur

l'architecture fonctionnelle des processus cognitifs sous-jacents à ces étapes de traitement. Nous proposons ainsi un modèle neurocognitif dynamique de la perception visuelle dans lequel il s'agirait plus de déterminer à quel moment ou selon quelles conditions certains traitements vont être réalisés prioritairement à d'autres, plutôt que de décrire une séquentialité de traitements définie. L'âge de l'observateur, ses buts, l'image (e.g., plus ou moins difficile à évaluer), son contenu émotionnel et la présence d'une autre image (également porteuse de contenus sémantique et émotionnel) dans le champ visuel peuvent moduler le déroulement des traitements visuels.

1) Lors de la perception d'une scène visuelle, la pertinence de ses différents constituants est automatiquement évaluée. En fonction du contenu détecté, des ressources attentionnelles peuvent être mobilisées pour évaluer plus précisément le stimulus, avec un impact plus fort, selon nos résultats, des stimuli négatifs. Ce traitement priorisé du contenu émotionnel peut alors interférer avec d'autres traitements, tels que les processus de catégorisation dans nos études. Lorsque plusieurs images sont présentes simultanément, la mise en relation de leurs contenus émotionnels dépend de l'orientation attentionnelle de l'observateur (et/ou du niveau de pertinence de l'image). Cette étape de mise en relation permet l'effet de cohérence émotionnelle.

Nous faisons l'hypothèse que les basses fréquences spatiales extraites de l'image permettent une première détection rapide de la présence de contenu émotionnel par l'amygdale ; la projection de ces informations grossières à l'amygdale peut se réaliser via des voies sous-corticales, via la voie corticale dorsale et le COF et/ou par la voie corticale ventrale et le cortex inférotemporal. Ces informations grossières sont traitées rapidement avant que l'identification de l'image ne soit réalisée. L'amygdale, par ses connexions multiples avec les aires corticales, va alors renforcer le traitement perceptif du stimulus émotionnel et orienter les ressources attentionnelles. Bien que nos résultats montrent un effet d'interférence lié au contenu émotionnel négatif uniquement, il est important de préciser que l'amygdale permet de détecter la présence de contenu pertinent pour l'individu. Or, cette notion de pertinence varie en fonction des buts et besoins de l'individu qui sont réactualisés constamment. Ainsi, suivant les situations, des contenus même non émotionnels, pourront être priorisés (voir Mazzietti, Sellem & Koenig, 201). Si le stimulus est jugé pertinent, des mécanismes sensori-moteurs, tels que les comportements oculomoteurs pour positionner le stimulus émotionnel en vision fovéale et les activations physiologiques de préparation à l'action, se déclencheront. Selon le modèle des processus composants (Scherer, 2001) le ressenti émotionnel de l'individu dépendra alors de

l'adéquation entre l'expression émotionnelle, les réponses autonomes, physiologiques et les tendances à l'action qui font suite à l'évaluation de la pertinence du stimulus (pour une description de ces composants voir Sander, Grandjean & Scherer, 2018). Le ressenti émotionnel peut ensuite impacter les traitements cognitifs (e.g., action sur l'attention, la mémoire, la prise de décision...).

2) Parallèlement à la dynamique d'opérations proposée dans l'étape 1, les traitements visuels ascendant et descendant permettent l'activation de la représentation en mémoire du stimulus et ainsi sa catégorisation sémantique. Chez l'adulte, la mise en correspondance des réseaux sémantiques des différents éléments présents dans le champ visuel va permettre de restreindre les ambiguïtés et d'améliorer la catégorisation sémantique. L'effet de cohérence sémantique n'est pas mis en évidence chez l'enfant d'âge scolaire dans nos études, ce qui suggère une trajectoire dans l'édification du processus de catégorisation sémantique dont le développement semble dépendre de la maturation cérébrale et des expériences de vie.

Les données de la littérature nous permettent de supposer que, dans le fonctionnement mature, une première vague d'informations grossières, portées par les basses fréquences spatiales, est transmise rapidement via les voies magnocellulaires. Ces basses fréquences spatiales sont projetées dans des aires de plus haut niveau cognitif (telles que le COF et le cortex para-hippocampique) afin de permettre l'activation de représentations et la génération d'hypothèses perceptives sur la scène visuelle et ses constituants. Cette première vague se déroule avant que le traitement des basses fréquences spatiales le long de la voie visuelle ventrale (V1, V2, V4, cortex inférotemporal) se termine. L'identification finale de l'image nécessitera la mise en commun dans le cortex inférotemporal des informations des basses fréquences spatiales avec les informations plus précises des hautes fréquences spatiales véhiculées via les voies parvocellulaires.

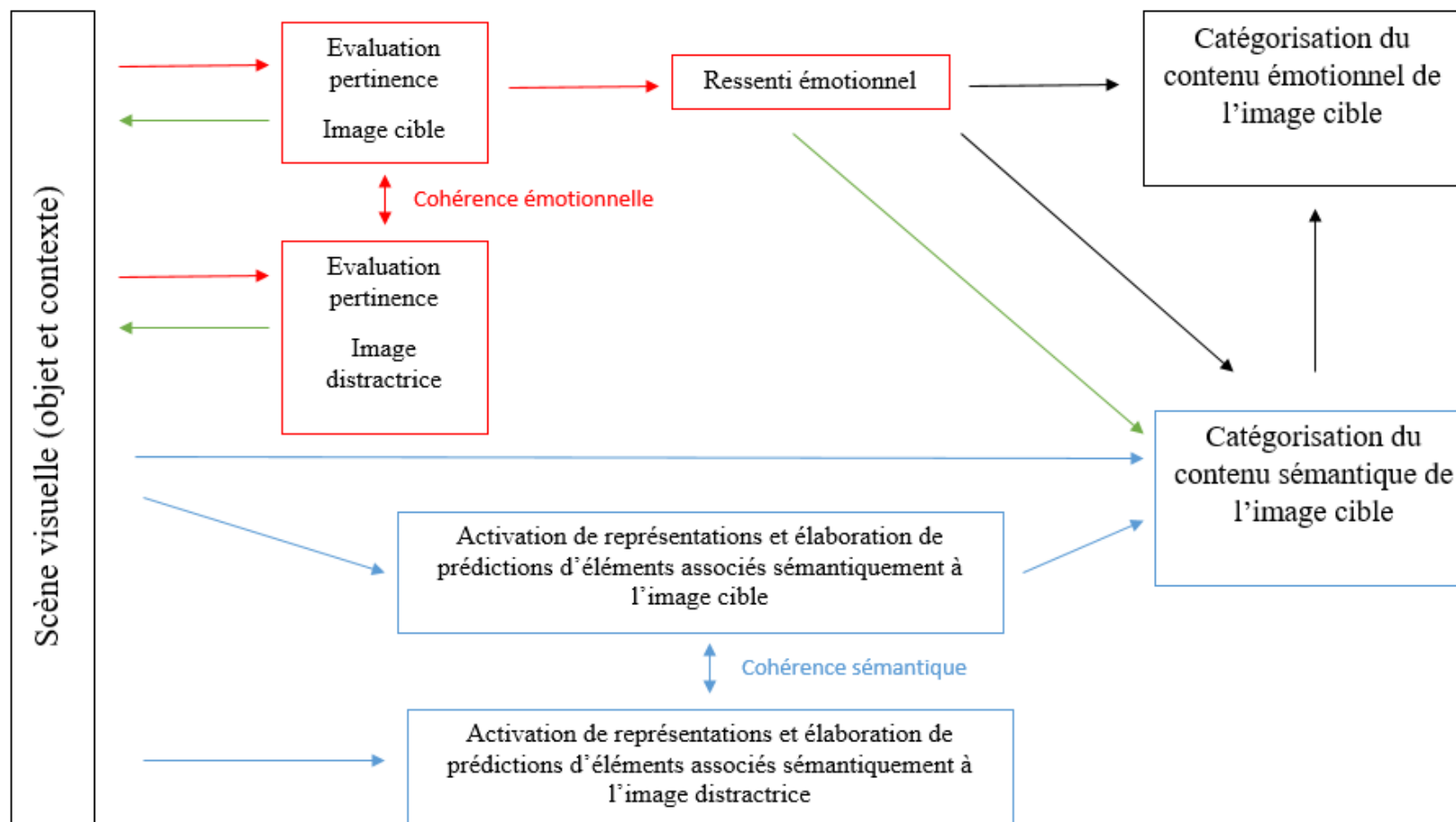


Figure 3 : Représentation schématique des traitements sémantiques et émotionnels lors de la présentation visuelle rapide de deux. L'étape 1 (en rouge) schématise l'évaluation (non volontaire) de la pertinence de l'image, les flèches en vert représentent l'influence du contenu émotionnel sur la catégorisation sémantique. L'étape 2 (en bleu) représente les traitements sémantiques ascendants et descendants, et l'étape 3 (en noir) l'évaluation volontaire du contenu émotionnel.

Le processus de catégorisation sémantique peut ainsi se dérouler directement (en modalité *feedforward*) ou être régulé par des retours (*feedback*) du cortex préfrontal vers les aires visuelles primaires (V1 et V2) afin de guider et restreindre la prise d'informations visuelles sur les éléments les plus pertinents nécessaires pour préciser l'identification de l'image. Chez l'enfant d'âge scolaire, les traitements semblent se dérouler majoritairement via les réseaux ascendants et bénéficient peu d'activation de réseaux sémantiques convergents.

- 3) *Enfin la catégorisation volontaire du contenu émotionnel semble demander plus de ressources attentionnelles ou de prise d'information perceptive que la catégorisation volontaire du contenu sémantique et peut parfois nécessiter, notamment pour les images de contextes négatifs, les plus complexes à évaluer, une identification sémantique. A contrario, la catégorisation des images positives semble facilitée, quel que soit le niveau d'image et ce de façon précoce dans le développement.*

L'étape de catégorisation, donc de jugement volontaire, des émotions peut se réaliser sur la base du ressenti émotionnel, probablement lorsque celui-ci a atteint un seuil permettant une prise en compte consciente de l'expérience émotionnelle (Grandjean, Sander & Scherer, 2008). Lorsque ce niveau n'est pas atteint et que l'expérience émotionnelle est ambiguë, il est probable que l'individu passe par une étape de reconnaissance de l'image afin d'aider son évaluation émotionnelle. Les images de contextes, constituées d'informations diffuses sans contenu social, provoquent peut-être ainsi des expériences émotionnelles moins fortes. Néanmoins ceci n'est pas le cas pour les images positives qui semblent bénéficier d'un traitement facilité car leur catégorisation dépendrait principalement du traitement des basses fréquences spatiales (Kumar & Srinivasan, 2011).

Pour résumer, les travaux réalisés dans le cadre de cette thèse ont permis de préciser certains aspects concernant la façon dont l'individu traite les différents éléments présents dans son champ visuel. Les traitements sémantiques et émotionnels semblent se dérouler indépendamment tout en étant en constante interaction. La mise en relation des contenus sémantiques et émotionnels entre l'objet et le contexte d'une scène visuelle se réalise à des stades différents du processus de traitement des informations. L'effet de cohérence émotionnelle semble se situer à une étape précoce du traitement visuel, probablement sous-corticale, alors que l'effet de cohérence sémantique se situe à une étape plus tardive dépendant de structures corticales de haut niveau.

Certaines limites à ces travaux sont néanmoins à prendre en considération. Les images à contenu émotionnel utilisées dans les trois études qui ont porté sur l'évaluation des traitements émotionnels (études 2, 3 et 4) ont été sélectionnées afin de ne pas être choquantes pour des enfants. En conséquence, les images choisies présentent un niveau d'éveil (*arousal*) faible ou modéré. Or, si l'on se réfère à certains modèles (notamment *l'arousal-biased competition model*, Mather & Sutherland, 2011), la détection de l'émotion par l'amygdale se ferait en fonction du niveau d'éveil du stimulus et indépendamment de sa valence. Notre utilisation de stimuli à niveau d'éveil faible, ou modéré, a pu réduire l'impact du contenu émotionnel. En outre, bien que nos images représentent des objets et des environnements du quotidien, elles n'en restent pas moins des photographies statiques présentées sur écran. Il est évident que leur impact émotionnel ne peut pas correspondre à celui de la rencontre avec des objets ou environnements similaires hors laboratoire. Il est fort probable, notamment, que nos tâches expérimentales minimisent les effets de capture attentionnelle.

3. Application en clinique

L'utilisation d'un paradigme de perception visuelle rapide de scènes complexes permet d'étudier la prise d'information initiale relative à des éléments de niveau local et de niveau global qui apparaissent simultanément et séparément dans le champ visuel. Nous pouvons ainsi analyser les traitements sémantiques et émotionnels initiaux afin de les différencier des traitements liés au déploiement des processus attentionnels plus tardifs et d'exploration visuelle. Notre étude auprès des enfants TDA/H suggère ainsi des capacités de traitement des informations émotionnelles préservées lors d'une prise d'information rapide. Les inclusions de cette étude sont toujours en cours, il est donc important d'être prudent quant à l'interprétation de ces résultats qui ne sont pas définitifs. Ces résultats, préliminaires, suggèrent que le défaut de traitement des informations émotionnelles, principalement mis en évidence dans la littérature dans des tâches où les stimuli sont présentés un temps assez long pour permettre l'exploration visuelle de la scène, serait probablement secondaire au déficit attentionnel constitutif de ce trouble. Il semble donc que lors de l'apparition du stimulus, les enfants TDA/H ont accès à, et traitent normalement le, contenu émotionnel extrait rapidement. En revanche, nous faisons l'hypothèse que dans un second temps ils n'exploreraient pas l'image comme le font les enfants au développement typique, ce qui ne leur permettrait pas d'affiner la compréhension de l'émotion.

L'effet de cohérence émotionnelle est également mis en évidence de façon semblable pour les enfants avec ou sans TDA/H, ce qui implique que les enfants TDA/H ont la capacité de traiter simultanément deux sources d'informations émotionnelles et de les associer même si elles sont présentées séparément. Ce résultat renforce l'hypothèse d'un traitement initial du contenu émotionnel préservé dans le TDA/H. Néanmoins, il est à envisager que des profils comportementaux semblables, entre enfants TDA/H et au développement typique, pour les traitements émotionnels, soient sous-tendus par des activations cérébrales différentes, comme nous l'avons mis en évidence à partir de certaines études décrites dans notre article de synthèse (article III, chapitre III). Des recherches futures devront étudier ainsi les réseaux cérébraux activés lors de la perception visuelle rapide dans le TDA/H.

Enfin, dans notre étude, les enfants TDA/H présentent des capacités de catégorisation sémantique inférieures à celles des enfants au développement typique, ce qui peut évoquer de plus faibles capacités d'accès aux représentations (déficit d'encodage et/ou de récupération). Nos résultats rejoignent ceux d'une étude en PE (Tye et al., 2013) qui a comparé le traitement de visages émotionnels dans deux groupes d'enfants, avec TDA/H ou avec trouble du spectre autistique (TSA). Cette étude a montré que les traitements initiaux (N170) du contenu émotionnel étaient préservés chez les enfants TDA/H, alors que les traitements plus tardifs (N400), liés à la compréhension sémantique de la scène, étaient perturbés (le profil inverse a été obtenu chez les enfants avec TSA). Néanmoins, ces résultats diffèrent de ceux des études présentées dans notre article de synthèse (article III, chapitre III), qui suggèrent un traitement sémantique des scènes visuelles comparable entre enfants TDA/H et enfants au développement typique. Ces résultats qui semblent contradictoires peuvent s'expliquer en partie par les méthodologies et les images utilisées. Pour investiguer plus précisément la présence de particularités de traitement du contenu sémantique dans le TDA/H il sera nécessaire de spécifier la nature des images utilisées (i.e., représentant des objets naturels ou manufacturés, manipulables ou non) ainsi que les niveaux de traitements demandés (i.e., traitements thématiques ou taxonomiques). Il a en effet été montré que ces traitements dépendent de réseaux cérébraux différents et sont sous-tendus par des trajectoires développementales distinctes (Collette, Bonnotte, Jacquemont, Kalénine & Bartolo, 2016 ; Kalénine & Bonthoux, 2008).

Conclusion

L'être humain doit pouvoir traiter efficacement, donc le plus rapidement possible et à moindre coût, son environnement visuel pour s'y adapter au mieux. Pour cela il doit pouvoir

détecter et sélectionner, parmi un flux d'information continu, les éléments les plus pertinents pour satisfaire à ses besoins et pour sa survie. Dans cette thèse, nous nous sommes intéressés à la façon dont les contenus sémantiques et émotionnels présents dans des scènes visuelles sont traités afin de permettre cette rapidité de sélection des informations pertinentes et de compréhension de l'environnement physique et social. L'ensemble des travaux suggère que l'efficacité de la cognition visuelle est permise par des traitements indépendants des contenus sémantiques et émotionnels qui interagissent en fonction des buts de l'observateur. Nous pouvons également faire l'hypothèse que les contraintes environnementales et les expériences de vie de l'individu influencent ces interactions. Enfin, l'étude des traitements sémantiques et émotionnels lors de la présentation rapide de scènes visuelles a montré son intérêt pour améliorer la compréhension des déficits de traitement des contenus émotionnels présents dans le trouble déficitaire de l'attention avec ou sans hyperactivité. Des recherches futures pourront poursuivre ces travaux auprès d'autres populations, par exemple dans le cadre du trouble du spectre autistique, afin de questionner la spécificité des particularités de traitement des émotions, et ainsi discuter leur pertinence pour des diagnostics différentiels. Au total, l'ensemble de ces travaux nous permet de décrire un modèle dynamique de la perception visuelle dans lequel les traitements sémantiques et émotionnels suivent des trajectoires développementales asynchrones et dont la priorisation d'une dimension de traitement sur une autre n'est pas préétablie mais varie en fonction des contraintes liées à l'environnement et à l'observateur.

RÉFÉRENCES

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, 12(2), 169–177. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12015233>
- Aldunate, N., López, V., & Bosman, C. A. (2018). Early Influence of Affective Context on Emotion Perception: EPN or Early-N400? *Frontiers in Neuroscience*, 12, 708. <https://doi.org/10.3389/fnins.2018.00708>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5)* (5th ed.). American Psychiatric Association.
- Andreou, P., Neale, B. M., Chen, W., Christiansen, H., Gabriels, I., Heise, A., ... Kuntsi, J. (2007). Reaction time performance in ADHD: improvement under fast-incentive condition and familial effects. *Psychological Medicine*, 37(12), 1703–1715. <https://doi.org/10.1017/S0033291707000815>
- Antes, J.R. (1977). Recognizing and localizing features in brief picture presentations. *Memory and Cognition*, 5, 155-161.
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5(8), 617–629. <https://doi.org/10.1038/nrn1476>
- Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280–289. <https://doi.org/10.1016/J.TICS.2007.05.005>
- Bar & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron*, 38 (2), 347-358.
- Barkley, R. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121, 65-94.
- Barrett, L. F., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1521), 1325–1334.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Rcpp, L. (2015). Package ‘lme4’. *Convergence*, 12(1).
- Batty, M., & Taylor, M. J. (2002). Visual categorization during childhood: An ERP study. *Psychophysiology*, 39(4), S0048577202010764. <https://doi.org/10.1017/S0048577202010764>
- Bayet, L., Pascalis, O., & Gentaz, É. (2014). Le développement de la discrimination des expressions faciales émotionnelles chez les nourrissons dans la première année. *L'Année Psychologique*, 114(03), 469–500. <https://doi.org/10.4074/S0003503314003030>
- Beffara, B., Wicker, B., Vermeulen, N., Ouellet, M., Bret, A., Molina, M. J. F., & Mermillod, M. (2015). Reduction of interference effect by low spatial frequency information priming in an emotional Stroop task. *Journal of Vision*, 15(6), 16. <https://doi.org/10.1167/15.6.16>
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 14(2), 143–177. [https://doi.org/10.1016/0010-0285\(82\)90007-X](https://doi.org/10.1016/0010-0285(82)90007-X)
- Bornstein, M. H., Mash, C., & Arterberry, M. E. (2011). Perception of object-context relations: eye-movement analyses in infants and adults. *Developmental Psychology*. <https://doi.org/10.1037/a0021059>

- Bourne, V. (2006). *The divided visual field paradigm: Methodological considerations*. (Psychology).
- Boyce, S.J., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *J Exp Psychol Hum Percept Perform*, *15* (3), 556-566.
- Bringmann, A., Syrbe, S., Görner, K., Kacza, J., Francke, M., Wiedemann, P., & Reichenbach, A. (2018). The primate fovea: Structure, function and development. *Progress in Retinal and Eye Research*, *66*, 49–84. <https://doi.org/10.1016/J.PRETEYERES.2018.03.006>
- Brosch, T., Scherer, K. R., Grandjean, D. M., & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss medical weekly*, *143*, w13786.
- Carniglia, E., Caputi, M., Manfredi, V., Zambarbieri, D., & Pessa, E. (2006). The influence of emotional picture thematic content on exploratory eye movements. *Journal of Eye Movement Research*, *5*(4), 1–9.
- Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(4), 1228–1258.
- Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, *54*(1–3), 241–257. [https://doi.org/10.1016/S0301-0511\(00\)00058-2](https://doi.org/10.1016/S0301-0511(00)00058-2)
- Chaumon, M., Kveraga, K., Barrett, L. F., & Bar, M. (2014). Visual Predictions in the Orbitofrontal Cortex Rely on Associative Content. *Cerebral Cortex*, *24*(11), 2899–2907. <https://doi.org/10.1093/cercor/bht146>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. <https://doi.org/10.1016/j.neuron.2008.04.017>
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. <https://doi.org/10.1038/nrn755>
- Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and Explicit Evaluation: fMRI Correlates of Valence, Emotional Intensity, and Control in the Processing of Attitudes. *Journal of Cognitive Neuroscience*, *16*(10), 1717–1729. <https://doi.org/10.1162/0898929042947919>
- Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*(8), 559–564. <https://doi.org/10.1111/j.0956-7976.2004.00719.x>
- de Borst, A. W., & de Gelder, B. (2016). Clear signals or mixed messages: inter-individual emotion congruency modulates brain activity underlying affective body perception. *Social Cognitive and Affective Neuroscience*, *11*(8), 1299–1309. <https://doi.org/10.1093/scan/nsw039>
- De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychol Res*, *52* (4), 317-329.
- Dimond, S., & Beaumont, G. (1972). Processing in perceptual integration between and within the cerebral hemispheres. *British Journal of Psychology*, *63*(4), 509–514. <https://doi.org/10.1111/j.2044-8295.1972.tb01300.x>

- Duh, S., & Wang, S. (2014). Infants detect changes in everyday scenes: The role of scene gist. *Cognitive Psychology*, 72, 142–161. <https://doi.org/10.1016/J.COGLPSYCH.2014.03.001>
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Fredrickson, B. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, 56(3), 218–226. doi:10.1098/rstb.2004.1512
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132(5), 692–731. <https://doi.org/10.1037/0033-2909.132.5.692>
- Gendron, M., Crivelli, C., & Barrett, L. F. (2018). Universality Reconsidered: Diversity in Making Meaning of Facial Expressions. *Current Directions in Psychological Science*, 27(4), 211–219. <https://doi.org/10.1177/0963721417746794>
- Grandjean, D., Sander, D., & Scherer, K. R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Consciousness and cognition*, 17(2), 484–495.
- Hale, T., Loo, S., Zaidel, E., Hanada, G., Maciona, J., & Smalley, S. (2009). Rethinking a Right Hemisphere Deficit in ADHD. *Journal of Attention Disorders*, 13(1), 3–17.
- Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, 113(8), 1339–1350.
- Harnad, S. (1987). *Categorical perception: The groundwork of cognition* (Cambridge:).
- Harnad, S. (2017). To cognize is to categorize: Cognition is categorization. In *Handbook of categorization in cognitive science* (Elsevier, pp. 21–54).
- Haute Autorité de Santé. Conduite à tenir en médecine de premier recours devant un enfant ou un adolescent susceptible d’avoir un trouble déficit de l’attention avec ou sans hyperactivité. Saint-Denis La Plaine: HAS; 2014.
- Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, 50(1), 243–271.
- Hietanen, J. K., & Astikainen, P. (2013). N170 response to facial expressions is modulated by the affective congruency between the emotional expression and preceding affective picture. *Biological Psychology*, 92(2), 114–124.
- Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology General*, 127, 398–415. <https://doi.org/10.1037/0096-3445.127.4.398>
- Ihssen, N., & Keil, A. (2013). Accelerative and decelerative effects of hedonic valence and emotional arousal during visual scene processing. *The Quarterly Journal of Experimental Psychology*, 66(7), 1276–1301.
- Joubert, O. R., Fize, D., Rousselet, G. a, & Fabre-Thorpe, M. (2008). Early interference of context congruence on object processing in rapid visual categorization of natural scenes. *Journal of Vision*, 8(13), 11.1–18. <https://doi.org/10.1167/8.13.11>

- Joubert, O. R., Rousselet, G. a., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, *47*(26), 3286–3297. <https://doi.org/10.1016/j.visres.2007.09.013>
- Kalénine, S., & Bonthoux, F. (2006). The formation of living and non-living superordinate concepts as a function of individual differences. *Current psychology letters. Behaviour, brain & cognition*, (19, Vol. 2, 2006).
- Kalénine, S., & Bonthoux, F. (2008). Object manipulability affects children's and adults' conceptual processing. *Psychonomic bulletin & review*, *15*(3), 667-672.
- Kalénine, S. (2009). Le rôle de l'action dans l'accès aux concepts d'objets: Apport de la neuropsychologie et des neurosciences cognitives. *Revue de neuropsychologie*, *1*(2), 150-158.
- Kalénine, S., Peyrin, C., Pichat, C., Segebarth, C., Bonthoux, F., & Baciú, M. (2009). The sensory-motor specificity of taxonomic and thematic conceptual relations: A behavioral and fMRI study. *Neuroimage*, *44*(3), 1152-1162.
- Karatekin, C., & Asarnow, R. F. (1999). Exploratory eye movements to pictures in childhood-onset schizophrenia and attention-deficit/hyperactivity disorder (ADHD). *Journal of Abnormal Child Psychology*, *27*(1), 35–49. <https://doi.org/10.1023/A:1022662323823>
- Kauffmann, L., Ramanoël, S., & Peyrin, C. (2014). The neural bases of spatial frequency processing during scene perception. *Frontiers in Integrative Neurosciences*, *8*(37).
- Kofler, M. J., Rapport, M. D., Sarver, D. E., Raiker, J. S., Orban, S. A., Friedman, L. M., & Kolomeyer, E. G. (2013). Reaction time variability in ADHD: A meta-analytic review of 319 studies. *Clinical Psychology Review*, *33*(6), 795–811. <https://doi.org/10.1016/j.cpr.2013.06.001>
- Kret, M. E., & de Gelder, B. (2010). Social context influences recognition of bodily expressions. *Experimental Brain Research*, *203*(1), 169–180. <https://doi.org/10.1007/s00221-010-2220-8>
- Kumar, D., & Srinivasan, N. (2011). Emotion perception is mediated by spatial frequency content. *Emotion*, *11*(5), 1144.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General*, *144*(2), 339–365. <https://doi.org/10.1037/xge0000040>
- Lai, V. T., Hagoort, P., & Casasanto, D. (2012). Affective Primacy vs. Cognitive Primacy: Dissolving the Debate. *Frontiers in Psychology*, *3*, 243. <https://doi.org/10.3389/fpsyg.2012.00243>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): affective ratings of pictures and instruction manual. Tech Rep A-8*. University of Florida, Gainesville.
- Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist*, *39*(2), 124–129. <https://doi.org/10.1037/0003-066X.39.2.124>
- LeDoux, J. (2003). The Emotional Brain, Fear, and the Amygdala. *Cellular and Molecular Neurobiology*, *23*(4/5), 727–738. <https://doi.org/10.1023/A:1025048802629>

- Leppänen, J. M., & Hietanen, J. K. (2004). Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychological Research Psychologische Forschung*, *69*(1–2), 22–29. <https://doi.org/10.1007/s00426-003-0157-2>
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *The Behavioral and Brain Sciences*, *35*(3), 121–143. <https://doi.org/10.1017/S0140525X11000446>
- LoBue, V., Rakison, D. H., & DeLoache, J. S. (2010). Threat Perception Across the Life Span. *Current Directions in Psychological Science*, *19*(6), 375–379. <https://doi.org/10.1177/0963721410388801>
- March, D. S., Gaertner, L., & Olson, M. A. (2018). On the Prioritized Processing of Threat in a Dual Implicit Process Model of Evaluation. *Psychological Inquiry*, *29*(1), 1–13. <https://doi.org/10.1080/1047840X.2018.1435680>
- Marchewka, A., Zurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*(2), 596–610. <https://doi.org/10.3758/s13428-013-0379-1>
- Martel, M. M. (2009). Research Review: A new perspective on attention-deficit/hyperactivity disorder: emotion dysregulation and trait models. *Journal of Child Psychology and Psychiatry*, *50*(9), 1042–1051. <https://doi.org/10.1111/j.1469-7610.2009.02105.x>
- Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on psychological science*, *6*(2), 114–133.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. <https://doi.org/10.3758/s13428-011-0168-7>
- Mazzietti, A., & Koenig, O. (2014). The relevance bias: Valence-specific, relevance-modulated performance in a two-choice detection task. *Cognition & emotion*, *28*(1), 143–152.
- Mazzietti, A., Sellem, V., & Koenig, O. (2014). From stimulus-driven to appraisal-driven attention: Towards differential effects of goal relevance and goal relatedness on attention?. *Cognition and Emotion*, *28*(8), 1483–1492.
- McManis, M., Bradley, M., Berg, W., Cuthbert, B., & Lang, P. (2001). Emotional reactions in children: Verbal, physiological, and behavioral responses to affective pictures. *Psychophysiology*, *38*(2), 222–231.
- Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, *102*(45), 16518–16523. <https://doi.org/10.1073/pnas.0507650102>
- Mermillod, M., Bourrier, Y., David, E., Kauffmann, L., Chauvin, A., Guyader, N., ... Peyrin, C. (2019). The importance of recurrent top-down synaptic connections for the anticipation of dynamic emotions. *Neural Networks*, *109*, 19–30. <https://doi.org/10.1016/J.NEUNET.2018.09.007>
- Mermillod, M., Droit-Volet, S., Devaux, D., Schaefer, A., & Vermeulen, N. (2010). Are Coarse Scales Sufficient for Fast Detection of Visual Threat? *Psychological Science*, *21*(10), 1429–1437. <https://doi.org/10.1177/0956797610381503>

- Mermillod, M., Kauffmann, L., Roux-Sibilon, A., Bret, A., Palluel-Germain, R., & Peyrin, C. (2017). Are low spatial frequencies or high contrasts the trigger of threat detection? Retrieved from <http://hal.univ-grenoble-alpes.fr/hal-01915872>
- Mondloch, C. J. (2012). Sad or fearful? The influence of body posture on adults' and children's perception of facial displays of emotion. *Journal of Experimental Child Psychology*, *111*(2), 180–196. <https://doi.org/10.1016/J.JECP.2011.08.003>
- Mondloch, C. J., Horner, M., & Mian, J. (2013). Wide eyes and drooping arms: Adult-like congruency effects emerge early in the development of sensitivity to emotional faces and body postures. *Journal of Experimental Child Psychology*, *114*(2), 203–216. <https://doi.org/10.1016/j.jecp.2012.06.003>
- Mueller, C. J., Fritsch, N., Hofmann, M. J., & Kuchinke, L. (2017). Differences in the dynamics of affective and cognitive processing – An ERP study. *Brain Research*, *1655*, 41–47. <https://doi.org/10.1016/J.BRAINRES.2016.11.018>
- Munneke, J., Brentari, V., & Peelen, M. V. (2013). The influence of scene context on object recognition is independent of attentional focus. *Frontiers in Psychology*, *4*(AUG), 1–10. <https://doi.org/10.3389/fpsyg.2013.00552>
- Musch, J., & Klauer, K. C. (2003). *The psychology of evaluation : affective processes in cognition and emotion*. Lawrence Erlbaum Associates.
- Muszkat, M., de Mello, C., Muñoz, P., Lucci, T., David, V., Siqueira, J., & Otta, E. (2015). Face scanning in autism spectrum disorder and attention deficit/hyperactivity disorder: Human versus dog face scanning. *Frontiers in Psychiatry*, *6*(150).
- Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Visual Cognition*, *25*(4–6), 563–574. <https://doi.org/10.1080/13506285.2017.1301615>
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2010). Semantic categorization precedes affective evaluation of visual scenes. *Journal of Experimental Psychology: General*, *139*(2), 222–246. <https://doi.org/10.1037/a0018858>
- Ofen, N., & Shing, Y. L. (2013). From perception to memory: Changes in memory systems across the lifespan. *Neuroscience & Biobehavioral Reviews*, *37*(9), 2258–2267. <https://doi.org/10.1016/J.NEUBIOREV.2013.04.006>
- Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, *77*(3), 247–265. <https://doi.org/10.1016/j.biopsycho.2007.11.006>
- Palmer, T. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519–526. <https://doi.org/10.3758/BF03197524>
- Pelc, K., Kornreich, C., Foisy, M.-L., & Dan, B. (2006). Recognition of emotional facial expressions in attention-deficit hyperactivity disorder. *Pediatric Neurology*, *35*, 93–97.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*(2), 148–158. <https://doi.org/10.1038/nrn2317>
- Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, *13*(4), 160–166. <https://doi.org/10.1016/j.tics.2009.01.006>
- Pessoa, L. (2010). Emotion and cognition and the amygdala: From “what is it?” to “what’s to be done?” *Neuropsychologia*, *48*(12), 3416–3429. <https://doi.org/10.1016/j.neuropsychologia.2010.06.038>

- Pessoa, L. & Adolphs, R. (2010). Emotion processing and the amygdala: from a “low road” to “many roads” of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773–782. <https://doi.org/10.1038/nrn2920>
- Pourtois, G., Schettino, A., & Vuilleumier, P. (2012). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology*, *92*(3), 492–512.
- Rémy, F., Vayssière, N., Pins, D., Boucart, M., & Fabre-Thorpe, M. (2014). Incongruent object/context relationships in visual scenes: Where are they processed in the brain? *Brain and Cognition*, *84*(1), 34–43. <https://doi.org/10.1016/J.BANDC.2013.10.008>
- Righart, R., & de Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(3), 264–272. <https://doi.org/10.3758/CABN.8.3.264>
- Righart, R. & de Gelder, B. (2006). Context Influences Early Perceptual Analysis of Faces—An Electrophysiological Study. *Cerebral Cortex*, *16*(9), 1249–1257. <https://doi.org/10.1093/cercor/bhj066>
- Rosen, M. L., Sheridan, M. A., Sambrook, K. A., Peverill, M. R., Meltzoff, A. N., & McLaughlin, K. A. (2018). The Role of Visual Association Cortex in Associative Memory Formation across Development. *Journal of Cognitive Neuroscience*, *30*(3), 365–380. https://doi.org/10.1162/jocn_a_01202
- Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get to the gist of real-world natural scenes. *Visual Cognition*, *12*(6), 852–877. <https://doi.org/10.1080/13506280444000553>
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. <https://doi.org/10.1037/0033-295X.110.1.145>
- Sander, D., & Koenig, O. (2002). No inferiority complex in the study of emotion complexity: A cognitive neuroscience computational architecture of emotion. *Cognitive Science Quarterly*, *2*(3–4), 249–272.
- Sander, D., Grandjean, D., & Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Networks*, *18*(4), 317–352. <https://doi.org/10.1016/J.NEUNET.2005.03.001>
- Sander, D., Grandjean, D., & Scherer, K. R. (2018). An appraisal-driven componential approach to the emotional brain. *Emotion Review*, *10*(3), 219–231.
- Schel, M. A., & Crone, E. A. (2013). Development of response inhibition in the context of relevant versus irrelevant emotions. *Frontiers in Psychology*, *4*, 383. <https://doi.org/10.3389/fpsyg.2013.00383>
- Scherer, K. (2001). *Appraisal considered as a process of multilevel sequential checking*. In K. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 99–120). New York, NY: Oxford University Press.
- Scheuerecker, J., Frodl, T., Koutsouleris, N., Zetsche, T., Wiesmann, M., Kleemann, A. M., ... Meisenzahl, E. M. (2007). Cerebral Differences in Explicit and Implicit Emotional Processing – An fMRI Study. *Neuropsychobiology*, *56*(1), 32–39. <https://doi.org/10.1159/000110726>
- Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal Structure Predicts Statistical Learning and Associative Inference

- Abilities during Development. *Journal of Cognitive Neuroscience*, 29(1), 37–51. https://doi.org/10.1162/jocn_a_01028
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, 84(1), 1.
- Serrano, V. J., Owens, J. S., & Hallowell, B. (2015). Where Children With ADHD Direct Visual Attention During Emotion Knowledge Tasks: Relationships to Accuracy, Response Time, and ADHD Symptoms. *Journal of Attention Disorders*, 22(8), 752–763. <https://doi.org/10.1177/1087054715593632>
- Shaw, P., Stringaris, A., Nigg, J., & Leibenluft, E. (2014). Emotion Dysregulation in Attention Deficit Hyperactivity Disorder. *American Journal of Psychiatry*, 171(3), 276–293. <https://doi.org/10.1176/appi.ajp.2013.13070966>
- Shin, D. W., Lee, S. J., Kim, B. J., Park, Y., & Lim, S. W. (2008). Visual attention deficits contribute to impaired facial emotion recognition in boys with attention-deficit/hyperactivity disorder. *Neuropediatrics*, 39(6), 323–327. <https://doi.org/10.1055/s-0029-1202286>
- Skodzik, T., Holling, H., & Pedersen, A. (2017). Long-Term Memory Performance in Adult ADHD. *Journal of Attention Disorders*, 21(4), 267–283. <https://doi.org/10.1177/1087054713510561>
- Sloutsky, V. M., & Fisher, A. V. (2011). The development of categorization. In *Psychology of learning and motivation* (Academic P, pp. 141–166).
- Sonuga-Barke, E. J. S. (2002). Psychological heterogeneity in AD/HD - A dual pathway model of behaviour and cognition. *Behavioural Brain Research*, 130, 29–36. [https://doi.org/10.1016/S0166-4328\(01\)00432-6](https://doi.org/10.1016/S0166-4328(01)00432-6)
- Spruyt, A., Hermans, D., De Houwer, J., Vandromme, H., & Eelen, P. (2007). On the nature of the affective priming effect: effects of stimulus onset asynchrony and congruency proportion in naming and evaluative categorization. *Memory & Cognition*, 35(1), 95–106. <https://doi.org/10.3758/BF03195946>
- Stock, J. Van den, & Gelder, B. de. (2014). Face identity matching is influenced by emotions conveyed by face and body. *Frontiers in Human Neuroscience*, 8, 53. <https://doi.org/10.3389/fnhum.2014.00053>
- Storbeck, J., Robinson, M. D., & McCourt, M. E. (2006). Semantic Processing Precedes Affect Retrieval: The Neurological Case for Cognitive Primacy in Visual Processing. *Review of General Psychology*, 10(1), 41–55. <https://doi.org/10.1037/1089-2680.10.1.41>
- Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, 15(11), 745–756. <https://doi.org/10.1038/nrn3838>
- Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, 11(10), 697–709. <https://doi.org/10.1038/nrn2889>
- Theurel, A., Witt, A., Malsert, J., Lejeune, F., Fiorentini, C., Barisnikov, K., & Gentaz, E. (2016). The integration of visual context information in facial emotion recognition in 5- to 15-year-olds. *Journal of Experimental Child Psychology*, 150, 252–271. <https://doi.org/10.1016/J.JECP.2016.06.004>

- Thomas, R. L., Nardini, M., & Mareschal, D. (2017). The impact of semantically congruent and incongruent visual information on auditory object recognition across development. *Journal of Experimental Child Psychology*, *162*, 72–88. <https://doi.org/10.1016/J.JECP.2017.04.020>
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.
- Trapp, S., & Bar, M. (2015). Prediction, context, and competition in visual recognition. *Annals of the New York Academy of Sciences*, *1339*(1), 190–198.
- Türkan, B. N., Amado, S., Ercan, E. S., & Perçinel, I. (2016). Comparison of change detection performance and visual search patterns among children with/without ADHD: Evidence from eye movements. *Research in Developmental Disabilities*, *49–50*, 205–215. <https://doi.org/10.1016/j.ridd.2015.12.002>
- Tye, C., Mercure, E., Ashwood, K. L., Azadi, B., Asherson, P., Johnson, M. H., ... McLoughlin, G. (2013). Neurophysiological responses to faces and gaze direction differentiate children with ASD, ADHD and ASD + ADHD. *Developmental Cognitive Neuroscience*, *5*, 71–85. <https://doi.org/10.1016/j.dcn.2013.01.001>
- Ullman, S. *High-Level Vision: Object Recognition and Visual Cognition*. (1996).
- van Stralen, J. (2016). Emotional dysregulation in children with attention-deficit/hyperactivity disorder. *ADHD Attention Deficit and Hyperactivity Disorders*, *8*(4), 175–187. <https://doi.org/10.1007/s12402-016-0199-0>
- Vanmarcke, S., & Wagemans, J. (2015). The influence of age and gender on ultra-rapid categorization. *Visual Cognition*, *23*(7), 894–916. <https://doi.org/10.1080/13506285.2015.1091801>
- VanRullen, R., & Thorpe, S. J. (2001). Is it a Bird? Is it a Plane? Ultra-Rapid Visual Categorisation of Natural and Artifactual Objects. *Perception*, *30*(6), 655–668. <https://doi.org/10.1068/p3029>
- Vesker, M., Bahn, D., Degé, F., Kauschke, C., & Schwarzer, G. (2018). Developmental changes in the categorical processing of positive and negative facial expressions. *PLOS ONE*, *13*(8), e0201521. <https://doi.org/10.1371/journal.pone.0201521>
- Villemonteix, T., Marx, I., Septier, M., Berger, C., Hacker, T., Bahadori, S., ... Massat, I. (2017). Attentional control of emotional interference in children with ADHD and typically developing children: An emotional N-back study. *Psychiatry Research*, *254*(July 2016), 1–7. <https://doi.org/10.1016/j.psychres.2017.04.027>
- Wickham, H. (2009). Elegant graphics for data analysis (ggplot2).
- Widen, S. C. (2013). Children's Interpretation of Facial Expressions: The Long Path from Valence-Based to Specific Discrete Categories. *Emotion Review*, *5*(1), 72–77. <https://doi.org/10.1177/1754073912451492>
- Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development*, *23*(2), 291–312. <https://doi.org/10.1016/J.COGDEV.2008.01.002>
- Xiao, J., Hays, J., Ehinger, K., Oliva, A., & Torralba, A. (2010). SUN Database: Large-scale Scene Recognition from Abbey to Zoo. In *IEEE Conference on Computer Vision and Pattern Recognition*.
- Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist*, *39*(2), 117–123. <https://doi.org/10.1037/0003-066X.39.2.117>

ANNEXES

Liste des annexes

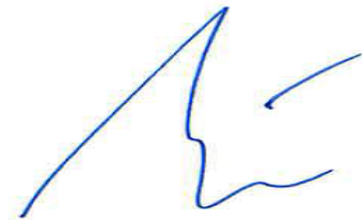
- Annexe 1 : Avis du Comité d’Ethique pour les Recherches Non Interventionnelles de l’Université Côte d’Azur
- Annexe 2 : Notice d’information sur le protocole expérimental à l’attention des parents (version contrôle)
- Annexe 3 : Formulaire d’inclusion (version contrôle)
- Annexe 4 : Notice d’information détaillée sur le protocole expérimental à l’intention des parents (version contrôle)
- Annexe 5 : Formulaire de consentement éclairé à l’attention des parents (version contrôle)
- Annexe 6 : Questionnaire « points forts, points faibles » (SDQ version parents)
- Annexe 7 : Notice d’information sur le protocole expérimental et consentement à l’intention des enfants (version contrôle)
- Annexe 8 : Avis du Comité de Protection des Personnes Sud Méditerranée I
- Annexe 9 : Enregistrement clinical trials
- Annexe 10 : Notice d’information détaillée sur le protocole expérimental et formulaire de consentement à l’intention des parents (version patient)
- Annexe 11 : Formulaire d’inclusion (version patient)
- Annexe 12 : Notice d’information sur le protocole expérimental et consentement à l’intention des enfants (version patient)
- Annexe 13 : Poster présenté à la Société Française de Psychologie
- Annexe 14 : Formulaire de consentement libre et éclairé des études 1 et 2 chez l’adulte
- Annexe 15 : Questionnaire latéralité et renseignements adulte

Annexe 1 : Avis du Comité d’Ethique pour les Recherches Non Interventionnelles de l’Université Côte d’Azur

CERNI-Dossier 2018-2

Soumis par S. FAURE

A la suite des réponses apportées par Sylvane FAURE le 17 décembre 2018 aux questions et suggestions des deux rapporteurs, le CERNI d’UCA donne un avis de conformité aux règles éthiques du projet de recherche « TDA/H » soumis à son appréciation.



Michel Rainelli
Professeur émérite UCA
Réfèrent éthique et intégrité scientifique
Président du CERNI

Etude sur le traitement des informations visuelles dans le développement normal et chez l'enfant présentant un trouble déficitaire de l'attention

Madame, Monsieur,

Nous proposons à votre enfant de participer à une **étude analysant comment les enfants perçoivent, traitent et mémorisent des images**. Nous analysons ces capacités chez les enfants qui ne présentent pas de problèmes particuliers dans leur scolarité, vie quotidienne et sociale, raison pour laquelle **nous sollicitons votre accord et la participation de votre enfant**. Les données obtenues permettront (1) de renforcer les méthodes d'apprentissages et (2) seront comparées aux données d'enfants présentant un Trouble Déficitaire de l'Attention avec ou sans Hyperactivité afin d'améliorer le diagnostic et la prise en charge de ce trouble.

Cette étude est réalisée par le laboratoire d'Anthropologie et de Psychologie Cliniques, Cognitives et Sociales de l'Université Nice Côte d'Azur et par le Centre de Référence des Troubles des Apprentissages des Hôpitaux Pédiatriques de Nice CHU-Lenval. Elle est proposée à des enfants de **7 à 12 ans, droitiers**, et elle se réalise en **deux séances d'environ 30 mn**, espacées d'une semaine. Les tâches consistent à juger des images présentées rapidement sur un écran d'ordinateur. Un questionnaire sur les points forts et les points faibles de votre enfant vous sera proposé ainsi qu'à son enseignant afin d'apporter des éléments supplémentaires. Les séances seront proposées au sein de l'établissement scolaire, en accord avec l'enseignant afin de ne pas gêner l'organisation de la classe.

L'ensemble des renseignements personnels et scientifiques collecté sera archivé et restera strictement confidentiel. **L'anonymat des enfants est assuré**. A terme, Un récapitulatif des résultats généraux de l'étude pourra être transmis sur demande.

Si votre enfant et vous-mêmes êtes intéressés par cette étude, nous vous demandons de **remplir le questionnaire d'inclusion** joint et le **rendre à l'enseignant cette semaine**, sous enveloppe cachetée pour garantir la confidentialité. **Cela n'engagera pas la participation de votre enfant**.

Nous vous transmettrons alors un descriptif plus détaillé du protocole, le questionnaire « points forts - points faibles » et les deux exemplaires d'un formulaire de consentement éclairé que vous devrez remplir et signer uniquement si vous acceptez que votre enfant participe à l'étude.

Nous vous remercions de l'intérêt que vous porterez à cette démarche, et vous invitons à **nous contacter** pour tout complément d'informations, et ce, **quelle que soit votre décision**.

Responsables et coordonnateurs associés et contacts :

Pr Sylvane FAURE
Université Côte d'Azur, LAPCOS,
sfaure@unice.fr

Dr Catherine FOSSOUD
Centre de référence des troubles des apprentissages
Hôpitaux Pédiatriques de Nice CHU-Lenval

Annexe 3 : Formulaire d'inclusion (version contrôle)

A compléter par les parents

En cas de doute, ne pas hésiter à demander des informations complémentaires

Nom de la personne remplissant ce questionnaire :

Date :

Lien avec l'enfant :

Numéro de téléphone et/ou adresse mail :

Nom de l'enfant :

Prénom :

Date de naissance :

Age :

Classe actuelle:

Etablissement scolaire :

Latéralité manuelle : Droitier Gaucher Indéterminé Ambidextre

Sexe : Masculin Féminin

Critères d'inclusion :

Age compris entre 7 et 12 ans au moment de l'étude : Oui Non

Attention ! Une réponse « non » à cette question ne permet pas d'inclure l'enfant dans l'étude.

Critères d'exclusion :

Naissance prématurée (avant 35 sem d'aménorrhée révolues) : Oui Non

Antécédents de Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H) diagnostiqué : Oui Non

Antécédents de pathologie neurologique acquise diagnostiquée : Oui Non

Antécédents de pathologie psychiatrique diagnostiquée : Oui Non

Déficience intellectuelle diagnostiquée : Oui Non

Trouble de la compréhension verbale pouvant pénaliser la passation : Oui Non

Trouble sensoriel élémentaire pouvant pénaliser la passation (*) : Oui Non

Absence de consentement des parents ou du tuteur légal : Oui Non

Attention ! Une réponse « oui » à l'une de ces questions ne permet pas d'inclure l'enfant dans l'étude.

S'il existe une contre-indication à l'inclusion de l'enfant dans l'étude, ne pas poursuivre le questionnaire.

(*) Notamment, déficience auditive ou visuelle

Informations complémentaires :

Problème de vision : Oui Non

si oui, quel est le type de correction ?.....

Traitement(s) médicamenteux en cours : Oui Non

si oui, lesquels ?.....

Scolarité : Redoublement Saut de classe

si oui, indiquez la classe concernée :.....

Aide scolaire (AVSI, RASED, ...) ?

si oui, précisez la nature, durée et fréquence de cette aide :.....

Diagnostic de trouble(s) des apprentissages : Oui Non

si oui précisez lequel (lesquels) :

Suivi paramédical : Orthophoniste Psychomotricien(ne) Psychologue/Neuropsychologue
Ergothérapeute Autre

si oui précisez la fréquence des suivis :

Niveau d'étude / Profession du père :.....

Niveau d'étude / Profession de la mère :.....

**NOTICE D'INFORMATION destinée aux Titulaires de l'autorité parentale pour la
participation d'une personne MINEURE**

Madame, Monsieur,

Votre enfant est invité à participer à une recherche. Avant d'accepter qu'il participe à l'étude, il est important que vous preniez le temps de lire, de comprendre et de considérer attentivement les renseignements qui suivent. Le présent document vous renseigne sur les modalités de cette recherche. S'il y a des mots que vous ne comprenez pas, n'hésitez pas à poser des questions. Pour cela vous pouvez contacter la personne responsable de l'étude, le Pr Sylvane Faure (sfaure@unice.fr), et/ou rencontrer à l'école de votre enfant (pendant ou après le temps scolaire) Mme Anaïs Leroy (anaïs.leroy@unice.fr), investigatrice, doctorante en psychologie au LAPCOS et psychomotricienne, qui proposera à votre enfant une tâche de jugement sur des images de la vie quotidienne et une tâche de mémoire. Votre enfant sera consulté, il peut accepter ou refuser de participer et vous pouvez à tout moment changer d'avis et décider de retirer votre enfant de l'étude.

Objectif de l'étude :

L'objectif est d'analyser comment les enfants traitent, c'est-à-dire reconnaissent et mémorisent, des images. Cette étude sera proposée à des enfants présentant un développement sans particularité et à des enfants présentant un Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H). Ce trouble occasionne des difficultés pour rester attentif, concentré, et peut entraîner un besoin constant de bouger. Ces difficultés peuvent se répercuter sur la scolarité et les relations sociales et affectives de ces enfants. Aussi, les enfants présentant un TDA/H pourraient être en difficulté pour traiter certaines informations contenues dans des images. C'est pourquoi nous étudions comment les enfants reconnaissent différentes informations présentes dans les images : des informations sur le sens global de l'image (par exemple, reconnaître une plage, un restaurant), des informations plus précises sur les objets au sein de l'image (comme un palmier, des assiettes, une personne...) et des informations sur les émotions que procurent ces images (émotion positive, négative ou images neutres sans émotion particulière). Pour cela, nous devons analyser le développement des compétences dans la reconnaissance de ces types d'informations chez des enfants présentant un développement sans particularité, raison pour laquelle nous sollicitons votre accord et la participation de votre enfant. Les données obtenues seront ensuite comparées à celles obtenues en milieu hospitalier auprès d'enfants présentant un TDA/H.

Déroulement de l'étude :

Cette étude est réalisée par le Laboratoire d'Anthropologie et de Psychologie Cliniques, Cognitives et Sociales de l'Université Nice Côte d'Azur et le Centre de Référence des Troubles des Apprentissages (CERTA) des hôpitaux pédiatriques Nice Lenval. Elle est proposée à des enfants de 7 à 12 ans et se réalise en deux séances, d'environ 30 mn chacune, espacées d'une semaine, sur lieu scolaire. La première consiste à juger des images : soit selon leur contenu sémantique (choisir entre différentes catégories, par exemple entre « personne », « animal » ou « objet inanimé »), soit selon leur contenu émotionnel (dire si les images sont « positives », « négatives » ou « neutres »). Les images sont présentées très rapidement sur un écran d'ordinateur. Durant la seconde séance, deux images sont présentées à chaque essai à l'enfant, qui doit reconnaître celle provenant de la série d'images qu'il a vue lors de la première séance. Etant donné que nous étudions la mémorisation non volontaire des images, il est nécessaire de ne pas informer l'enfant qu'il y aura une seconde séance où il devra reconnaître les images que nous lui avons présentées lors de la première séance. Enfin, un questionnaire évaluant la gestion des émotions, les relations sociales et le comportement de votre enfant vous sera proposé, ainsi qu'à l'enseignant de votre enfant, afin d'apporter des éléments

supplémentaires. Ce questionnaire intitulé « points forts, points faibles » existe en version parents et en version enseignant : si vous souhaitez consulter au préalable le questionnaire qui sera adressé à l'enseignant de votre enfant, vous pouvez en faire la demande auprès du Pr Sylvane Faure (sfaure@unice.fr), responsable de l'étude.

Quels sont les bénéfices attendus ?

Concernant les enfants qui présentent un développement sans particularité, cette étude va permettre une meilleure connaissance de comment se développent les capacités de traitement des informations visuelles. Nous pourrions ainsi envisager comment utiliser les informations visuelles pour favoriser les apprentissages (compréhension et mémorisation des informations). Les résultats de ces travaux contribueront également à l'acquisition de nouvelles connaissances sur les particularités de traitement des informations visuelles par les enfants présentant un TDA/H afin d'améliorer les outils diagnostiques et de prise en charge.

Quels sont les risques prévisibles ?

Il n'existe aucun inconvénient majeur potentiel pour les participants à cette étude. Les risques potentiels mineurs sont une fatigue visuelle et/ou des maux de tête provenant du temps passé à regarder l'écran d'ordinateur ; pour diminuer ces risques chaque séance dure moins de 45 minutes et inclut trois pauses. Le risque que l'enfant soit choqué par certaines images a été limité au maximum : nous avons sélectionné les images de façon à ce qu'aucune ne soit choquante et nous nous sommes assurées que des parents d'enfants âgés de 7 à 9 ans confirmaient unanimement notre jugement.

Quelles sont les contraintes ?

L'étude ne présente pas de contraintes particulières. Votre enfant sera rencontré sur son lieu scolaire à des créneaux ne gênant pas sa scolarité (avec accord de son enseignant).

Serez-vous informé des résultats globaux et donc des connaissances acquises grâce à cette recherche ?

A l'issue de cette recherche, ses résultats globaux vous seront communiqués par courrier si vous le souhaitez. Merci de communiquer votre souhait et celui de votre enfant au Pr Sylvane Faure (sfaure@unice.fr).

Votre enfant pourra-t-il participer simultanément à une autre recherche ? A la fin de cette étude pourra-t-il immédiatement participer à une autre recherche ?

Oui, la participation à cette étude n'aura aucune répercussion sur les capacités de votre enfant, qui pourra participer à d'autres recherches.

Vous décidez que votre enfant participe à l'étude, que devez-vous faire concrètement ?

Si vous acceptez que votre enfant participe à cette étude, vous devez obligatoirement lire, dater et signer l'autorisation de participation accompagnant cette notice. Vous avez le droit de refuser que votre enfant participe à l'étude ou de retirer votre consentement à tout moment, et sans encourir aucune responsabilité, ni aucun préjudice de ce fait ; vous devrez alors en informer le Pr Sylvane Faure (sfaure@unice.fr).

Confidentialité des données :

Un traitement des données relatives à votre enfant va être mis en œuvre pour analyser les résultats de la recherche. A cette fin, les informations recueillies auprès de vous et de l'enseignant au moyen du questionnaire « points forts-points faibles » et ses réponses aux tâches des séances 1 et 2 seront accessibles aux personnes travaillant sur cette étude (responsable Pr S.Faure ; investigateurs C.Fossoud, CERTA, A.Leroy et S.Spotorno, LAPCOS). Ces données seront identifiées par un code

Annexes

qui inclura des informations relatives à l'étude et le numéro aléatoire d'inclusion de votre enfant. Toutes ces informations seront strictement confidentielles et protégées par le secret professionnel ainsi que par la loi informatique et libertés n°2004-801 du 6/08/2004 et le décret d'application numéro 2005-1309 paru le 20/10/2005. Vous aurez le droit de vous opposer au traitement automatisé des données concernant votre enfant. Les résultats de l'étude pourront être publiés dans des journaux médicaux et scientifiques sans que l'identité de votre enfant soit révélée.

« Processus attentionnels et émotionnels dans le traitement des scènes visuelles complexes chez l'enfant: étude du fonctionnement normal et application au trouble déficitaire de l'attention. »

AUTORISATION DE PARTICIPATION destinée aux Titulaires de l'autorité parentale pour la participation d'une personne MINEURE

Les données de cette étude, incluant celles portant sur le cas de votre enfant, sont couvertes par le secret professionnel.

Je soussigné(e)-----

demeurant à (adresse complète)-----

Père Mère Titulaire de l'exercice de l'autorité parentale

Et

Je soussigné(e)-----

demeurant à (adresse complète)-----

Père Mère Titulaire de l'exercice de l'autorité parentale

De l'enfant : -----

Déclare avoir compris le but et les modalités de cette étude proposée à mon enfant, qui m'ont été pleinement expliquées.

Les informations relatives au principe de l'étude, et son intérêt m'ont bien été communiquées dans la Note d'Information. J'ai eu la possibilité de l'étudier attentivement. Des réponses ont été apportées à toutes mes questions. J'ai disposé d'un délai de réflexion avant de prendre ma décision.

J'accepte que mon enfant participe à l'étude « Processus attentionnels et émotionnels dans le traitement des scènes visuelles complexes chez l'enfant: étude du fonctionnement normal et application au trouble déficitaire de l'attention ». Il m'a bien été précisé que je pouvais refuser que mon enfant participe à cette étude et que dans le cas d'une participation à celle-ci, je pouvais revenir sur ma décision à tout moment. On m'a également expliqué que j'ai la possibilité de contacter la responsable l'étude, Sylvane Faure (sfaure@unice.fr) ou le Dr Fossoud au 04 92 03 92 00 pour poser des questions à tout moment avant et en cours d'étude.

J'ai été informé(e) :

- que certaines données nominatives concernant mon enfant feront l'objet d'un traitement informatisé dans le respect du secret médical.
- que les données de cette étude, incluant celles portant sur mon enfant, sont couvertes par le secret professionnel.
- de mon droit de m'opposer au traitement automatisé des données nominatives concernant mon enfant.
- que, conformément à la loi du 2012-300 du 05/03/2012, à la fin de l'étude je peux demander à l'investigateur une synthèse des résultats globaux de la recherche. Il m'a été garanti que toute information nouvelle survenant en cours d'essai me sera transmise.

Annexes

- de mon droit d'accès et de rectification à ces données directement ou indirectement par l'intermédiaire d'un médecin de mon choix désigné à cet effet ; je n'autorise leur consultation que par les personnes qui collaborent à la recherche, aux personnes chargées de contrôler la qualité de l'étude ainsi que par un représentant des autorités de santé.

J'atteste que mon enfant ne s'oppose pas à participer à cette étude.

<input type="checkbox"/> Mère	<input type="checkbox"/> Père	<input type="checkbox"/> Titulaire de l'autorité parentale
Nom:.....		
Prénom :.....		
Signature		Date : / /

<input type="checkbox"/> Mère	<input type="checkbox"/> Père	<input type="checkbox"/> Titulaire de l'autorité parentale
Nom:.....		
Prénom :.....		
Signature		Date : / /

<i>L'expérimentateur</i>	
Nom:	
<i>En signant ce formulaire, je confirme que j'ai expliqué l'étude avec exactitude au(x) signataire(s) susmentionnés</i>	
Signature	Date : / /

Fait en deux exemplaires dont un remis au participant.

Questionnaire Points forts - Points faibles (SDQ-Fra)

Pa4-17

Cochez pour chaque énoncé la case "Pas vrai", "Un peu vrai" ou "Très vrai". Cela nous aiderait si vous répondiez, pour chaque énoncé, du mieux que vous pouvez, même si vous n'êtes pas absolument sûr ou si l'énoncé vous paraît inadéquat! Répondez, s'il vous plaît, en vous basant sur le comportement de l'enfant au cours des six derniers mois.

Nom de votre enfant

Garçon/Fille

Date de naissance.....

	Pas vrai	Un peu vrai	Très vrai
Attentif(ve) aux autres, tient compte de ce qu'ils ressentent	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Agité(e), hyperactif(ve), ne tient pas en place	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Se plaint souvent de maux de tête ou de ventre ou de nausées	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Partage facilement avec les autres enfants (friandises, jouets, crayons, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fait souvent des crises de colère ou s'emporte facilement	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Plutôt solitaire, a tendance à jouer seul(e)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
En général obéissant(e), fait habituellement ce que les adultes demandent	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
S'inquiète souvent, paraît souvent soucieux(se)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Aide volontiers quand quelqu'un s'est fait mal ou ne se sent pas bien	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ne tient pas en place ou se tortille constamment	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A au moins un(e) ami(e)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Se bagarre souvent avec les autres enfants ou les tyrannise	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Souvent malheureux(se), abattu(e) ou pleure souvent	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Généralement aimé(e) des autres enfants	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Facilement distrait(e), a du mal à se concentrer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Anxieux(se) ou se cramponne aux adultes dans les situations nouvelles, perd facilement ses moyens	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gentil(le) avec les enfants plus jeunes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ment ou triche souvent	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Harcelé(e) ou tyrannisé(e) par d'autres enfants	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Toujours prêt(e) à aider les autres (parents, professeurs, autres enfants)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Réfléchit avant d'agir	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vole à la maison, à l'école ou ailleurs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
S'entend mieux avec les adultes qu'avec les autres enfants	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A de nombreuses peurs, facilement effrayé(e)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Va jusqu'au bout des tâches ou devoirs, maintient bien son attention	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Avez-vous d'autres préoccupations ou remarques à faire?

Tournez la page - vous trouverez d'autres questions au verso

Dans l'ensemble, estimez-vous que votre enfant éprouve des difficultés dans l'un ou l'autre de ces domaines: émotion, concentration, comportement ou relations avec les autres?

Non	Oui- mineures	Oui- importantes	Oui- sérieuses
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Dans le cas d'une réponse affirmative, veuillez continuer avec les questions suivantes:

«Ces difficultés ont été présentes depuis?

Moins d'un mois	1-5 mois	6-12 mois	Plus d'un an
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

«Est-ce qu'elles dérangent ou gênent votre enfant?

Pas du tout	Un peu	Assez	Beaucoup
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

«Est-ce que ces difficultés interfèrent avec la vie quotidienne de votre enfant dans les domaines suivants?

	Pas du tout	Un peu	Assez	Beaucoup
La vie à la maison	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Les amitiés	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Les apprentissages à l'école	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Les loisirs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

«Est-ce que ces difficultés pèsent sur vous ou sur la famille en général?

Pas du tout	Un peu	Assez	Beaucoup
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Votre signature

Date

Mère/Père/Autre (précisez, SVP:)

Avec tous nos remerciements pour votre aide

©Robert Goodman, 2005

« Processus attentionnels et émotionnels dans le traitement des scènes visuelles complexes chez l'enfant : étude du fonctionnement normal et application au trouble déficitaire de l'attention. »

NOTICE D'INFORMATION destinée aux participants mineurs de 6 à 12 ans

« A LIRE A VOIX HAUTE A L'ENFANT »

Certains enfants rencontrent des difficultés à l'école, parfois ils n'arrivent pas à se concentrer ou à ne pas bouger. Pour mieux comprendre pourquoi certains enfants ont ces difficultés, nous faisons une étude, c'est-à-dire que nous avons créé des activités que l'on va proposer à beaucoup d'enfants, des enfants qui ont des difficultés pour rester concentrés et des enfants qui n'en n'ont pas comme toi. Nous regarderons ensuite si les réponses sont pareilles ou différentes entre les enfants pour pouvoir aider au mieux les enfants qui ont des difficultés à se concentrer. Je te propose de participer à cette étude où il faut regarder et reconnaître des images qui vont apparaître très vite sur un écran d'ordinateur. Il y aura deux séances. Tu peux refuser de participer à l'étude sans que cela ne change rien pour toi. Dis-le-moi simplement ou dis-le à tes parents. Tu pourras changer d'avis à tout moment, il suffira de me le dire ou de le dire à tes parents. L'étude a été expliquée à tes parents, tu peux en discuter avec eux avant de prendre ta décision. S'il y a des mots ou des choses que tu ne comprends pas, tu peux me demander de répéter ou de t'expliquer encore.

Comme tes parents, tu as le choix d'accepter ou de refuser de participer.

« Processus attentionnels et émotionnels dans le traitement des scènes visuelles complexes chez l'enfant : étude du fonctionnement normal et application au trouble déficitaire de l'attention. »

CONSENTEMENT destiné aux participants mineurs de 6 à 12 ans

Donne ta réponse en remplissant la partie ci-dessous (tes parents peuvent t'aider).

Je m'appelle

..... m'a proposé de participer à l'étude et m'a expliqué à quoi ça servira.

Entoure la phrase qui correspond à ton choix :

J'accepte de participer

Je refuse de participer

Date : |_|_|_|_|

Nom de la personne qui a donné l'information :

Date : |_|_|_|_|

Signature

NB : Fait en deux exemplaires

COMITE DE PROTECTION DES PERSONNES SUD MÉDITERRANÉE I

Président
Professeur Stéphane RANQUE

AVIS

Collège technique
Personnes qualifiées en recherche

Yves JANNES
Aurélië MORAND (Pédiatre)
Stéphane RANQUE (méthodologiste)
Agnès ROYER CHAMARD

Marc GANNIER
Thierry BEGE
Jean GALDARY (méthodologiste)
Karim BENDIANE

Médecins généralistes
Jean Charles REYNIER
Michel ROTILY

Pharmaciens hospitaliers
Nicole ROATTINO
Charléric BONNET

Infirmières
Dominique CHANAUD

Collège social

Personnes qualifiées en éthique
Christine ASSALANTE
Célio CHISCHPORTICH

Psychologues

Pierre DE ALCALA
Christian ROCCARDI

Travailleurs sociaux
Lucie CAMILLI

Juristes

Jean-Pierre BINON
Corinne SIMEONE
Mélissa DROUCHE

Représentants d'associations et usagers

Marie-Odile MEYER
Odile DALMONT

Le Comité de Protection des Personnes Sud-Méditerranée I, agréé par arrêté ministériel en date du 1^{er} juin 2018, constitué selon l'arrêté du Préfet de la Région Provence Alpes Côte d'Azur en date du 18 décembre 2018,

en application du code de la santé publique et de la réglementation en vigueur relative aux recherches mentionnées au 2° de l'article L. 1121-1 du code de la santé publique ne portant pas sur un produit mentionné l'article L.5311-1 du Code de la Santé Publique

ayant été saisi par un courrier de la Fondation Lenval - Hôpitaux Pédiatriques de Nice CHU-LENVAL promoteur d'un dossier de recherche intitulée :

« PROCESSUS ATTENTIONNELS ET EMOTIONNELS DANS LE TRAITEMENT DES SCENES VISUELLES COMPLEXES CHEZ L'ENFANT PRESENTANT UN TROUBLE DEFICITAIRE DE L'ATTENTION AVEC OU SANS HYPERACTIVITE (TDA/H).»

Identifiée sous le numéro ID RCB : 2019-A00102-55 et dont l'investigateur (coordinateur) est Mme le Dr. Catherine FOSSOUD

Ayant, après vérification de la conformité réglementaire, enregistré ce dossier le 28 janvier 2019 sous la référence interne 19 19

Lors de sa séance plénière du 13 février 2019 au cours de laquelle Mesdames : L.CAMILLI, D. CHANAUD, C. CHISCHPORTICH, O. DALMONT, M. DRIOUCHE, M-O MEYER, A. MORAND, N. ROATTINO, Messieurs : P. DE ALCALA, K. BENDIANE, M. GAINNIER, S. RANQUE, M. ROTILY

Après avoir entendu le rapporteur du collège technique, le rapporteur du collège social et l'avis du méthodologiste ont délibéré,

Le Comité ayant examiné le dossier de recherche ainsi constitué :

- Courrier de demande d'avis daté et signé du 21 janvier 2019
- Formulaire de demande d'avis daté et signé du 21 janvier 2019
- Document additionnel daté et signé du 21 janvier 2019
- Protocole de recherche, Version n° 0.0 du 21 janvier 2019
- Résumé du protocole en français Version n° 0.0 du 21 janvier 2019
- Document d'information de l'autorité parentale/ enfant de 6 à 12 ans Version n° 0.0 du 21 janvier 2019
- Formulaire de consentement de l'autorité parentale/ enfant de 6 à 12 ans Version n°0.0 du 21 janvier 2019
- Attestation d'assurance du 18 janvier 2019
- Lettre justifiant l'adéquation des moyens humains, matériels et techniques au projet de recherche et leur compatibilité avec les impératifs des personnes qui s'y prêtent daté du 21 janvier 2019

Hôpital Sainte Marguerite
270 Bd Sainte-Marguerite 13274 MARSEILLE
Tél. : 04.91.74.42.56 - Fax : 04.91.74.42.25 - E-mail : cppsudmed1@gmail.com
Site Internet : <http://www.cppsudmediterranee1.fr>

1

COMITE DE PROTECTION DES PERSONNES SUD MÉDITERRANÉE I

Président
Professeur Stéphane RANQUE

- Liste des investigateurs daté du 21 janvier 2019
- CV du Dr Catherine FOSSOUD
- CVs des investigateurs associés

a émis un AVIS FAVORABLE

à la mise en œuvre de cette recherche, considérant que les conditions de validité de la recherche, notamment celles définies dans l'article L. 1123-7 du code de la santé publique, étaient réunies.



Professeur Stéphane RANQUE

ClinicalTrials.gov PRS **DRAFT Receipt (Working Version)**

Last Update: 03/01/2019 04:54

ClinicalTrials.gov ID: NCT03861585

Study Identification

Unique Protocol ID: 18-HPNCL-07

Brief Title: Emotional Information Processing in Attention Deficit Disorder With or Without Hyperactivity (TIVE-TDA)

Official Title: Attentional and Emotional Processing of Complex Visual Scenes in Children With Attention Deficit Disorder With or Without Hyperactivity

Secondary IDs:

Study Status

Record Verification: February 2019

Overall Status: Not yet recruiting

Study Start: March 2019 [Anticipated]

Primary Completion: March 2020 [Anticipated]

Study Completion: September 2020 [Anticipated]

Sponsor/Collaborators

Sponsor: Fondation Lenval

Responsible Party: Sponsor

Collaborators:

Oversight

U.S. FDA-regulated Drug: No

U.S. FDA-regulated Device: No

U.S. FDA IND/IDE: No

Human Subjects Review: Board Status: Approved

Approval Number: CPP Sud Med I

Board Name: Comité de protection des personnes Sud-Méditerranée I

Board Affiliation: Comité de protection des personnes Sud-Méditerranée I

Phone: +33 (0) 491744256

Email: cppsudmed1@gmail.com

Address:

270 bd Sainte Marguerite

13274 Marseille Cedex 9

Data Monitoring: No

FDA Regulated Intervention: No

Study Description

Brief Summary: The aim of this study is to analyse explicit and implicit emotional information processing abilities in children with attention deficit disorder with or without hyperactivity

Detailed Description: The main symptoms of Attentional Deficit Hyperactivity Disorder (ADHD) are inattention, motor agitation and impulsivity. However, other dysfunctions affecting the quality of life remain poorly studied: lack of understanding and management of emotions, focus on the local aspects of a visual scene limiting the ability to assign a general meaning to the scene and alteration of long-term memory encoding. This study aims to analyse these difficulties using different tasks requiring processing of rich and varied everyday images, having high ecological validity. It involves the participation of 56 boys and girls with ADHD, aged 7 to 12 years. A first phase examines the immediate understanding of images using two tasks: semantic categorization (Experiment 1) and emotional evaluation (Experiment 2) of images with positive, negative or neutral emotional valence, and depicting real environments (natural vs. manufactured contexts) or foreground objects pasted into a noise background (inanimate objects vs. animals vs. people. In each trial, one context image and one object image are presented briefly and simultaneously, one in each visual field. In order to be appropriately understood in both their semantic and emotional contents, context images will require more global processing, while object images will require more local, detailed processing. Their semantic (Exp. 1) and emotional (Exp. 2) consistency is manipulated. A week later, the participants have to perform a memory task requiring old/new recognition in a two-alternative forced choice (2AFC) paradigm that presents in each trial a pair of images (one old, one new) having the same emotional valence (Exp. 3). The study will characterize the specificities of processing and representing visual emotional information in ADHD children. The results will be compared with those from a previous study we conducted with the same methodology on neuro-typical children (controls).

Conditions

Conditions: Attention Deficit Disorder With Hyperactivity
Attention Deficit Disorder Without Hyperactivity

Keywords:

Study Design

Study Type: Interventional

Primary Purpose: Basic Science

Study Phase: N/A

Interventional Study Model: Parallel Assignment

Number of Arms: 2

Masking: None (Open Label)

Allocation: Randomized

Enrollment: 56 [Anticipated]

Arms and Interventions

Arms	Assigned Interventions
Experimental: semantic group Half of the participants (named the "semantic group") perform a semantic categorisation task.	Semantic group The "semantic task" has to categorise semantically (i.e., type of contexts or type of objects) one image at each trial.
Experimental: emotional group Half of the participants (named the "emotional group") perform an emotional evaluation task.	Emotional group The "emotional task" has to evaluate the emotional content (i.e., neutral, positive or negative) of one image at each trial.

Outcome Measures

Primary Outcome Measure:

1. Accuracy and response time at an explicit emotional evaluation task of visual scenes, by children with Attentional Deficit Hyperactivity Disorder (ADHD)
 The performance at the emotional evaluation task is analyzed by association of two separate outcomes measures: measure of efficiency by the accuracy (i.e. proportion of correct response) and measure of the effectiveness by the response time at the correct trials (in milliseconds). This experiment analyzed the explicit processing of visual scenes' emotional content .
 [Time Frame: at baseline Day 0]
2. Accuracy and response time at a semantic categorization task of visual scenes, by children with Attentional Deficit Hyperactivity Disorder (ADHD)
 The performance at the semantic categorization task is analyzed by association of two separate outcomes measures: measure of efficiency by the accuracy (i.e. proportion of correct response) and measure of the effectiveness by the response time at the correct trials (in milliseconds). This experiment analyzed the implicit processing of visual scenes' emotional content .
 [Time Frame: at baseline Day 0]

Secondary Outcome Measure:

3. Proportion and response time of correct response for the contextual images compared to the local images at the semantic categorization task for ADHD's participants
 The performance is analyzed by association of two separate outcomes measures: measure of efficiency by the accuracy (i.e. proportion of correct response) and measure of the effectiveness by the response time at the correct trials (in milliseconds).
 [Time Frame: at baseline Day 0]
4. Proportion and response time of correct response for the contextual images compared to the local images at the emotional evaluation task for ADHD's participants
 The performance is analyzed by association of two separate outcomes measures: measure of efficiency by the accuracy (i.e. proportion of correct response) and measure of the effectiveness by the response time at the correct trials (in milliseconds).
 [Time Frame: at baseline Day 0]
5. Influence of the image's emotional content (i.e. neutral, positive or negative) on long term memory in ADHD participants
 The performance, in a two-alternative forced choice (2 AFC) image recognition, is analyzed by two separate outcomes measures: measure of efficiency by the accuracy (i.e. proportion of correct response) and measure of the effectiveness by the response time (in milliseconds).
 [Time Frame: at 1 week from the first evaluation]
6. Performance on long term memory in ADHD participants

The performance, in a two-alternative forced choice (2 AFC) image recognition, is analyzed by two separate outcomes measures: measure of efficiency by the accuracy (i.e. proportion of correct response) and measure of the effectiveness by the response time (in milliseconds).

[Time Frame: at 1 week from the first evaluation]

Eligibility

Minimum Age: 7 Years

Maximum Age: 12 Years

Sex: All

Gender Based: No

Accepts Healthy Volunteers: No

Criteria: Inclusion Criteria:

- Children between 7 and 12 years old.....
- Normal or corrected visual acuity
- ADHD diagnostic given by a doctor according to DSM 5 criteria (Diagnostic and Statistical Manual of Mental Disorders)
- Free from drug treatment until 24 hours before the experiment session.....
- Affiliation to a social security system.....
- Signature of the authorization documents from one parent or the holder of parental authority

Exclusion Criteria:

-History of neurological troubles, dysphasia, autistic spectrum disorder, intellectual disabilities

Contacts/Locations

Central Contact Person: FOSSOUD Catherine, MD
Telephone: +33(0)492039201
Email: fossoud.c@pediatrie-chulenal-nice.fr

Central Contact Backup: Faure Sylvane, Pr
Telephone: +33(0) 4 89 88 14 60
Email: Sylvane.FAURE@univ-cotedazur.fr

Study Officials: FOSSOUD Catherine, MD
Study Principal Investigator
Fondation Lenal - Nice Children Hôpitaux Pédiatriques de Nice

Locations: **France**
Hôpitaux Pédiatriques de Nice CHU-Lenal
Nice, France
Contact: Fossoud Catherine, MD fossoud.c@pediatrie-chulenal-nice.fr

IPDSharing

Plan to Share IPD:

References

Citations:

Links:

Available IPD/Information:

Annexe 10 : Notice d'information détaillée sur le protocole expérimental et formulaire de consentement éclairé à l'intention des parents (version patient)

« Processus attentionnels et émotionnels dans le traitement des scènes visuelles complexes chez l'enfant présentant un Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H) »
Promoteur de l'étude : Fondation LENVAL, Hôpitaux Pédiatriques de Nice CHU-LENVAL,
57, Avenue de la Californie – 06200 Nice,
Tel : 04 92 03 03 92 ; Fax : 04 92 03 03 44, fondation@lenval.com
Investigateur principal: Dr. Fossoud– CERTA – Hôpitaux pédiatriques de Nice CHU-Lenval – 57
av. de la Californie 06200 Nice
Tel: (04.92.03.92.00) - Fax: (04.92.03.92.09)

NOTICE D'INFORMATION destinée aux Parents/Titulaires de l'autorité parentale pour la participation d'une personne MINEURE

Madame, Monsieur,

Votre enfant est pris en charge par le Centre de Référence des Troubles des Apprentissages (CERTA) pour des difficultés d'apprentissages, liées notamment à un trouble déficitaire de l'attention avec ou sans hyperactivité.

A ce titre, nous vous proposons de faire participer votre enfant à un projet de recherche impliquant la personne humaine intitulé « *Processus attentionnels et émotionnels dans le traitement des scènes visuelles complexes chez l'enfant présentant un Trouble Déficitaire de l'Attention avec ou sans Hyperactivité* ».

Avant d'accepter qu'il(elle) participe à l'étude, il est important que vous preniez le temps de lire, de comprendre et de considérer attentivement les renseignements qui suivent.

Le présent document vous renseigne sur les modalités de ce projet de recherche. Si certaines choses ne vous semblent pas claires, n'hésitez pas à poser des questions ou demander davantage d'explications.

Si vous refusez que votre enfant participe à l'étude, cela ne modifiera pas la qualité de la prise en charge médicale de votre enfant. Avant de prendre votre décision, vous bénéficiez d'un délai de réflexion.

Si vous acceptez que votre enfant participe à l'étude, nous souhaitons que sa participation soit **volontaire et libre de toute contrainte**. Afin d'éclairer votre décision, vous devez avoir reçu et bien compris les informations suivantes :

OBJECTIFS DE L'ETUDE :

Le but de l'étude à laquelle nous vous proposons de faire participer votre enfant est d'analyser comment les enfants qui présentent un TDA/H traitent, c'est-à-dire reconnaissent et mémorisent, des images porteuses ou non de contenu émotionnel. Le TDA/H occasionne des difficultés pour rester attentif, concentré et peut entraîner un besoin constant de bouger. Ces difficultés peuvent se répercuter sur la scolarité et les relations sociales et affectives de ces enfants. Aussi, les enfants présentant un TDA/H pourraient être en difficulté pour traiter certaines informations contenues dans des images. C'est pourquoi dans cette étude nous analysons comment les enfants reconnaissent différentes informations présentes dans les images : des informations sur le sens global de l'image (par exemple, reconnaître une plage, un restaurant), des informations plus précises sur les objets au sein de l'image

(comme un palmier, des assiettes, une personne...) et des informations sur les émotions que procurent ces images (émotion positive, négative ou images neutres sans émotion particulière).

DEROULEMENT DE L'ETUDE :

→ Quel sera le déroulement de cette étude ? Selon quelle méthodologie sera-t-elle mise en œuvre ?

Cette étude est réalisée par le Centre de Référence des Troubles des Apprentissages (CERTA) des Hôpitaux Pédiatriques de Nice CHU-Lenval. Elle est proposée à des enfants de 7 à 12 ans et se réalise en deux séances d'environ 30 mn chacune, espacées d'une semaine, au CERTA. La première séance consiste à juger des images. L'enfant doit soit choisir entre différentes catégories (par exemple entre « une personne », « un animal » ou « un objet ») celle que l'image représente pour lui (jugement du contenu sémantique), soit juger si il trouve que cette image fait ressentir une émotion « positive » (joyeuse, agréable), « négative » (de tristesse, de peur ou de colère) ou « neutre » (qu'elle ne fait pas ressentir d'émotion particulière). Les images sont présentées très rapidement sur un écran d'ordinateur. Durant la seconde séance qui aura lieu une semaine après la 1ère séance, deux images sont présentées à chaque essai à l'enfant, qui doit reconnaître celle provenant de la série d'images qu'il a vue lors de la première séance. Etant donné que nous étudions ici la mémorisation non volontaire des images, il est nécessaire de ne pas informer l'enfant qu'il y aura une seconde séance où il devra reconnaître les images que nous lui avons présentées lors de la première séance. Enfin, un questionnaire évaluant la gestion des émotions, les relations sociales et le comportement de votre enfant vous sera proposé, ainsi qu'à l'enseignant de votre enfant, afin d'apporter des éléments supplémentaires. Ce questionnaire intitulé « points forts, points faibles » existe en version parents et en version enseignant, si vous souhaitez consulter au préalable ce questionnaire, vous pouvez en faire la demande auprès d'Anaïs Leroy (anaïs.leroy@univ-cotedazur.fr), investigatrice, doctorante en psychologie et psychomotricienne, qui réalisera les séances avec enfant.

Tableau synoptique

	<i>T0 – Inclusion</i>	<i>Séance 1</i>	<i>Séance 2</i>
<i>Examen médical</i>	X		
<i>Vérification des critères d'inclusion et de non inclusion</i>	X		
<i>Signature du consentement éclairé</i>	X		
<i>Jugement des images</i>		X	
<i>Reconnaissance des images</i>			X
<i>Questionnaires</i>	X	X	

→ **Quelles seront les précautions particulières que votre enfant devra prendre et les impératifs qu'il (elle) aura à respecter ?**

Si votre enfant suit un traitement médicamenteux (méthylphénidate) dans le cadre de la prise en charge du TDA/H, il ne devra pas prendre son traitement durant les 24 heures précédant chacune des séances.

→ **Quel est le nombre de sujets participant à l'étude ?**

Cinquante-six enfants présentant un trouble déficitaire de l'attention avec ou sans hyperactivité seront inclus dans cette étude.

→ **Quelle sera la durée de cette étude ?**

La durée prévisionnelle de l'étude est de 18 mois. Avec une durée d'inclusion de 1 an.

La participation de votre enfant à l'étude durera 7 jours.

RAPPORT BENEFICE/RISQUE DE L'ETUDE :

→ **Quels sont les bénéfices attendus ?**

Les résultats de l'étude contribueront à l'acquisition de nouvelles connaissances sur les particularités de traitement des informations visuelles par les enfants présentant un TDA/H afin d'améliorer les outils diagnostiques et de prise en charge.

→ **Quels sont les risques prévisibles?**

Il n'existe aucun inconvénient majeur potentiel pour les participants à cette étude. Les risques potentiels mineurs sont une fatigue visuelle et/ou des maux de tête provenant du temps passé à regarder l'écran d'ordinateur ; pour diminuer ces risques chaque séance dure moins de 45 minutes et inclut trois pauses. Le risque que l'enfant soit choqué par certaines images a été limité au maximum : nous avons sélectionné les images de façon à ce qu'aucune ne soit choquante et nous nous sommes assurées que des parents d'enfants âgés de 7 à 8 ans confirmaient unanimement notre jugement.

→ **Quelles sont les contraintes?**

Vous devrez vous déplacer avec votre enfant au CERTA ; pour réduire cette contrainte, lorsque cela sera possible nous proposerons un rendez-vous attendant à un autre rendez-vous déjà prévu au CERTA.

ALTERNATIVES MEDICALES ET ARRET PREMATURE :

→ **Quelles seront les modalités de sa prise en charge à la fin de cette recherche, en cas d'arrêt prématuré de la recherche ou en cas d'exclusion ?**

Sa participation à l'étude pourra être prématurément arrêtée avant la visite finale en cas de survenue d'un événement indésirable grave, par décision du promoteur, de l'autorité compétente (ANSM) ou du médecin investigateur qui peut aussi décider de le (la) faire sortir d'étude si il (elle) n'en respecte pas les impératifs.

A la fin de l'étude, il (elle) reprendra le cours habituel de sa prise en charge.

MODALITES ADMINISTRATIVES:

→ **Qui a autorisé la mise en œuvre de ce projet de recherche ?**

Cette étude a été soumise au Comité de Protection des Personnes (**intitulé exact du CPP**) et a reçu son avis favorable le (**date de l'avis favorable**).

Cette étude a été transmise pour information à l'Agence Nationale de Sécurité du Médicament et des produits de santé (ANSM) selon la réglementation en vigueur. (Catégorie 2)

→ **Assurance**

Conformément aux dispositions des articles L.1121-10 et R.1121-10 du CSP, la Fondation LENVAL, Hôpitaux Pédiatriques de Nice CHU-LENVAL assume des conséquences dommageables de la recherche pour la personne qui s'y prête et ses ayants droit.

A cet effet, la Fondation LENVAL, Hôpitaux Pédiatriques de Nice CHU-LENVAL a souscrit auprès de la Société Hospitalière d'Assurances Mutuelles un contrat d'assurance (sous le n°105.617) garantissant sa responsabilité civile ainsi que celle de tout intervenant à la recherche.

RESULTATS DE L'ETUDE:

→ **Serez vous informé des résultats globaux et donc des connaissances acquises grâce à cette recherche ?**

A l'issue de cette recherche, les résultats globaux de l'étude après leur analyse, vous seront communiqués par courrier si vous le souhaitez.

Vous devrez faire savoir votre souhait au médecin de votre enfant ou à Mme Anaïs Leroy (anaïs.leroy@univ-cotedazur.fr).

PARTICIPATION :

→ **Pourra-t-il (elle) participer simultanément à une autre recherche et à la fin de cette étude pourra-t-il (elle) immédiatement participer à une autre recherche ?**

Oui, la participation à cette étude n'aura aucune répercussion sur les capacités de votre enfant, qui pourra participer à d'autres recherches.

→ **Vous décidez de le faire participer à l'étude, que devez-vous faire concrètement ?**

Si vous acceptez que votre enfant participe à cette étude, vous devez obligatoirement lire, dater et signer l'autorisation de participation accompagnant cette notice d'information. Vous avez le droit de refuser que votre enfant participe à l'étude ou de retirer votre consentement à tout moment, et sans encourir aucune responsabilité, ni aucun préjudice de ce fait ; vous devrez alors en informer le Dr Fossoud investigateur de l'étude.

PRISE EN CHARGE DES FRAIS DE DEPLACEMENTS :

Un forfait de 25€ vous est proposé comme indemnisation de vos frais de déplacement, pour en bénéficier il vous suffit de nous transmettre votre relevé d'information bancaire.

CONFIDENTIALITE DES DONNEES :

Dans le cadre de la recherche médicale à laquelle le Docteur Fossoud vous propose de participer, un traitement de vos données personnelles va être mis en œuvre pour permettre d'analyser les résultats de la recherche au regard de l'objectif de cette dernière.

A cette fin, les données médicales vous concernant et les données relatives à vos habitudes et votre qualité de vie, seront transmises au promoteur de la recherche ou aux personnes agissant pour son compte, en France. Ce recueil concerne vos données de santé et seront identifiées par un numéro de code et la première lettre de votre nom ainsi que la première lettre de votre prénom.

Ces données pourront également, dans des conditions assurant leur confidentialité, être transmises aux autorités de santé françaises ou étrangères, à d'autres entités des Hôpitaux Pédiatriques de Nice CHU-Lenval. Le traitement de ces données est conforme à la loi Informatique et Libertés n° 78-17

du 6 janvier 1978 modifiée et ainsi qu'au règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 applicable à compter du 25 mai 2018 (règlement général sur la protection des données).. La base juridique de ce traitement est l'intérêt légitime poursuivi par les Hôpitaux Pédiatriques de Nice CHU-Lenval pour améliorer la prise en charge de cette pathologie. Le responsable du traitement informatique est le Directeur Général des Hôpitaux Pédiatriques de Nice CHU-Lenval. La Fondation Lenval a désigné un Délégué à la Protection des Données (DPO) que vous pouvez joindre au contact.dpo@lenval.com et qui pourra répondre à toutes vos questions concernant la protection des données personnelles.

Vous pouvez exercer votre droit d'accès et de rectification aux informations qui vous concernent, ou une limitation du traitement relatif à votre personne en écrivant au Délégué à la Protection des Données (DPO) des Hôpitaux Pédiatriques de Nice CHU-Lenval. De même vous aurez le droit de vous opposer au traitement automatisé des données vous concernant et vous avez également le droit de porter une réclamation auprès de la Commission Nationale Informatique et Libertés si vos données ne sont pas traitées conformément à la réglementation des données personnelles

Les données de l'étude pourront être publiées dans des journaux médicaux et scientifiques sans que votre identité soit révélée. Le promoteur est responsable du traitement des données de l'étude Le promoteur est responsable du traitement des données de l'étude et a signé un engagement de conformité à la méthodologie de référence MR001 de la CNIL.

Les données de l'étude pourront être publiées dans des journaux médicaux et scientifiques sans que son identité soit révélée. Le promoteur est responsable du traitement des données de l'étude et a signé un engagement de conformité à la méthodologie de référence MR001 de la CNIL.

COORDONNEES DU OU DES MEDECINS RESPONSABLES :

Pour toute question concernant l'étude, ainsi que vos droits, vous pourrez contacter le:

Dr Catherine Fossoud

Téléphone : 04.92.03.92.00

Télécopie : 04.92.03.92.09

« Traitement des informations visuelles émotionnelles dans le trouble déficitaire de l'attention avec ou sans hyperactivité »

*Promoteur de l'étude : Fondation LENVAL, Hôpitaux Pédiatriques de Nice CHU-LENVAL,
57, Avenue de la Californie – 06200 Nice,
Tel : 04 92 03 03 92 ; Fax : 04 92 03 03 44, fondation@lenval.com*

*Investigateur principal : Dr. Fossoud – CERTA –
Hôpitaux pédiatriques de Nice CHU-Lenval – 57 av. de la Californie 06200 Nice
Tel: (04.92.03.92.00) - Fax: (04.92.03.92.09)*

CONSENTEMENT ECLAIRE – TITULAIRES DE L'AUTORITE PARENTALE

Je déclare avoir compris le but et les modalités de cette étude, qui m'ont été pleinement expliqués par le médecin signataire.

Les informations relatives au principe de l'étude, et son intérêt m'ont bien été communiquées dans la Note d'Information. J'ai eu la possibilité de l'étudier attentivement. Des réponses ont été apportées à toutes mes questions. J'ai disposé d'un délai de réflexion avant de prendre ma décision.

J'accepte de faire participer volontairement mon enfant au **protocole de recherche impliquant la personne humaine « TIVE_TDA »** dont le promoteur est la Fondation LENVAL, Hôpitaux Pédiatriques de Nice CHU-LENVAL et l'investigateur coordonnateur est le Dr Catherine Fossoud du CERTA. Il m'a bien été précisé que je pouvais refuser de faire participer mon enfant à cette étude et que si je souhaitais le faire participer, je pouvais revenir sur ma décision à tout moment. Il m'a également été expliqué que j'ai la possibilité de contacter l'un des investigateurs de l'étude (coordonnées à la fin du consentement) pour poser des questions à tout moment avant et en cours d'étude. J'ai bien compris que le traitement approprié à sa maladie (état de santé) sera mis en place, sans qu'à aucun moment sa participation ou sa non-participation au protocole de recherche « **TIVE_TDA** » n'ait une quelconque influence sur son suivi ou sur son traitement.

J'ai été informé(e) :

- que, pour cette étude, le promoteur a souscrit une assurance en Responsabilité Civile auprès de la SHAM, sous le numéro de contrat 145.017.
- que, conformément à l'article L.1121-11 du Code de la Santé Publique l'affiliation à un régime de Sécurité Sociale sont obligatoires
- que certaines données nominatives le concernant feront l'objet d'un traitement informatisé dans le respect du secret médical.
- que les données de cette étude, incluant celles portant sur son cas, sont couvertes par le secret professionnel.
- de mon droit de m'opposer au traitement automatisé des données nominatives le concernant.
- que toutes les données resteront confidentielles.
- que, conformément à la loi 2012-300 du 05/03/2012, à la fin de l'étude je peux demander à l'investigateur une synthèse des résultats globaux de la recherche. Il m'a été garanti que toute information nouvelle survenant en cours d'essai me sera transmise.

Annexes

- de mon droit d'accès et de rectification à ces données directement ou indirectement par l'intermédiaire d'un médecin de mon choix désigné à cet effet ou auprès du Délégué à la Protection des Données. Je n'autorise leur consultation que par les personnes qui collaborent à la recherche, aux personnes chargées par le promoteur de contrôler la qualité de l'étude ainsi que par un représentant des autorités de santé.
- que cette étude a été déclarée à la Commission Nationale Informatique et Liberté selon l'engagement de conformité MR001.
- Les procédures médicales de cette étude sont conformes aux recommandations nationales et l'étude a obtenu l'accord du Comité de Protection des Personnes Sud Méditerranée I, avis délivré le 28/01/2019

J'atteste que mon enfant (nom/ prénom), ne s'oppose pas à participer à cette étude clinique.

<input type="checkbox"/> Mère	<input type="checkbox"/> Père	<input type="checkbox"/> Titulaire de l'autorité parentale
Nom:.....		
Prénom :.....		
Signature		Date : / /

LE MEDECIN INVESTIGATEUR Nom et prénom du médecin (en majuscules) :.....	
Téléphone :	
<i>En signant ce formulaire, je confirme que j'ai expliqué l'étude avec exactitude au(x) signataire(s) susmentionné(s)</i>	
Signature	Date : / /

Fait en deux exemplaires dont un remis au(x) parent(s).

Annexe 11 : Formulaire d'inclusion (version patient)

Questionnaire d'inclusion

*A compléter par les parents
En cas de doute, ne pas hésiter à demander des informations complémentaires*

Nom de l'examineur :
Date du jour :
Code identification :

} A remplir par
l'examineur

Date de naissance :
Age :
Classe actuelle:
Etablissement scolaire :
Latéralité manuelle : Droitier Gaucher Indéterminé Ambidextre
Sexe : Masculin Féminin

Critères d'inclusion :

Age compris entre 7 et 12 ans au moment de l'étude : Oui Non
Diagnostic établi par un centre spécialisé : Oui Non
Forme diagnostiquée : Mixte (TDAH) Inattentif (TDA) Hyperactif/impulsif

Critères d'exclusion :

Antécédents de pathologie neurologique acquise diagnostiquée : Oui Non
Antécédents de pathologie psychiatrique diagnostiquée : Oui Non
Déficience intellectuelle diagnostiquée : Oui Non
Trouble de la compréhension verbale pouvant pénaliser la passation : Oui Non
Trouble sensoriel élémentaire pouvant pénaliser la passation (*) : Oui Non
Absence de consentement des parents ou du tuteur légal : Oui Non
Attention ! Une réponse « oui » à l'une de ces questions ne permet pas d'inclure l'enfant dans l'étude.
(*) Notamment, déficience auditive ou visuelle

Informations complémentaires :

Problème de vision : Oui Non
si oui, quel est le type de correction ?.....
Traitement(s) médicamenteux en cours : Oui Non
si oui, lesquels ?.....
Scolarité : Redoublement Saut de classe
si oui, indiquez la classe concernée :.....
Aide scolaire (AVSI, RASED, ...) ?
si oui, précisez la nature, durée et fréquence de cette aide :.....
Diagnostic d'autre(s) trouble(s) des apprentissages : Oui Non
si oui précisez lequel (lesquels) :
Suivi paramédical : Orthophoniste Psychomotricien(ne)
Psychologue/Neuropsychologue Ergothérapeute Autre
si oui précisez la fréquence des suivis :
Niveau étude/ profession du père :.....
Niveau d'étude/ profession de la mère :.....
Numéro de téléphone et/ou adresse mail :

Annexe 12 : Notice d'information sur le protocole expérimental et consentement à l'intention des enfants (version patient)

«Processus attentionnels et émotionnels dans le traitement des scènes visuelles complexes chez l'enfant présentant un Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H) »

Promoteur de l'étude: Fondation LENVAL, Hôpitaux Pédiatriques de Nice CHU-LENVAL,
57, Avenue de la Californie – 06200 Nice,
Tel : 04 92 03 03 92 ; Fax : 04 92 03 03 44, fondation@lenval.com

Investigateur principal : Dr Fossoud.– CERTA
Hôpitaux Pédiatriques de Nice CHU-LENVAL
57, avenue de la Californie- 06200 Nice
Tel: 01 92 03 05 27 - Fax: 04 92 03 04 87

Notice d'Information destinée aux Patients Mineurs de 6 à 12 ans

Bonjour,

Tu viens au CERTA car tu as parfois des difficultés à te concentrer longtemps à l'école ou à écouter sans bouger. C'est pour cela que je te propose de participer à une étude.

Lis bien les explications qui suivent. Tu pourras poser toutes les questions que tu veux au médecin.

Nous avons créé des activités que l'on va proposer à beaucoup d'enfants, pour essayer de mieux comprendre et aider les enfants qui ont des difficultés à se concentrer.

Si tu veux participer, il faudra venir 2 fois. Je te proposerai des activités sur un ordinateur avec des images.

Tu n'es pas obligé(e) de participer à cette étude.

Si tu choisis de participer à cette étude tu pourras aussi changer d'avis et arrêter à n'importe quel moment en le disant à ton médecin et à tes parents. Personne ne t'en voudra. Ton médecin continuera à bien s'occuper de toi et gardera de très bonnes relations avec tes parents, comme auparavant.

Annexes

Saches que si cela arrive il t'en expliquera les raisons.

Dans le cadre de cette étude, toutes les informations qui ont été recueillies à ton sujet seront maintenues secrètes et ton nom ne figurera pas sur les documents de l'étude.

Tu n'es pas obligé(e) de donner ta réponse tout de suite, tu peux réfléchir quelques jours et en discuter avec tes parents avant de donner ton accord définitif.

«Processus attentionnels et émotionnels dans le traitement des scènes visuelles complexes chez l'enfant présentant un Trouble Déficitaire de l'Attention avec ou sans Hyperactivité (TDA/H) »

Promoteur de l'étude: *Fondation LENVAL, Hôpitaux Pédiatriques de Nice CHU-LENVAL, 57, Avenue de la Californie – 06200 Nice, Tel : 04 92 03 03 92 ; Fax : 04 92 03 03 44, fondation@lenval.com*

Investigateur principal: *Dr Fossoud.– CERTA- Hôpitaux Pédiatriques de Nice CHU-LENVAL - 57, avenue de la Californie- 06200 Nice Tel: 01 92 03 05 27 - Fax: 04 92 03 04 87*

CONSENTEMENT destiné aux patients mineurs de 6 à 12 ans

Donne ta réponse en remplissant la partie ci-dessous (tes parents peuvent t'aider).

Je m'appelle

Le Docteur m'a proposé de participer à l'étude et m'a expliqué à quoi ça servira.

Entoure la phrase qui correspond à ton choix :

J'accepte de participer

Je refuse de participer

Date : |__|__|____|

Nom du médecin qui a donné l'information :

Date : |__|__|____|

Fait en deux exemplaires dont un remis au(x) parent(s).

Signature du médecin

Annexe 13 : Poster présenté à la Société Française de Psychologie



Mémoire à long terme d'images perçues brièvement : influence du contenu émotionnel et des niveaux global et local
Anais Leroy,* Sylvane Faure* & Sara Spotorno[§]



*Laboratoire d'Anthropologie et de Psychologie Cliniques, Cognitives et Sociales, Université Côte d'Azur, France
[§]Institute of Neuroscience and Psychology, University of Glasgow, United Kingdom

A.L. est financée par le ministère français de l'enseignement supérieur et de la recherche

Contact: anais.leroy@unice.fr

Introduction

L'influence mnésique de la valence émotionnelle d'images visuelles a été abordée surtout en considérant la mémoire à court terme [1-2], alors que l'impact sur la mémoire à long terme demeure peu étudié. De plus, les études ont typiquement présenté des images seules en vision centrale lors de l'encodage, sans prendre en compte que dans notre vie quotidienne nous traitons de façon simultanée et très rapide plusieurs informations, aussi en vision périphérique, qui peuvent être porteuses de même valence émotionnelle ou non.

Cette étude a donc pour objectif d'analyser l'influence du contenu émotionnel et de l'orientation attentionnelle sur la reconnaissance à long terme d'images initialement présentées brièvement, par paires et latéralisées.

Participants : 46 (25 femmes) ; âge : $M=28.82$ ans, $SD=8$.

Matériel : 288 images porteuses de valence émotionnelle positive ou négative et 144 images neutres, provenant de l'IAPS [11], de la NAPS [12], de la base de données SUN [13] et de Google images.

Phase d'encodage : paires d'images (un contexte visuel, par ex. une forêt, et une image avec un objet animé ou inanimé en premier plan, par ex. une biche ; Fig. 1), présentées rapidement lors de 4 tâches différentes (11 participants pour T1 et T4, 12 pour T2 et T3) :

- T1 : catégorisation sémantique d'une image ;
- T2 : jugement de la valence émotionnelle d'une image ;
- T3 : jugement de la cohérence sémantique de la paire ;
- T4 : jugement de la cohérence émotionnelle de la paire.

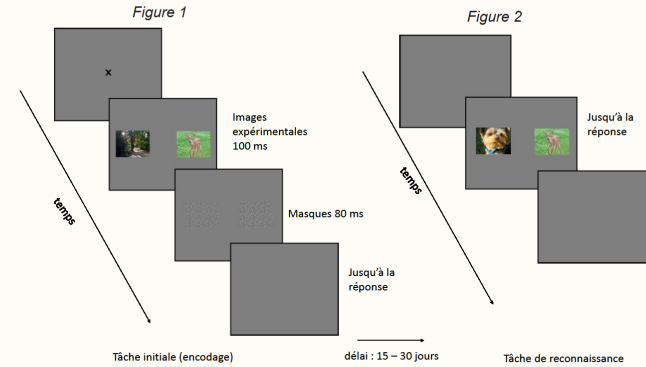
Ces tâches permettent d'orienter l'attention de façon implicite/involontaire (T1 et T3) ou explicite/volontaire (T2 et T4) sur la valence émotionnelle des images.

Conditions de valence émotionnelle de la paire d'images lors de l'encodage : 2 image neutres, 2 images positives, 2 images négatives, 1 image négative et 1 positive.

Phase de reconnaissance en mémoire à long terme, dont les participants étaient informés seulement après les tâches initiales d'encodage.

A chaque essai, présentation de deux images-contextes ou deux images-objets, dont une image ancienne (présentée lors de la tâche initiale) et une nouvelle, ayant la même valence émotionnelle (Fig. 2). Réponse manuelle en choix forcé : sélection sur clavier de la touche correspondant à l'image ancienne.

Méthode

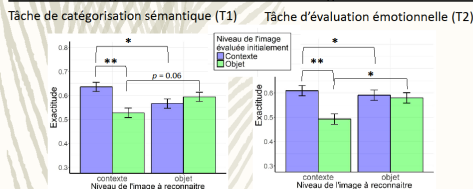


Résultats

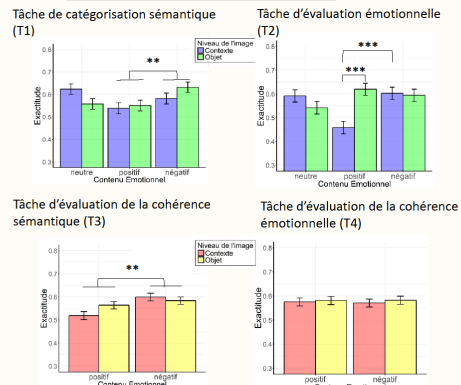
Analyses : VD : exactitude de la reconnaissance. Modèles mixtes linéaires généralisés (GLMMs ; séparés pour chaque tâche initiale), avec le niveau de l'image à reconnaître (contexte ou objet : T1-T4), sa valence émotionnelle (positive vs. négative vs. neutre : T1 et T2 ; positive vs. négative : T3 et T4), le niveau de l'image initialement jugée (contexte ou objet : T1 et T2) en prédicteurs, et les participants et essais en facteurs aléatoires (T1-T4). Tests t à un échantillon pour comparaisons au niveau hasard (0.50).

L'exactitude est supérieure au niveau hasard dans toutes les conditions expérimentales ($t_s > 2.5$, $p_s < 0.009$) à part pour les contextes positifs dans T2 ($t < 1$, $p = 0.220$) et T3 ($t < 1$, $p = 0.70$).

Niveau de l'image à reconnaître x niveau de l'image à évaluer initialement



Valence émotionnelle x niveau de l'image



Discussion

- En premier lieu, notre étude montre des **capacités mnésiques surprenantes**, car la tâche de mémoire n'était pas attendue et la présentation à l'encodage des paires d'images était très brève et latéralisée.
- Nos résultats suggèrent une **compétition à la fois perceptive et émotionnelle entre les deux images perçues simultanément**. Les images contenant un objet en premier plan sont priorisées dans cette compétition par rapport à celles contenant essentiellement des informations de contexte, distribuées de façon plus homogène, globale au sein de l'image donc moins distinctes sur les plans perceptif et émotionnel et par conséquent, moins prégnantes au niveau représentationnel.
- Nous mettons en avant des effets différents entre **attention volontaire**, portée sur une image-cible et sur les aspects spécifiques de son contenu pertinent pour la tâche courante, et **attention involontaire**, dirigée vers l'image non-cible et les aspects non pertinents pour la tâche (mais potentiellement importants pour l'individu, comme les contenus émotionnels):
 - En particulier, la meilleure reconnaissance des images négatives par rapport aux positives lorsque l'attention volontaire n'a pas été dirigée vers la valence émotionnelle lors de l'encodage (T1 et T3), suggère un **traitement automatique et implicite des contenus négatifs**.
 - L'orientation de l'attention volontaire sur les émotions (et sur une seule image lors de l'encodage : T2) améliore la reconnaissance des objets à valence positive, alors que l'exactitude pour les contextes à valence positive diminue dans ces conditions spécifiques. Cette chute de performance pourrait s'expliquer par une compétition attentionnelle avec l'objet émotionnel associé dans la paire initialement encodée.
 - Enfin, le fait de porter attention volontairement aux deux images et à leur valence émotionnelle (T4) permet de traiter et de reconnaître de façon identique les images positives et négatives, présentant ou non un objet en premier plan.

Perspectives : notre étude montre la nécessité de prendre en considération les aspects perceptifs et attentionnels pour mieux comprendre l'influence des émotions sur la mémoire. La comparaison des résultats à ceux obtenus lors des tâches initiales d'encodage permettra d'analyser plus finement les effets d'interférence et de coût cognitif de traitement de l'information. De plus, les études futures devraient comparer les images en fonction de leur niveau d'éveil (« arousal ») et distinguer les différentes émotions de base (comme peur, tristesse et colère).

Références

[1] Dolcos, F., Wang, L., & Mathew, M. (2014). Current research and emerging directions in emotion-cognition interactions. *Front. Integr. Neurosci.*, 8:83.
 [2] Koenigs, F. A. (2009). Remembering the details: Effects of emotion. *Emotion Review*, 12(2), 59e113.
 [3] Lang, P.J., Bradley, M.M., Cuthbert, B.N.(2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. Univ. of Florida, Gaines, FL, US.
 [4] Marchewka A., Żurawski L., Jednoróg K., Grabowska A. (2014) The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, 46(2), 596-610.
 [5] Xiao, J., Ehinger, K. A., Hays, J., Torralba, A., & Oliva, A. (2016). Sun database: Exploring a large collection of scene categories. *International Journal of Computer Vision*, 219(1), 3-22.

**Formulaire de Consentement libre et éclairé
pour une expérimentation en psychologie expérimentale**

Je certifie avoir donné mon accord pour participer à cette étude de psychologie. J'accepte volontairement d'y participer et je comprends que ma participation n'est pas obligatoire et que je peux stopper ma participation à tout moment sans avoir à me justifier ni encourir aucune responsabilité. Mon consentement ne décharge pas les organisateurs de la recherche de leurs responsabilités et je conserve tous mes droits garantis par la loi.

Au cours de cette expérience, j'accepte que soient recueillies des données comportementales. Je comprends que les informations recueillies sont strictement confidentielles et à usage exclusif des investigateurs concernés.

J'ai été informé que mon identité n'apparaîtra dans aucun rapport ou publication et que toute information me concernant sera traitée de façon confidentielle. J'accepte que les données enregistrées à l'occasion de cette étude puissent être conservées dans une base de données et faire l'objet d'un traitement informatisé non nominatif. J'ai bien noté que le droit d'accès prévu par la loi « informatique et libertés » s'exerce à tout moment auprès des expérimentateurs.

Je consens à participer à cette étude réalisée au sein du Laboratoire d'Anthropologie et de Psychologie Cognitive et Sociale (LAPCOS) de l'université de Nice Sophia-Antipolis, sous la supervision du Pr. Faure Sylvane et du Dr. Spotorno Sara.

Date :

Nom du volontaire :

Signature du volontaire (précédée de la mention « lu et approuvé ») :

Nom de l'expérimentateur:

Signature de l'expérimentateur :

Nom de l'expérience :

Annexe 15 : Questionnaire latéralité et renseignements adulte

- Participant n°
- Age:
- Sexe: Femme Homme
- Œil directeur :
- Œil testé :

Indiquez vos préférences dans l'utilisation de vos mains pour les activités suivantes en inscrivant une croix « X » dans la colonne appropriée. Lorsque la préférence est tellement forte que vous ne tenteriez jamais d'utiliser l'autre main à moins d'y être contraint, inscrivez deux croix « XX ». Dans le cas où vous n'avez pas de préférence, lorsque la main vous est indifférente, inscrivez une croix dans chaque colonne.

Certaines activités requièrent les deux mains. Dans ce cas, la partie de la tâche, ou de l'objet pour laquelle la préférence vous est demandée est inscrite entre parenthèses.

Merci d'essayer de répondre à toutes les questions et de ne laisser un blanc que lorsque vous n'avez aucune expérience de l'objet ou de la tâche.

N°	Tâche	GAUCHE	DROITE
1	Ecrire		
2	Dessiner		
3	Lancer		
4	Utiliser une paire de ciseaux		
5	Brosse à dents		
6	Couteau (sans fourchette)		
7	Cuillère		
8	Balai (la main du haut)		
9	Allumer une allumette (main qui tient l'allumette)		
10	Ouvrir une boîte (couvercle)		
1	Avec quel pied préférez-vous taper ?		
2	Quel œil utilisez-vous lorsque vous n'en utilisez qu'un seul ?		

Informations complémentaires (entourer la réponse)

- Etes-vous droitier ou gaucher? Droitier Gaucher
- Y a-t-il au sein de votre famille directe (père, mère, frères, sœurs) des personnes gauchères ou ambidextres ? Oui Non

Si oui, qui ?.....

- Etes-vous de langue maternelle française ? Oui Non
- Avez-vous des problèmes de vision ? Oui Non

Si oui, quel est le type de correction ?.....

- Avez-vous déjà subi un traumatisme crânien ou toute autre atteinte cérébrale ?

Oui Non

- Si oui, leurs éventuelles conséquences vous gênent-elles encore aujourd'hui ?

Oui Non

- Etes-vous actuellement sous traitement antidépresseur ou anxiolytique ?

Oui Non