

MOVING TOWARDS THE UNDERSTANDING OF DYNAMIC HUMAN FACE PROCESSING

Doctoral Thesis

Presented at the Faculty of Letters of the University of Fribourg, Switzerland
Submitted for the degree of Doctor in Philosophy
Specialized in Cognitive Neuroscience

Presented by
Anne-Raphaëlle RICHÖZ
(from Ursy, Switzerland)

Approved by the Faculty of Letters on the proposal of Prof. Dr. Roberto CALDARA
(first examiner), Prof. Dr. Olivier PASCALIS (second examiner) and Prof. Dr. Bruno
ROSSION (third examiner)

Fribourg, 12th January 2018
Prof. Dr. Bernadette CHARLIER (Dean of the Faculty)

This thesis was directed by Prof. Dr. Roberto CALDARA and co-directed by Prof. Dr. Olivier PASCALIS as part of a joint supervision between the University of Fribourg in Switzerland and the University of Grenoble-Alpes in France.

The University of Fribourg and a grant F14/06 from the Rectors' Conference of Swiss Universities (CRUS) supported this thesis.

Thesis examination committee:

Prof. Dr. Roberto **CALDARA**, University of Fribourg, thesis director, examiner

Prof. Dr. Olivier **PASCALIS**, University of Grenoble-Alpes, thesis co-director, examiner

Prof. Dr. Bruno **ROSSION**, University of Louvain, examiner

Prof. Dr. Valérie **CAMOS**, University of Fribourg, President of the examination committee

Prof. Dr. Chantal **MARTIN-SOELCH**, University of Fribourg, assessor

Prof. Dr. Nicolas **RUFFIEUX**, University of Fribourg, assessor

Fribourg, 12th January 2018



swissuniversities



Abstract

The human visual system is steadily stimulated by dynamic cues. Faces provide crucial information about identity, gender, ethnicity, speech, attention, eye-gaze direction, or emotions, all important for adapted social interactions. From an evolutionary perspective, humans have been more extensively exposed to dynamic faces, as *static* face images have only appeared recently with the advent of photography and the expansion of digital tools and social networks. Yet, most studies investigating face perception have relied on *static* faces and only a little is known about the mechanisms involved in dynamic face processing.

To clarify this issue, this thesis aimed to use *dynamic* faces to investigate different aspects of face processing in different populations and age groups. In Study 1, we used *dynamic* faces to investigate whether the ability of infants aged 6, 9, and 12 months in matching audible and visible attributes of gender is influenced by the use of adult-directed (ADS) vs. infant-directed (IDS) speech. Our results revealed that from 6 months of age, infants matched female but not male faces and voices when presented with adult-directed speech. This ability emerged at 9 months of age when presented with infant-directed speech. Altogether, these findings support the idea that the perception of multisensory gender coherence is influenced by the very nature of social interactions.

In Study 2, we used a novel 4D technique to reconstruct the *dynamic* internal representations of the six basic expressions in a *pure* case of acquired prosopagnosia (i.e., a brain-damaged patient severely impaired in recognizing familiar faces). This was done in order to re-examine the debated issue of whether identity and expression are processed independently. Our results revealed that patient PS used all facial features to represent basic expressions, contrasting sharply with her suboptimal use of facial information for identity recognition. These findings support the idea that different sets of representations underlie the processing of identity and expression. We then examined patient PS's ability to recognize *static* and *dynamic* expressions using her internal representations as stimuli. Our results revealed that she was selectively impaired in recognizing many of the *static* expressions; an impairment that could be explained by her specific set of lesions in the right inferior occipital gyrus. In contrast, she displayed maximum accuracy in recognizing all the *dynamic* emotions with the exception of fear. The latter findings support recent evidence suggesting that separate cortical pathways, originating in early visual areas and not in the inferior occipital gyrus, are responsible for the processing of *static* and *dynamic* face information. Altogether, our findings suggest that the temporal properties of human facial expressions influence their processing, particularly in fragile face processing systems, questioning the evidence obtained by neuropsychological studies investigating expression recognition with *static* images only.

Moving on from our second study, in Study 3, we investigated whether *dynamic* cues offer processing benefits for the recognition of facial expressions in other populations with fragile face processing systems; for instance, young children whose systems are yet to fully mature or elderly people whose systems are declining. To this aim, we conducted a large sample cross-sectional study with more than 400 participants aged between 5 to 96 years, investigating their ability to recognize the six basic expressions presented under different temporal conditions. Consistent with previous studies, our findings revealed the highest recognition performance for happiness, regardless of age and experimental condition, as well as marked confusions among expressions with perceptually similar facial signals (e.g., fear and surprise). By using Bayesian modelling, our results further enabled us to quantify, for each expression and condition individually, the steepness of increase and decrease in recognition performance, as well as the peak efficiency, the point at which observers' performance reaches its maximum before declining. Finally, our results offered new evidence for a *dynamic* advantage for facial expression recognition, stronger for some expressions than others and more important at specific points in development.

Overall, the results highlighted in this thesis underlie the critical importance of research featuring *dynamic* stimuli in face perception and expression recognition studies; not only in the field of prosopagnosia, but also in other domains of developmental and clinical neuroscience.

Résumé

Les visages animés que nous voyons quotidiennement fournissent des informations cruciales sur les expressions émotionnelles d'autrui, la direction de l'attention ou du regard, toutes déterminantes pour des interactions sociales adaptées. Au cours des dernières décennies, la plupart des études investiguant la reconnaissance des visages ont utilisé des photographies *statiques*. Or dans notre environnement naturel, les visages et leurs expressions émotionnelles sont des phénomènes *dynamiques* qu'il est difficile de communiquer écologiquement avec des images *statiques*. Les visages *dynamiques* auxquels nous sommes constamment exposés sont plus riches et plus réalistes et ils permettent de transmettre plus finement et naturellement les sentiments, émotions et intentions d'autrui. Cette exposition quotidienne et répétée à des visages en mouvement pourrait-elle avoir un effet sur notre système visuel, favorisant le traitement de stimuli *dynamiques* au détriment des *statiques* ?

Afin d'éclairer cette problématique, les recherches présentées dans cette thèse avaient pour but d'utiliser des stimuli *dynamiques* pour étudier différents aspects du traitement des visages à travers plusieurs groupes d'âge et populations. Dans notre première recherche, nous avons utilisé des visages animés pour voir si la capacité de nourrissons âgés de 6-, 9- et 12 mois à associer des attributs audibles et visibles à un genre est influencée par l'utilisation d'un discours de type adulte (ADS) par opposition à un langage de type enfantin (IDS). Nos résultats ont montré qu'à partir de 6 mois, lorsqu'ils étaient soumis à un discours de type adulte, les nourrissons associaient les voix et visages de femmes, mais pas d'hommes. Par contre, lorsqu'ils étaient confrontés à un langage de type enfantin, cette capacité apparaissait seulement à l'âge de 9 mois. Ces premiers résultats soutiennent l'idée selon laquelle le développement de la perception multisensorielle chez les nourrissons est influencé par la nature même des interactions sociales.

Dans notre deuxième recherche, nous avons utilisé une nouvelle technique 4D pour reconstruire les représentations mentales des six émotions de base d'une patiente présentant un cas unique et *pure* de prosopagnosie acquise (i.e., une incapacité à reconnaître les visages), afin de réexaminer une question bien débattue, à savoir si les modules cérébraux sous-jacents à la reconnaissance de l'identité et des expressions faciales sont séparés ou communs. Les résultats ont montré que notre patiente a utilisé toutes les caractéristiques faciales pour identifier les émotions de base, ce qui contraste fortement avec son utilisation déficitaire de l'information faciale pour la reconnaissance de l'identité. Ces résultats confortent l'idée selon laquelle différents systèmes de représentations sous-tendent le traitement de l'identité et de l'expression. Par la suite, nous avons pu démontrer que notre patiente était capable de reconnaître adéquatement les expressions émotionnelles *dynamiques*, mais pas les émotions *statiques* provenant de ses propres représentations internes. Ces résultats qui pourraient être expliqués par un ensemble spécifique de lésions dans son gyrus occipital inférieur droit, soutiennent l'idée selon laquelle le traitement des stimuli *statiques* et *dynamiques* se produit dans des régions cérébrales différentes. Finalement, notre étude souligne l'importance de l'utilisation de stimuli *dynamiques* dans la recherche, notamment chez des patients cérébro-lésés, et questionne les résultats obtenus par d'autres travaux étudiant la reconnaissance des expressions uniquement avec des stimuli *statiques*.

Dans notre troisième recherche, nous avons investigué si d'autres populations ayant un système visuel neurologiquement fragile ou en développement bénéficient également de la présentation d'expressions *dynamiques*. Nous avons demandé à plus de 400 sujets âgés de 5 à 96 ans de catégoriser les six expressions de base en versions *statique*, *dynamique* ou bruitée. En utilisant un modèle Bayésien, nos résultats nous ont permis de quantifier la pente d'amélioration et de déclin pour chaque expression dans chaque condition, ainsi que d'estimer l'âge auquel l'efficacité est maximale. En résumé, nos résultats montrent la supériorité des stimuli *dynamiques* dans la reconnaissance des expressions faciales, de manière plus marquée pour certaines expressions que d'autres et de façon plus importante à certains moments spécifiques du développement.

Dans l'ensemble, les résultats de cette thèse soulignent l'importance d'investiguer la reconnaissance des visages avec des stimuli *dynamiques*, non seulement en neuropsychologie, mais aussi dans d'autres domaines des neurosciences développementales et cliniques.

*Someone once asked me: Why do you love music so much? I replied: Because it's the only thing that
stays when everything and everyone is gone.*
Unknown

Avoir envie de réaliser un rêve, c'est le talent. Le reste, c'est de la sueur.
Jacques Brel

Acknowledgements

Four years ago, as I started my PhD at the University of Fribourg, I did not know what to expect. When asking people around me about their own experiences, the most frequent answer that came up was the following: a solitary and really hard journey; a tiring and emotionally draining commitment that requires one to read innumerable articles, write papers and reviews, program codes to carry out experiments and analyze data, try out new ideas and prepare talks and posters.

Looking back, I agree that doing a PhD is indeed very hard. However, my PhD was not a solitary track. I had the chance to meet countless outstanding people, make new friends all around the world, and attend conferences in places I had never been before. It left marks on me, changed my perception and even my mind. My PhD journey was long, but a journey that was worth the effort as I gained invaluable experiences and enjoyed (nearly) every single moment of it. My warmest thanks go to all the people who have accompanied me throughout this adventure. Professors, colleagues, family members and friends that have helped me and supported me directly or indirectly in the completion of my work.

My deepest gratitude goes first and foremost to my supervisor Prof. Roberto Caldara for giving me the opportunity to do a PhD in his research group at the University of Fribourg. I would like to thank you sincerely for having taken me under your wings when I was a Masters student, guiding me on the right path with patience, wisdom and humor throughout these years. You inspired me to work hard every day, to believe in my strengths and skills, to grow and broaden my horizons. Thank you so much for all you have done, for your advice, support and encouragements. I treasure every discussion we had and everything I have learnt from you. Your enthusiasm and commitment for research motivated me to look forward to working every day.

I would also like to express my sincere thanks to my co-supervisor Prof. Olivier Pascalis for welcoming me in his research laboratory at the University of Grenoble four years ago. Thank you for your time, your support, your encouragements, and your constructive comments. My warmest thanks go to all the people I had the chance to meet in your team: Anne, Carole, Chloé, David, Eve, Fabrice, Guillaume Hélène, Laurie, Marjorie, and all the others. A special thankful thought goes to Marie for her precious assistance with the coding of the videos. I will never forget the interesting discussions, your jokes and laughter and the good times spent in Grenoble.

My sincere gratitude goes to all members of my dissertation committee: Prof. Bruno Rossion, Prof. Daniel Fiset, Prof. Valérie Camos, Prof. Chantal Martin-Sölch, and Prof. Nicolas

Ruffieux. I am honored and grateful to have you all in my examination committee. Thank you for your time and comments.

I would like to thank my colleagues and friends Charles (Junpeng) and Nayla for their time, help, teaching guidance and suggestions with the statistical questions. I owe you so much. Many thanks, Charles, for your commitment and dedication to Bayesian analysis and the great time we spent in ABIM three years ago. I am also grateful to all my other present and former colleagues and friends of the iBMLab: Lisa, Marie-Paule, Meike, Michaël, Peter, Sébastien, Valentina, Xinyi, Yindi, with a special thought to my supportive and unforgettable office mates Caroline, Helen and Nicolas for their daily conversations and help. Work would not have been the same without you! Thanks also to my friend and colleague Maria-Teresa for welcoming me in Florence two years ago and for helping out with testing.

Although my name stands alone on the cover of this thesis, my work would not have been possible without the guidance and help of many collaborators. I take this opportunity to express my heartfelt gratitude to Junpeng Lao, Rachel Jack, Olivier Garrod, Philippe Schyns, Paul Quinn, Anne Hillairet de Boisferon, Carole Berger, Hélène Loevenbruck, David Lewkowicz, Kang Lee and Marjorie Dole. I am also grateful to all the anonymous reviewers for their helpful and insightful comments.

My sincere gratitude goes to Pierrette for taking part in one of my studies. This thesis would not exist without your motivated contribution. I would like to thank you also for all the beautiful moments we shared together, train rides, discussions, and the good (and the less good) restaurants that we had the chance to discover. I also extend my gratitude to all the other 582 participants involved in one of the three experimental contributions of this thesis: infants, children, adolescents, students, adults and elderly people. I could not have completed my thesis without your precious participation.

I would like to thank all my students at the University of Fribourg for their challenging questions, comments, suggestions, and for giving me the chance to get better. Special thanks go to Pauline, Linda, Lauriane, Benedetta, Christel, Martina, Adrien and Serena, for their help, motivation and commitment to research. I am also grateful to Andrée Girard, Claudia Vonlanthen, Laurence Pitton, Madeleine Olachea and Claudia Dietrich for their continuous help with the administrative questions. Thanks to the University of Fribourg and the grant F14/06 from the Rectors' Conference of Swiss Universities for funding my PhD studies.

Very special thanks go to my parents, Jean-Claude and Annika. I am sincerely grateful to my father, for reading my work, providing insightful and constructive comments, and helping me improve my writing skills. My deep sense of gratitude also goes to my mother, for her

creative thinking and endless imagination. Thank you both for your kindness, love, encouragements and unconditional emotional support. You pushed me when it was necessary to be pushed, helped me to go rapidly through unproductive and difficult periods, brought me to the mountains when some fresh air and new ideas were needed. You have transmitted me your passion for teaching and your intellectual curiosity.

I would like to thank my sisters who were always present to suggest some help with presentations, teaching, writing, also providing emotional support and ideas of activities to distract my mind when necessary. A very special thought also goes to my grandfather who has always encouraged and supported me. Thanks to my dearest friends Aubry, Futura, Jutta, Lauriane, Lisa and Sylvain for their support and interesting discussions.

In addition to those mentioned above, several other people have contributed to my thesis, providing help, motivation and encouragement. Unfortunately, there is not enough space to mention everyone here. Nevertheless, I still have a special thought for a few of them who have actively contributed to my work: Adèle, Brenda, Christopher, Jason, Franck, Michel, O.S.F., Jo, François, Isabelle, Chloé, Paul, Louis, Daniel, Dounia, J.M.F., Rosetta, Jean, Mouche and his brother. Thank you for being so unique. I had the chance to get to know you over my PhD studies and will never forget you.

Finally, my warmest gratitude goes to Music, with a special note of thanks to the following musicians who played a very important role during these four years of hard work: Bach, Beethoven, Chopin, Debussy, Händel, Haydn, Liszt, Mendelssohn, Mozart, Puccini, Rachmaninoff, Saint-Saëns, Schubert, Schumann, Tchaikovsky and Verdi (writing and working playlists), and to Avidan, Battiato, Bécaud, Brel, Callas, Carboni, Christine and The Queens, Christophe, Coldplay, De Gregori, Fauve, Ferrat, Florence and the Machine, Goldman, Grand Corps Malade, Hardy, Pavarotti, Pravo, Taha, Sardou, The Weeknd and Rossi, (references and motivation playlists) – to name a few among all the others. Your master pieces accompanied me every single day. When inspiration failed, music spoke.

Table of Contents

Abstract	iii
Résumé	v
Acknowledgements	ix
Table of Contents.....	xiii
List of Figures	xvii
List of Tables.....	xx
List of Boxes.....	xxi
Introduction and thesis overview.....	2
CHAPTER 1 Literature review.....	6
1.1 Face Processing.....	8
1.1.1 Is face processing special?.....	8
1.1.2 The early development of face processing.....	11
1.1.3 The role of experience in face processing.....	16
1.1.4 Gender processing.....	19
1.1.5 Faces as a multiple source of information	21
1.2 Facial expressions of emotion.....	24
1.2.1 Origins and categories of emotions.....	24
1.2.2 Transmitting and decoding facial expressions	26
1.2.3 Communicating emotions: the activation of different facial muscles	29
1.2.4 The development of expression recognition: a lifespan perspective.....	32
1.2.5 Theoretical explanations for age-related expression recognition deficits.....	40
1.2.6 Adding motion to emotion.....	45
1.3 Face processing in the brain.....	53
1.3.1 A distributed neural network.....	53
1.3.2 Prominent models of face processing.....	62
1.3.3 A critical review of early influential models of face perception.....	65
1.3.4 The processing of static and dynamic faces.....	76
CHAPTER 2 – Introduction to the experimental contribution.....	82
2.1 A developmental study.....	83
2.2 A neuropsychological study.....	86
2.3 A cross-sectional study.....	89
CHAPTER 3 – Audio-visual perception of gender by infants emerges earlier for ADS.....	92
3.1 Abstract.....	93
3.2 Introduction	94

3.3 Materials and method	97
3.4 Results	101
3.5 Discussion	102
3.6 Conclusions	107
3.7 Acknowledgements	108
CHAPTER 4 – Facial expression recognition and prosopagnosia	110
4.1 Abstract	111
4.2 Introduction	112
4.3 Participants	115
4.4 Experiments	117
4.4.1 Preliminary experiment	117
4.4.2 Experiment 1	118
4.4.3 Experiment 2	124
4.4.4 Experiment 3	126
4.5 General discussion	128
4.5.1 Does the face system rely on common or distinct representations?	129
4.5.2 Advantage in decoding dynamic facial expression in prosopagnosia	131
4.5.3 The special case of “fear”	133
4.6 Conclusions	134
4.7 Acknowledgements	135
4.8 Supplementary movies	136
CHAPTER 5 – Expression recognition across the life span	140
5.1 Abstract	141
5.2 Introduction	142
5.3 Material and methodes	146
5.3.1 Participants	146
5.3.2 Stimuli	147
5.3.3 Procedure	149
5.3.4 Data Analysis	150
5.4 Results	154
5.5 Discussion	161
5.5.1 Recognition trajectory across development	161
5.5.2 Static versus dynamic expressions	167
5.6 Conclusions	173
5.7 Acknowledgements	175
5.8 Supplementary figures	176
5.9 Supplementary results	182

CHAPTER 6 – General discussion.....	184
6.1 Main findings.....	185
6.2 Discussion	187
6.2.1 A key involvement of dorsal face-selective brain regions.....	187
6.2.2 Our findings in the light of revised frameworks	188
6.3 Limitations.....	190
6.3.1 Limitations to our first experimental contribution	190
6.3.2 Limitations to our second experimental contribution.....	190
6.3.3 Limitations to our third experimental contribution.....	192
6.4 Future directions	193
6.5 Conclusions.....	196
References.....	198
Curriculum Vitae.....	248
Publication List	252

List of Figures

Figure 1.1	The face-composite paradigm
Figure 1.2	The whole-face paradigm
Figure 1.3	Sampling visual information reveals culture-specific perceptual strategies
Figure 1.4	Infants' visual preference for faces
Figure 1.5	Prenatal preference for face-like visual patterns
Figure 1.6	Experimental paradigms used with preverbal infants
Figure 1.7	The role of experience in the development of face processing abilities
Figure 1.8	The different stimuli used to investigate perceptual narrowing during infancy
Figure 1.9	Gender processing in infants
Figure 1.10	Developmental change in categorization abilities
Figure 1.11	Multisensory integration of faces and voices
Figure 1.12	Multisensory integration of faces and voices is restricted to female faces
Figure 1.13	Transmitting and decoding facial expressions of emotion
Figure 1.14	Bubbles
Figure 1.15	Facial expression signals are weakly correlated with one another
Figure 1.16	AUs for disgust
Figure 1.17	Muscles for facial expressions
Figure 1.18	fNIRS procedure in developmental studies
Figure 1.19	Differences in hemodynamic responses to happy and angry faces evidenced by a near-infrared spectroscopic study
Figure 1.20	Fine-grained developmental mapping of emotional face processing
Figure 1.21	Neurobiological changes in the amygdala throughout development
Figure 1.22	Affective influences in shaping expression recognition
Figure 1.23	Positivity theory
Figure 1.24	Eye movement patterns in young and elderly people
Figure 1.25	Brain activation to fearful faces in young and old adults
Figure 1.26	Point-light static displays
Figure 1.27	Investigating the recognition of subtle expressions with static and dynamic stimuli
Figure 1.28	Neuropsychological evidence for a dynamic advantage
Figure 1.29	Three face-selective patches in the ventral prefrontal and temporal cortex of macaque monkeys
Figure 1.30	A comparison of face-selective clusters in macaque and human brains
Figure 1.31	Five single-cases of acquired prosopagnosia, LR, NS, GG, GD, and PS and their corresponding brain lesions
Figure 1.32	Lesion localizations

Figure 1.33	Numerous face-selective clusters in individual brains
Figure 1.34	Face-selective brain regions
Figure 1.35	Electrical brain stimulations applied to face-selective regions in the right- and left-sided fusiform gyrus
Figure 1.36	Bruce and Young's cognitive model of face processing
Figure 1.37	The three core regions of the face processing system
Figure 1.38	The influential neural model of human face processing proposed by Haxby and colleagues in 2000
Figure 1.39	The processing of invariant vs. changeable aspects of faces in dorsal and ventral streams
Figure 1.40	Distinct cortical pathways for static and dynamic face information
Figure 1.41	The size of the TBS disruption effect in the right pSTS, rOFA and rFFA for static and dynamic faces and objects
Figure 1.42	A schematic representation of lesioned and preserved face-selective brain regions in four patients
Figure 1.43	The reverse hierarchical view of face perception proposed by Rossion (2008; 2014)
Figure 1.44	Face-selective neural activation in the rOFA and rFFA to upright and inverted Monney faces and Arcimboldo face-like paintings
Figure 1.45	Differential neural activation to faces and cars in the rFFA and rOFA
Figure 1.46	Facial expression recognition occurs simultaneously in the rOFA and pSTS as evidenced by a TMS study (Pitcher, 2014)
Figure 1.47	The FFA is involved in the processing of emotional faces: Evidence from an fMRI adaption study
Figure 1.48	The revised neural framework of the face processing system proposed by Duchaine and Yovel (2015)
Figure 1.49	Percentage of signal change to dynamic and static faces, bodies, scenes, objects and scrambled faces in the rFFA, rOFA, right pSTS and right aSTS
Figure 1.50	Cluster size with static and dynamic localizers
Figure 1.51	Static and dynamic functional localizers
Figure 2.1	Infants preference for infant-directed speech (IDS)
Figure 2.2	IDS conveys cues for social selection
Figure 2.3	FACS synthesis
Figure 2.4	The procedure used by Yu et al. (2012) to reconstruct the internal representation of the six basic expressions in human observers
Figure 2.5	The recognition of static and dynamic expressions using an ideal observer approach
Figure 3.1	Schematic representation of the procedure
Figure 3.2	Mean proportion of total looking time to the face that matched the heard voice in each condition (IDS, ADS) and for each age group (6-, 9-, 12-month-olds)
Figure 4.1	FACS-based generator of random facial muscle activations

Figure 4.2	PS's and controls' reconstructed mental models of happiness, surprise, fear, disgust, anger, and sadness
Figure 4.3	Deviation maps of PS's and controls' reconstructed mental models of happiness, surprise, fear, disgust, anger, and sadness
Figure 4.4	Assessing categorization of facial expressions of emotion with PS's reconstructed mental models
Figure 4.5	Average static reconstructed mental model for the expressions of happiness, surprise, fear, disgust, anger, and sadness
Figure 5.1	Examples of the three different kinds of stimuli used in our study
Figure 5.2	Schematic representation of the procedure
Figure 5.3	A conceptual representation of the step linear model
Figure 5.4	Accuracy across age groups for each expression in the three different conditions
Figure 5.5	Anger
Figure 5.6	Disgust
Figure 5.7	Fear
Figure 5.8	Happiness
Figure 5.9	Sadness
Figure 5.10	Surprise
Figure 5.11	Confusion matrices – Response classification errors
Figure 6.1	Suboptimal facial information use in two brain-damaged patients

List of Tables

Table 1.1	The basic emotions
Table 1.2	Main studies investigating facial expression recognition across life span
Table 1.3	Differences in brain activation between young and old adults during the processing of fearful faces
Table 1.4	Main studies comparing the recognition of static and dynamic expressions
Table 1.5	Neuroimaging studies comparing the processing of static and dynamic facial expressions
Table 3.1	Fundamental frequency characteristics in Hertz (Hz)
Table 3.2	Fundamental frequency characteristics in Semitones (st)
Table 4.1	Recognition accuracy of PS and age-matched controls on Ekman and Friesen's facial expression recognition test
Table 4.2	PS's AUs preferences fell within the normal range of controls
Table 4.3	Percentage accuracy score of PS and age-matched controls for PS's static and dynamic reconstructed mental models
Table 4.4	Percentage accuracy score of PS and age-matched controls for average static and dynamic reconstructed mental models
Table 5.1	Recognition performance for surprise between the dynamic and static conditions at peak efficiency and above the age of 80

List of Boxes

Box 1 – Key terminology

Box 2 – Face-composite paradigm

Box 3 – Whole-face paradigm

Box 4 – Flexible face processing strategies

Box 5 – Experimental paradigms used in infants' studies

Box 6 – Bubbles

Box 7 – Three cases of acquired prosopagnosia – Case descriptions

Box 8 – Three cases of acquired prosopagnosia – Lesion localizations

What is more important than reading books? Reading people's faces.
Anis Mansour

Introduction and thesis overview

*The most interesting aspect of other people,
the point where we go inside them, is the face.
It tells all.*

David Hockney

“It happens almost every day. When I pick up my children from school, when I drink a coffee at the coffee shop next to my office, when I go shopping... Suddenly someone bumps into me, a woman, a man or a group of people. Obviously, they know me, because they call me by my name. We exchange a few words, they laugh, I wince, embarrassed. Usually it takes me only a few seconds to end the conversation. Why? Mainly because I don’t know who on earth they are. Are they close friends or more distant acquaintances? Colleagues? Relatives? Neighbors? I don’t have the slightest idea. For me being impaired in recognizing the faces of people around me, of public figures, or even of my own children, is like not possessing the adapted screwdriver to unscrew the resistant bolts that open the doors to identity. My universe is filled with thousands of people, thousands of people with no faces nor identities” (ARC7/7).

These were the words reported to me a few months ago, by a patient severely impaired in recognizing faces. A face is not only the gateway to identity, but as rightly pointed out by the British painter David Hockney, it is probably one of the richest and most fascinating stimuli we encounter in our daily lives. Faces are a wide variety of powerful information sources that are critical for survival and adapted social interactions. For instance, lip motion facilitates speech perception, gaze direction provides information about attention and interest, hair and skin texture convey cues for gender and race categorization, whereas facial expressions communicate the internal states of other individuals. From an evolutionary perspective, it seems reasonable to assume that our brain is tuned to rapidly achieve categorization of these different facial features and detect subtle changes in the emotional states of our conspecifics. Although some previous reports have evidenced that the human visual system displays a certain sensitivity to facial motion – for instance for expression and identity recognition (Ambadar, Schooler, & Cohn, 2005; Bould, Morris, & Wink, 2008; Butcher & Lander, 2017; Lander & Butcher, 2015; O’Toole, Roark, & Abdi, 2002), the majority of studies investigating face processing have overlooked the importance of this factor.

To clarify this issue, this thesis aims to use *dynamic* faces to investigate different aspects of face processing in different populations and age groups. More precisely, this work examines three distinct questions:

- 1) Is the extraction of gender from *dynamic* faces and voices influenced by the nature of speech signal in 6-, 9-, and 12-month-old infants?
- 2) Are the *dynamic* internal representations of the six basic expressions in a *pure* case of acquired prosopagnosia (i.e., a brain-damaged patient severely impaired in recognizing familiar faces) similar to those of healthy controls? Further, are these *dynamic* expressions better recognized by patient PS than *static* images?
- 3) Besides prosopagnosic individuals, do other populations with immature or fragile face processing systems also benefit from the presentation of *dynamic* emotional signals?

In order to answer these questions, this thesis is divided into six chapters, each including different sections.

CHAPTER 1 – Literature review. The first chapter provides an overview of the literature related to face processing. After a general introduction outlining the reasons why human faces are special, I will first review some developmental studies revealing how face processing develops throughout infancy. The second section gives an overview of some aspects of facial expressions of emotion, how they are perceived throughout development, and whether temporal dynamics facilitate their recognition. Finally, the last section of the first chapter offers a general overview of the neuroanatomical bases of human face processing based on studies from the neuropsychological and functional neuroimaging literature.

CHAPTER 2 – Experimental contribution. The second chapter introduces the experimental contribution of this thesis and provides an overview of some methodological aspects.

CHAPTER 3 – A developmental research. This chapter presents the first study of this thesis investigating the multisensory perception of gender in 6- to 12-month-old infants. In this study, conducted in collaboration with colleagues from the University of Grenoble-Alpes, we used dynamic faces to assess whether the extraction of gender from faces and voices is influenced by the very nature of speech signal, namely by infant- or adult-directed speech.

CHAPTER 4 – A neuropsychological research. The second experimental contribution of this thesis, presented in Chapter 4, is a neuropsychological research conducted in collaboration with researchers from the University of Glasgow. This study re-examines the widely debated question regarding whether distinct or common face representational systems underlie the processing of identity and facial expressions. In order to examine the existence of a dichotomy

between identity and expression recognition, we reconstructed the dynamic mental models of the six basic expressions in a *pure* case of acquired prosopagnosia. In a follow-up experiment, we assessed patient PS's recognition abilities of the reconstructed dynamic and static expressions.

CHAPTER 5 - A cross-sectional research. The last experimental contribution conducted at the University of Fribourg is a cross-sectional study. In this research presented in Chapter 5, we aim to investigate whether the dynamic advantage for facial expression recognition that we found in our prosopagnosic patient (CHAPTER 4) extends to other populations with fragile face processing systems, such as very young children or elderly people.

CHAPTER 6 - General discussion. In the last chapter, I provide a brief and concise summary of the work done and discuss the main findings, their implications, limits, and perspectives.

CHAPTER 1

LITERATURE REVIEW

Box 1 – Key Terminology

Face-selective brain regions

OFA:	Occipital Face Area located in the inferior occipital gyrus
FFA:	Fusiform Face Area located in the middle fusiform gyrus
antFUS:	Anterior part of the fusiform gyrus
pFUS:	Posterior part of the fusiform gyrus
STS:	Superior Temporal Sulcus
pSTS:	Posterior part of the Superior Temporal Sulcus
aSTS:	Anterior part of the Superior Temporal Sulcus
ATL:	Anterior Temporal Lobe
OFC:	Orbitofrontal Cortex
IT:	Infero-Temporal Cortex
PO:	Prefrontal Orbital Cortex
PL:	Prefrontal Lateral Cortex
PA:	Prefrontal Arcuate
AMG:	Amygdala
PG:	Parahippocampal Gyrus

Brain-recording techniques

PET:	Positron Emission Tomography
fMRI:	Functional Magnetic Resonance Imaging
fMRIa:	Functional Magnetic Resonance Imaging adaptation studies
DTI:	Diffusion Tensor Imaging
ERP:	Event-Related Potentials
TMS:	Transcranial Magnetic Stimulation
TBS:	Thetaburst Transcranial Magnetic Stimulation

Optogenetic: A recent technique that combines optical and genetic methods to optically control the electrical activity of excitable neurons or cells (Fenno, Yizhar, & Deisseroth, 2011; Kalanithi & Henderson, 2012).

Optical imaging: A technique that measures simultaneously the activity of a group of neurons in a specific cortical region by detecting changes in blood oxygenation triggered by neural activity (Wang, Tanifuji, & Tanaka, 1998).

fNIRS: A non-invasive neuroimaging technique that allows the localization of brain activity in specific cortical regions by monitoring relative changes in oxy-, deoxy-, and total-haemoglobin concentrations (Nakato et al., 2011).

1.1 FACE PROCESSING

1.1.1 Is face processing special?

Human faces are one of the richest and most powerful tools among social stimuli, playing a fundamental role in even the most basic human social interactions. Observers can reliably and rapidly derive a wide variety of information from faces such as identity, sex, race, gender, sexual orientation, attractiveness, speech patterns, direction of attention, emotional states, physical health, and even personality traits such as trustworthiness, competence dominance, or likeability. Our ability to perform relevant categorizations of these different facial features is a highly developed human skill, strongly encoded in our biology and critical for survival. This powerful expertise has raised a central question, guiding the work of thousands of studies over the last century: Is face processing special? Does this ability involve specific strategies and functionally and anatomically discrete brain regions?

Historically, several studies have provided evidence supporting the claim that face processing is special. First, developmental studies have revealed that within minutes after birth, neonates will preferentially look at face-like geometrical stimuli and discriminate faces in a remarkable manner. Morton and Johnson (1991) have hypothesized that newborns possess an innate mechanism, selectively tuned to process face patterns. This very early responsiveness to faces supports the view that faces are special, because such early preference does not seem to be present for the processing of non-face objects (McKone & Robbins, 2011; but see Arcaro et al., 2017).

Second, behavioral studies in adults relying on specific experimental paradigms, such as the face-composite paradigm (Young, Hellawell, & Hay, 1987) (for a review, see Rossion, 2013, see [Box 2](#)) or the whole-face paradigm (Farah et al., 1998; Tanaka & Farah, 1993, see [Box 3](#)), have revealed that the processing of faces elicits specific perceptual strategies, more holistic/configural (Farah et al., 1998; Rossion, 2013; Sargent, 1984; Tanaka & Farah, 1993) than the strategies used for other objects (Maurer, Le Grand, & Mondloch, 2002; Rossion, 2008, but see [Box 4](#)). As mentioned by Rossion (2008), holistic or configural processing strategies imply the simultaneous combination of the multiple distinctive features of a face (e.g., eyebrows, nose, mouth, eyes) into a unique representation. Three different types of configural processing strategies are used to recognize faces (Maurer et al., 2002). The first one considers information regarding the features of the face (i.e., two eyes located above a mouth). The second one relies on the space between the different features (second-order configural relations), whereas the last one focuses on the global structure of the face (holistic processing). Experimental tasks that disrupt configural

processing, such as inverting images (the face inversion effect, Rossion & Gauthier, 2002; Yin, 1969), lead to a poorer recognition performance for faces than any other classes of objects. This latter evidence supports the assumption that the processing of faces entails specific perceptual strategies compared to non-face objects.

In addition to behavioral evidence, electrophysiological studies in primates (e.g., Hasselmo, Rolls, & Baylis, 1989; Tsao et al., 2003; Tsao et al., 2008b), neuroimaging studies in adults (e.g., Gauthier et al., 2000b; Haxby et al., 1999; Hoffman & Haxby, 2000; Kanwisher, McDermott, & Chun, 1997; Sergent, Ohta, & MacDonald, 1992), as well as neuropsychological studies in brain-damaged patients (Busigny et al., 2010a; Dalrymple et al., 2011; Rossion, 2014; Rossion et al., 2003) have also provided strong evidence in support of the claim that faces are special and that a particular neural circuitry, characterized by structural and neural specialization, is dedicated to their processing. The following sections of this chapter will overview some of these studies, starting with developmental evidence.

Box 2 – The Face-Composite Paradigm

The face-composite paradigm is based on the visual illusion that two physically identical top halves of a face are erroneously perceived as being different when aligned with different bottom halves. According to Rossion (2013), this paradigm provides strong evidence that human faces are perceived as integrated wholes rather than a collection of individual discrete features and thus processed in a holistic way. Interestingly, the composite face illusion disappears when faces are inverted (Figure 1.1). When presented upside-down, two identical top halves are no longer perceived as being different. That is due to a reduction of the perceptual field, leading to a disruption of configural processing strategies (Rossion, 2013). In other words, when inverted, top halves are no longer influenced by bottom halves, as the individual parts of a specific face are processed independently, in an analytical way (Rossion, 2013).

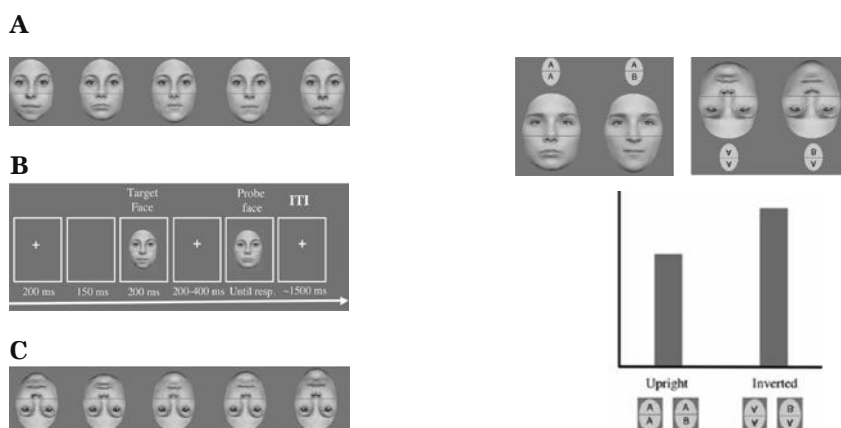


Figure 1.1. The face-composite paradigm. *Left: (A)* When five physically identical top halves are presented with different bottom halves, participants misperceived them as being different. *(B)* A delayed-matching task designed to evaluate the face-composite illusion. Participants have to determine whether the top half of the probe face is identical or different to the top half of the target face. *(C)* When presented upside-down, top halves are being correctly perceived as identical. *Right:* The expected matching performance for upright and inverted faces. Reprinted from Rossion (2013).

Box 3 – The Whole-Face Paradigm

The whole-face paradigm tests participants' performance at recognizing either the entire or individual parts of previously studied faces (Figure 1.2). In line with the holistic view of human face processing, recognition rates have been found to be greatly superior in the whole face compared to the condition in which the individual features are presented in an isolated way (Tanaka & Simonyi, 2016).

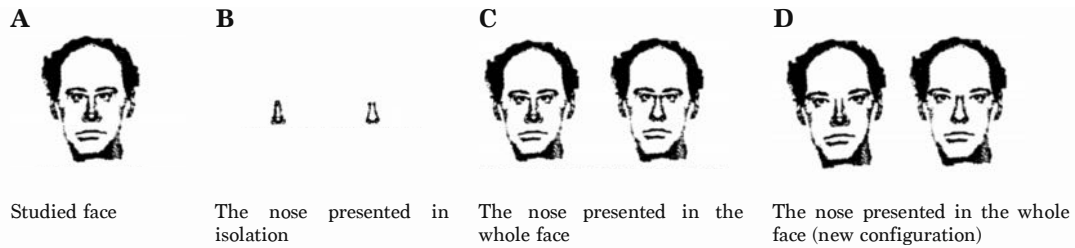


Figure 1.2. The whole-face paradigm. (A) Observers first study a face. Their memory for individual features presented (B) in isolation (C) in the studied face as a whole or (D) in a new face is subsequently tested in a two-alternative forced-choice task. Adapted from Tanaka and Simonyi (2016).

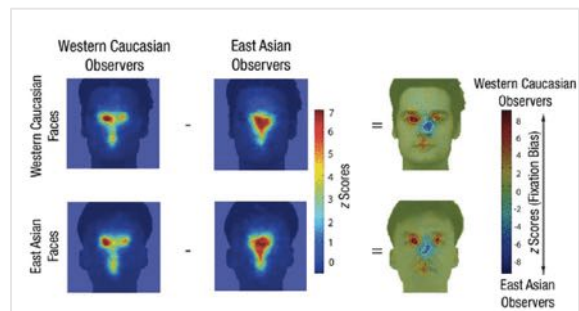
Box 4 – Flexible Face Processing Strategies

Although the dominant view in the literature posits that face processing is governed by universal holistic perceptual strategies (for a review, see Rossion, 2013), recent eye movement studies, investigating visual information sampling during face processing have put into question this prevalent assumption (for a review, see Caldara, 2017). For instance, Miellet, Caldara and Schyns (2011) have found that face identification was supported by both local and global strategies, depending on the location of the first fixation. Different strategies were even adopted by the same observer when the same face was presented in different trials.

Moreover, cross-cultural studies have provided further evidence for distinct face identification processing strategies. Blais et al. (2008) have demonstrated that Westerners identified faces by fixating the regions around the eye and the mouth, whereas Easterners favored central fixations towards the nose region (Figure 1.3).

Altogether these findings challenge the prevailing view that universal mechanisms support face identification, suggesting rather the existence of distinct and flexible strategies for face recognition.

Figure 1.3. Sampling visual information reveals culture-specific perceptual strategies. Left: Fixation maps of Western Caucasian and East Asian observers. Right: Fixation biases are evidenced by subtracting the East Asian z-scores from the Western Caucasian z-scored fixation maps. Reprinted from Caldara (2017).



1.1.2 The early development of face processing

Is there an innate preference for faces? The first stimuli that neonates perceive when they open their eyes are the faces of other human beings. Previous studies have revealed that within a few minutes after birth, newborns preferentially attend to schematic face-like configurations compared to scrambled stimuli that do not resemble a face (Goren, Sarty, & Wu, 1975; Morton & Johnson, 1991; Valenza et al., 1996). Other developmental studies have provided evidence that within a few hours after birth, newborns recognize the face of their mother (Bushnell, Sai, & Mullin, 1989; Pascalis et al., 1995; Sai, 2005). Such remarkable abilities already present at birth can be either supported by general perceptual mechanisms or by intrinsic predispositions to process faces (Pascalis et al., 1995). By relying on the latter proposition, Morton and Johnson (1991; for a recent review, see Johnson, Senju, & Tomalski, 2015) proposed a two-process theory of the development of face processing, defined primarily by an innate subcortical mechanism called CONSPEC. This face-sensitive mechanism is selectively tuned to geometrical face-like configurations (Figure 1.4A) and to the detection of biologically relevant conspecifics (Johnson et al., 2015; Morton & Johnson, 1991). The second mechanism, a stimulus-specific cortical mechanism, called CONLEARN replaces the initial CONSPEC mechanism approximatively at the age of two months and specializes in the recognition and processing of individual faces. For example, this cortical mechanism enables infants to distinguish between two individual faces based on internal fine-detailed features, an ability that develops between two to four months of age after extensive visual experiences with faces (Deruelle & de Schonen, 1998; Pascalis et al., 1998). According to this two-process theory, the subcortical mechanism guides the developing domain-specific cortical circuitry to acquire important information about faces. This complex interaction between both mechanisms, in turn, leads to the gradual emergence of the face-specialized neural network (Johnson et al., 2015). Since 1991, this two-process theory has been at the core of numerous debates and criticisms in the field of developmental psychology. Perhaps one of the most controversial proposals of this theory is the notion that intrinsic predispositions support infants' initial preference for faces (e.g., Gauthier & Nelson, 2001; Macchi Cassia, Simion, & Umiltà, 2001; Nelson, 2001). Doubts regarding this assumption along with new results have led some authors to propose alternative views.

These alternative views are based on 1) findings suggesting that the psychophysical properties of the stimuli account for infants' visual preferences for faces, 2) evidence that infants possess complex – non-specific – processing abilities at birth, 3) observations that infants exhibit visual preferences for up-down asymmetry, and 4) indications that cortical face-selective brain regions are activated shortly after birth. These different views and findings will be now briefly presented.

The sensory hypothesis. An alternative view proposes that particular characteristics of faces, though not face-specific, can simply account for newborns' visual preferences (Lee et al., 2011), because their psychophysical properties happen to best match the sensitivity of newborns' visual system (Turati et al., 2002). For example, Kleiner (1987) demonstrated that neonatal preferences for face-like patterns over other paired stimuli appeared only if the face-like patterns had an amplitude spectrum that better suited the contrast sensitivity of newborns' sensory channels as compared to the other stimuli.

Complex face processing abilities at birth. Previous experimental studies have shown that newborns possess more complex face representations than initially presumed (Quinn et al., 2008a). For example, infants prefer attractive faces over non-attractive ones (Slater et al., 1998; Slater et al., 2000); they are particularly sensitive to the eye region of a face (Batki et al., 2000) and prefer to look at smiling faces (Farroni et al., 2007) or at faces that directly gaze at them (Farroni et al., 2002). Such findings support the idea that newborns' face recognition abilities are more developed than originally assumed (Quinn et al., 2008a).

Visual preferences for up-down asymmetry. Other studies suggested that newborns' preference for face-like pattern is due to the presence of more elements on the upper portion compared to the lower portion of a configuration (Turati, 2004; Turati et al., 2002). By using the preferential looking paradigm (for more details on this paradigm, see [Box 5](#)), Turati et al. (2002) reported that newborns preferentially attended to stimuli with more elements in the upper region (Figure 1.4B), regardless of whether or not these stimuli resembled a face. By extending these findings to real faces, Macchi-Cassia et al. (2004) found that (1) newborns preferred real faces characterized by the presence of more elements in the upper part when contrasted with their reversed versions, (2) up-down asymmetry was not face specific, as non-face stimuli with more elements in the upper part triggered visual preference in infants, and (3) face-like configurations paired with non-face stimuli with more elements in the upper region did not trigger increased visual preference (Figure 1.4C). These experimental findings suggest that a non-specific top-heavy vertical-bias rather than a face-bias elicits infants' preference for face-like patterns.

Interestingly, a very recent prenatal study by Reid et al. (2017) evidenced the existence of a preference for face-like visual patterns present in human fetuses. More specifically, the authors investigated human fetuses' behavioral responses to upright and inverted face-like stimuli using 4D ultrasound technology. The data collected revealed that human fetuses, in the last trimester of pregnancy, were more likely to turn their heads to upright configurations as compared to inverted stimuli (Figure 1.5). These findings, consistent with previous results obtained with newborns or infants (e.g., Turati et al., 2002), suggest that postnatal experience

is not necessary for infants to exhibit a preference for face-like visual configurations (Reid et al., 2017). Finally, top-down asymmetry is a structural property that has also been proven to modulate the response of face-selective regions in adults. A neuroimaging study by Caldara and colleagues (2006) revealed an increased activation of the right Fusiform Face Area (rFFA), a key region among the face-selective brain regions, when geometrical patterns with high-contrasted elements in the upper part were exhibited to adult observers. These findings corroborate the idea that top-heaviness (i.e., up-down asymmetry) plays a crucial role in face processing, not only evoking face preferences in newborns, but also activating neural computations in the rFFA in adult observers.

Cortical activations at birth. A very recent developmental functional near-infrared spectroscopy study (fNIRS, for a description of this technique, see [Box 1](#)) with newborns, revealed face-selective activation over the bilateral posterior temporal cortex shortly after birth when neonates were exposed to human dynamic faces, but not when they were exposed to human arms (Farroni et al., 2013). This study provides evidence that certain cortical areas contribute to the processing of faces, even before major postnatal sensory experiences occur, corroborating the idea that both the subcortical and cortical pathways are already functional at birth (Acerra, Burnod, & de Schonen, 2002) and further develop through complex interactions (Nakano & Nakatani, 2014). Although this study clearly demonstrates cortical activation in response to dynamic faces, it is less evident whether the regions activated also control neonates' initial preferential looking orientation towards face-like patterns (Farroni et al., 2013).

Challenged by these findings, an expanded version of the CONSPEC mechanism has been elaborated (Johnson, 2011). This updated version suggests that the primary functions of this mechanism are not only to support the detection and preferential orientation of infants' attention towards faces, but also to enable them to choose adapted faces for social interactions (i.e., faces with a direct gaze or smiling). Thus, according to this view, the CONSPEC mechanism would serve as a basis for both face perception and social interactions (Johnson, 2011).

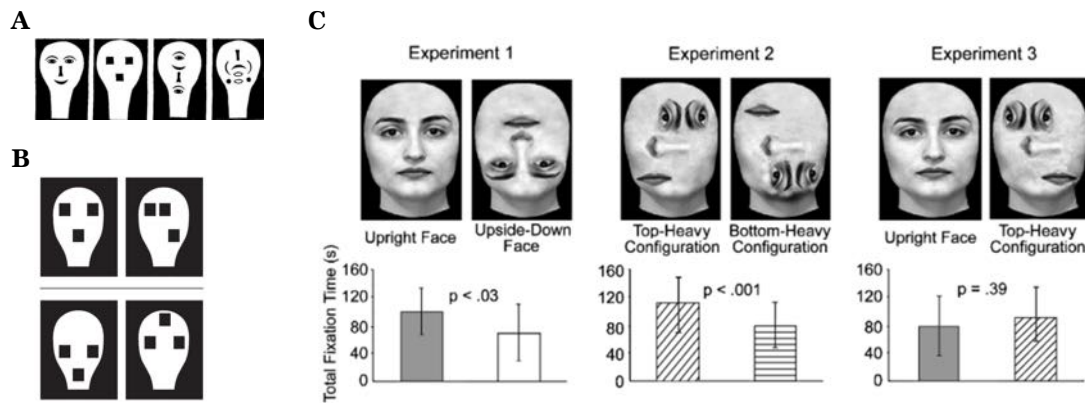


Figure 1.4 Infants' visual preference for faces. (A) Schematic configurations used in studies investigating infants' preference for face-like patterns. The two stimuli on the left are characterized by face-like configurations, although the second one is presented without the typical facial features. The two stimuli on the right are non-face stimuli, possessing facial features that are arranged in wrong configurations. Reprinted from Morton and Johnson (1991). (B) Schematic configurations used by Turati et al. (2002) to investigate whether up-down asymmetry (higher number of elements in the upper part) influences infants' visual preference for schematic face-like patterns. Reprinted from Turati et al. (2004). (C) Results of the study conducted by Macchi-Cassia et al. (2004) investigating whether newborns' face preference is driven by a top-heavy vertical bias. In experiment 1, pairs of stimuli presenting upright or inverted faces elicited greater preference for the upright versions. In experiment 2, non-face top-heavy configurations were preferred to bottom-heavy configurations. In experiment 3, upright faces were not preferred to top-heavy configurations. Reprinted from Macchi-Cassia et al. (2004).

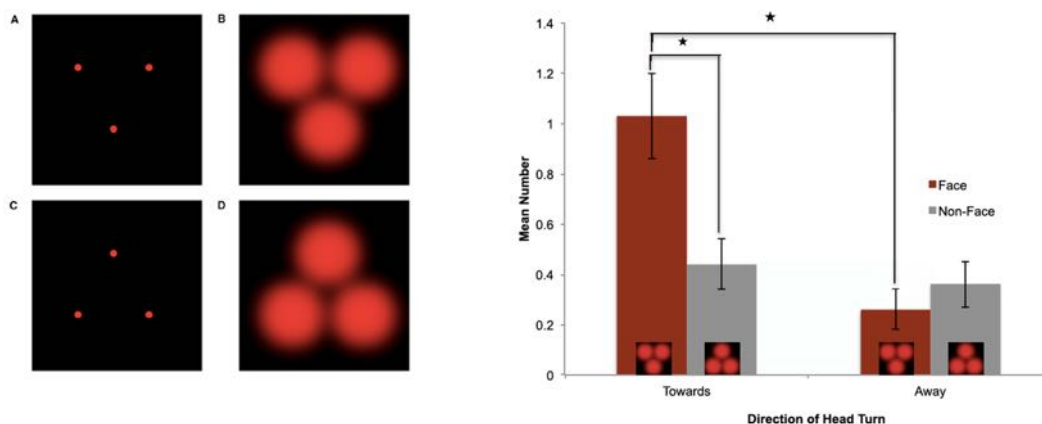


Figure 1.5. Prenatal preference for face-like visual patterns. Left: Upright and inverted examples of the stimuli presented on the maternal abdomen to the fetuses. (A) and (C) are examples of the stimuli before any contact with the abdominal tissue. (B) and (D) illustrate how the stimuli are perceived by the fetuses. Right: Mean number of head turns made by human fetuses towards and away the face-like (red) and non-face-like (gray) visual configurations. Error bars show standard errors, while stars indicate statistical significance. Adapted from Reid et al. (2017).

Box 5 – Experimental Paradigms Used in Infants’ Studies

The preferential looking paradigm. The preferential looking paradigm is one of the most commonly used technique to study preverbal newborn and infant behavior. It was first introduced by Fantz in the 1960s (Fantz, 1961, 1964, 1965). In preferential looking tasks, infants are presented with pairs of stimuli that differ in a specific manner (e.g., race, gender, specie, emotional expressions). The amount of looking time spent at each stimulus in the display is measured, indicating the relative level of interest for either one. In the original paradigm, a person looking through a hole measured looking time to either stimulus (Figure 1.6A). Nowadays, a low-light video camera located above stimulus presentation monitor digitally records infants’ looking behavior. The video-recordings are then digitized and manually coded frame-by-frame. This technique has been applied in the first study presented in this thesis (for more details, see CHAPTER 3).

The familiarization-novelty preference paradigm. The familiarization-novelty preference paradigm is a variation of the habituation-dishabituation method, a technique that uses heart rate (e.g., Canestrini, 1913), sucking (e.g., Bronshtein et al., 1958), and visual fixations (e.g., McCall & Kagan, 1970) as measures. The habituation-dishabituation technique measures the natural attentional decrease occurring in infants when repeatedly exposed to the same stimuli and their reaction to the presentation of a novel contrasting stimulus. If infants can distinguish between both stimuli, experimenters expect their attention to increase with the presentation of a novel stimulus (Houston-Price & Nakai, 2004). The familiarization-novelty preference technique employs a very similar procedure. Infants are first familiarized with a certain stimulus presented on display for a limited period of time. This is followed by a preferential-looking task during which a pair of stimuli (the familiar one and a novel one) are displayed to the infants (Figure 1.6C). If infants are able to differentiate between both stimuli, they should demonstrate a novelty-preference for the non-familiarized stimulus (Quinn & Eimas, 1986).

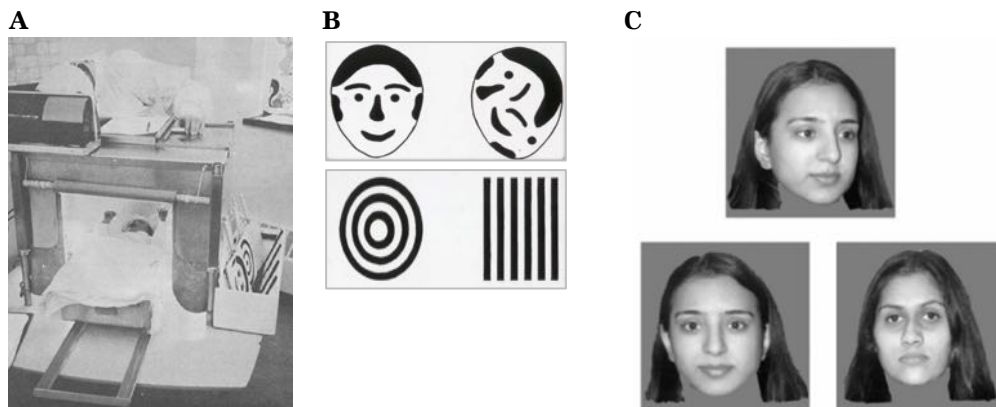


Figure 1.6. Experimental paradigms used with preverbal infants. (A) The preferential looking paradigm. In the original task, a newborn was placed in a “looking chamber” on whose ceiling different kinds of stimuli were presented by holding them with a handle over the newborns’ head. The experimenter measured infants’ looking behavior by peeping through a tiny hole. Reprinted from Fantz (1965). (B) Examples of the stimuli used in the original preferential looking task (Fantz, 1961, 1964). Newborns and infants were presented with several trials of paired stimuli. Their looking time for each stimulus was monitored, yielding a relative measure of their preference. Reprinted from Fantz (1965). (C) The familiarization-novelty procedure used by Kelly et al. (2007). Infants were familiarized to a face (in this example the face shown on top of the triad) until their attention was turned away for more than 2s. During the test phase, two faces (novel and familiar) were presented side-by-side. If infants were able to discriminate between both faces, their attention was directed to the novel face. Reprinted from Kelly et al. (2007).

1.1.3 The role of experience in face processing

In contrast to the widely debated phylogenetic view pertaining to the development of face processing, the CONLEARN mechanism emphasizing the importance of experience has been more readily accepted over the last two decades of developmental research. For instance, Le Grand et al. (2003) have demonstrated that early deprivation of visual inputs to the right hemisphere, due to congenital cataract in the left eye, strongly affected configural or second-order relational face processing abilities (i.e., distance between different facial features). Interestingly however, Sugita (2008) revealed that infant monkeys reared in isolation since birth in a visually rich environment, but with no exposure to faces or face-like visual stimuli, preferentially engaged with human and monkey faces compared to other non-face objects (see also Reid et al., 2017). A fine-grained discrimination advantage for human and monkey faces over non-face objects appeared however later, after an exposure to faces. A recent fMRI study by Arcaro and colleagues (2017) has provided further evidence that early exposure to faces is critical for developing face-domain brain circuits. Clusters of face-selective neurons are present in macaque monkeys exposed to faces since birth. In this study, such clusters were absent in macaques temporarily deprived from an exposure to faces during early development (Figure 1.7). Altogether, these findings emphasize the critical role played by early visual experience on the development of face processing abilities. Intriguingly however, in humans, a very recent study revealed that the ventral temporal cortex – a critical region for the processing of visual information – of individuals who were born blind elicited strong discriminatory responses to the auditory- presentation of stimuli from four distinct categories (e.g., faces, objects, body parts, scenes). Although congenitally blind individuals never processed any visual information, their category-selective visual map was remarkably similar to the cortical layout observed in normal-sighted controls. These findings suggest that early visual experience is not mandatory for the brain to develop category selectivity in the extrastriate cortex (Van den Hurk, Van Baelen, & Op de Beek, 2017). Altogether, these studies suggest that the ventral visual cortex is functionally shaped to develop category-selective processes independently from visual experience, but that such experience is mandatory for developing the more elaborated fine-tuned mechanisms dedicated to the recognition of individual exemplars (i.e., face recognition).

Developmental studies in early infancy have demonstrated that increased experiences with faces lead to stronger preferences for more sophisticated face properties. For example, Mondloch and colleagues (1999) discovered that newborns and 6-week-old infants were not able to distinguish schematic faces with a positive contrast (darker elements on lighter background) from faces with a negative contrast (lighter elements on darker background). By a marked contrast, 12-week-old infants, having had more experience with faces, exhibited preferences for faces with a positive contrast, a typical characteristic of the faces encountered in daily life

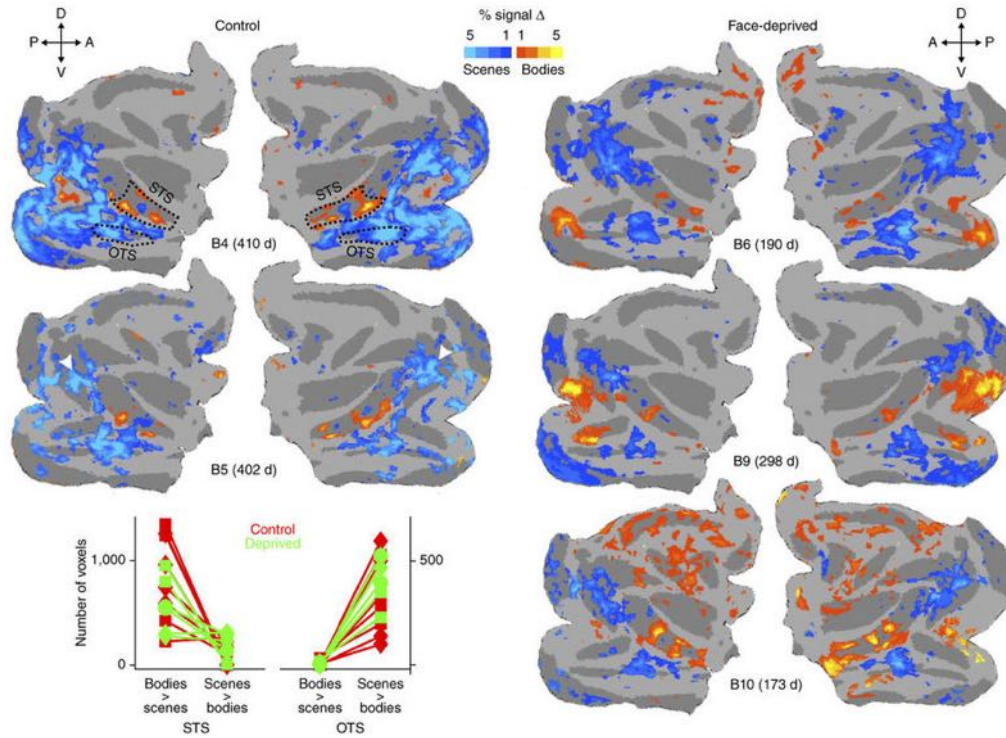
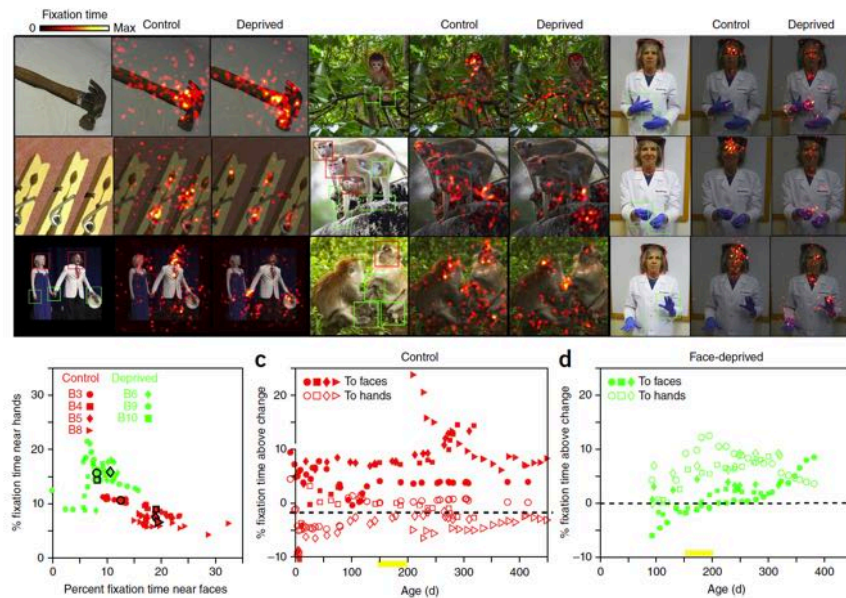
A**B**

Figure 1.7. The role of experience in the development of face processing abilities. (A) Face selective brain regions (i.e., the superior temporal sulcus in the example) were absent in monkeys who were reared with no exposure to faces. Brain regions responding preferentially to hands were however disproportionately large in these monkeys. **(B)** Interestingly, face-deprived monkeys preferentially attended to hands compared to any other objects while control monkeys preferentially looked at human and monkey faces. These findings suggest that early visual experience selectively affects brain connectivity and functional selectivity for processing objects of different categories. Adapted from Arcaro et al. (2017).

(Farroni et al., 2005). Other studies have revealed that the lack of experiences with specific types of faces leads to poorer recognition of these kinds of faces. Pascalis and Bachevalier (1998), for example, have evidenced that adult observers reliably recognized other human faces, but exhibited a decreased recognition performance with monkey faces. Interestingly, monkey faces demonstrated the opposite recognition pattern, being proficient with monkey, but impaired with human faces (Pascalis & Bachevalier, 1998). Such species-specific abilities have been explained by the increased experiences humans have with members of their community. Indeed, in their daily life, human observers are more exposed to other human faces than to monkey faces, whereas the opposite is true for monkeys.

During the early development stage, face processing abilities appear to be less affected by particular face types (Nelson, 2001). For example, Pascalis, de Haan, and Nelson (2002) revealed that 6-month-old infants were able to discriminate the identity of both monkey and human faces (Figure 1.8A). This was not the case of 9-month-olds, unless they were exposed to such faces daily (Figure 1.8B) (Pascalis et al., 2005). This progressive loss of ability with advancing age has also been observed with human faces from different ethnicities. By using the familiarity-novelty preference technique (for more details, see [Box 5](#)), Kelly et al. (2007), for example, evidenced that 3-month-old Caucasian infants were able to distinguish faces from their own and three other racial groups (African, Middle Eastern, and Chinese), 6-month-olds recognized Chinese faces as well as faces from their own group, whereas the recognition abilities of 12-month-olds were restricted to their own racial group (for an example of the stimuli, see Figure 1.8C). These observations are in line with the findings found while testing 3-, 6-, and 9-month-old Chinese infants (Kelly et al., 2009). This pattern of results, pointing to the crucial role played by experience in shaping face processing abilities, suggests that the face processing system progressively tunes to the faces encountered in our daily life (e.g., Kelly et al., 2007; Pascalis et al., 2002; Pascalis et al., 2005; Simpson et al., 2000). This process, named perceptual narrowing, parallels the course of language development (e.g., Kuhl, 2004; Kuhl et al., 2006; Kuhl et al., 1992; Werker & Tees, 1984 ; 2005); voice (e.g., Friendly, Rendall, & Trainor, 2013) and even music perception (e.g., Hannon & Johnson, 2005; Trainor, 2005; Trehub & Hannon, 2006). For example, previous studies have demonstrated that during the six first months of life, infants could discriminate the phonetic variations in almost every language in the world. Between 6 and 10 months however, their speech processing abilities progressively specialized in the perception of phonetic changes in their native language only (Kuhl et al., 2006; Werker & Tees, 1984).

Scott et al. (2007) suggest that the progressive tuning of domain-specific abilities can be explained by a gradual consolidation of the neural network supporting the processing of relevant properties that are encountered daily. According to this view, the decline in sensitivity, typical of perceptual narrowing, is not a developmental regression, but a gradual improvement in becoming

more proficient at processing biologically relevant features. Corroborating this idea, Humphreys and Johnson (2006) revealed that 7-month-olds, as compared to 4-month-olds, relied on fewer physical information to distinguish between two faces of their own race, suggesting that perceptual narrowing enhances fine-grained recognition abilities of experienced face subtypes.

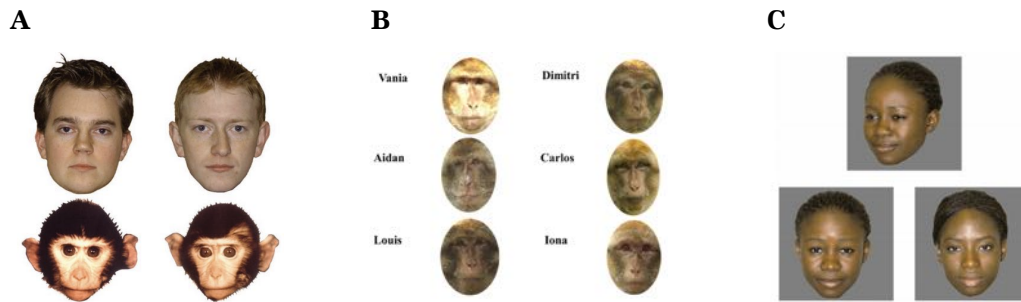


Figure 1.8. The different stimuli used to investigate perceptual narrowing during infancy. (A) Pascalis et al. (2000) tested 6- and 9-month-old infants' abilities to recognize human and monkey faces. Reprinted from Pascalis et al. (2000). (B) The daily exposure of infants between 6 and 9 months of age to non-native faces of Barbary macaques facilitates the discrimination of these monkey faces at 9 months, an ability that is otherwise lost by this age. Reprinted from Pascalis et al. (2005). (C) An example of the stimuli used by Kelly et al. (2007, 2009). At 3 months of age, infants were able to discriminate faces from all race groups (Caucasian, African, Middle Eastern, and Chinese), an ability that was restricted to Chinese faces and faces from their own group at 6 months. By the age of 9 months, the discrimination abilities were restricted to faces of their own race. On each trial, infants were first habituated to a particular face (in this example the face shown at the top of the triad). This presentation was followed by a test phase consisting of a paired presentation of the familiar and a novel face (faces at the bottom). Reprinted from Kelly et al. (2009).

1.1.4 Gender processing

Daily perceptual experiences play a crucial role in shaping face processing abilities. For instance, they modulate identity discrimination for faces of the same race as compared to faces of people from other race or species and also extend to the categorization of socially-relevant classes, such as gender. For example, Quinn et al. (2002) revealed that 3- to 4-month-old infants were able to discriminate novel female faces from familiar faces, but showed deficits in recognizing familiar male faces (see also, Ramsey-Rennels & Langlois, 2006). Infants' advantage in the processing of female over male faces is considered driven by the predominant visual experience with female faces as compared to male faces during early infancy. Interestingly, Quinn and colleagues (2008b) revealed that this fundamental preference for female faces is present only in the same but not in faces from other races (Quinn et al., 2008b), and the opposite pattern of preference is observed when the primary caregiver is a male (Quinn et al., 2002). Moreover, Liu et al. (2015) evidenced that the preference disappears in older infants after increasingly experiencing male faces.

Research with older infants have displayed that their ability to categorize face gender develops between five months and one year of age, although the precise age of emergence remains unclear (e.g., Cohen & Strauss, 1979; Cornell, 1974; Fagan & Singer, 1979; Leinbach & Fagot, 1993; Younger & Fearing, 1999). For example, Cohen and Strauss (1979) demonstrated that by the age of 7 months, infants could perform inclusive categorization of female faces. The ability to form exclusive categories of female and male faces appeared to emerge between 9 and 12 months when body-related aiding cues such as hair contours or clothes were provided (Leinbach & Fagot, 1993). In another study, Younger and Fearing (1999) used a familiarization-novelty preference procedure to investigate infants' ability to distinguish faces judged as gender-typical from faces difficult to categorize with respect to gender. Infants were familiarized to both female and male faces that clearly matched a gender-specific category (Figure 1.9A). Following familiarization, infants were exposed to novel faces that either matched the female-male face category previously exhibited, or novel faces that were gender ambiguous (Figure 1.9B). As shown in Figure 1.10, the findings revealed that 10-month-olds, but not 7-month-olds, were able to differentiate gender-typical from gender ambiguous faces (i.e., their attention increased towards faces whose gender was ambiguous). Thus, the current literature suggests that the ability to perform categorization of female faces emerges between 5 and 7 months, whereas the ability to classify both female and male faces into separate categories emerges later, between 9 and 10 months.



Figure 1.9 Gender processing in infants. Stimuli used by Younger and Fearing (1999) to investigate infants' ability to separate faces into specific categories (gender-typical vs. gender-atypical). **(A)** Stimuli presented during the familiarization phase. During the familiarization phase, two groups of infants (7- and 10-month-olds) were habituated to 8 gender typical faces (4 females, 4 males). **(B)** Stimuli presented during the test phase. The first row represents male stimuli, the second female stimuli. The first column displays gender typical faces, whereas the second displays gender-ambiguous faces. During the test phase, in addition to the eight familiarized faces, half of the infants of each age group saw a novel male gender-typical as well as a novel male gender-ambiguous face. The other half of the infants saw the novel gender-typical and gender-ambiguous female faces. Adapted from Younger and Fearing (1999).

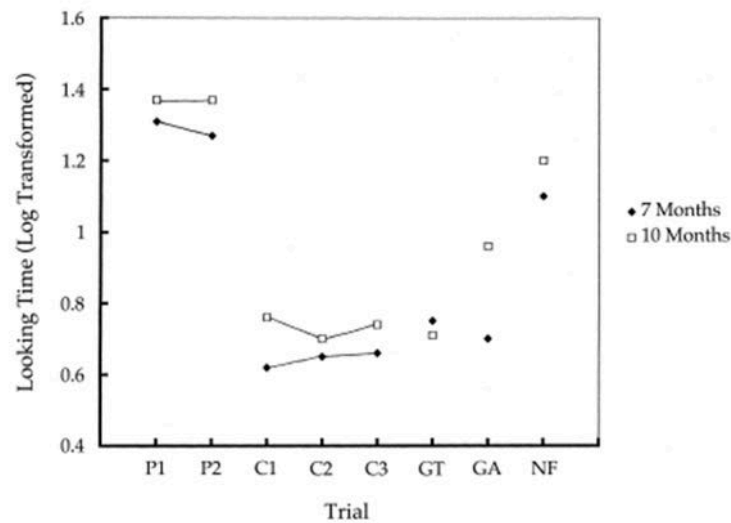


Figure 1.10. Developmental change in categorization abilities. Results reported by Younger and Fearing (1999). During the familiarization phase, the mean looking time to peak habituation trials (P1 and P2) as well as criterion habituation trials (C1, C2, C3) were measured. During the test phase, mean looking time for gender-typical (GT), gender-ambiguous (GA) as well as non-face trials (NF) were measured. The 7-month-olds were not able to differentiate female and male faces into specific categories. Moreover, they showed habituation to both gender-typical and gender-ambiguous faces. In contrast, 10-month-olds were able to form specific gender-related categories during the familiarization phase, as well as discriminate gender-typical (decreased attention to a gender-category previously shown during the familiarization phase) from gender-ambiguous faces (increased attention to a gender-category that did not clearly fit a category to which they were familiarized). Reprinted from Younger and Fearing (1999).

1.1.5 Faces as a multiple source of information

The integration of audible and visible information. Previous intermodal matching studies investigating infants' abilities to integrate various multisensory properties (e.g., a voice with a particular face) revealed that 4.5-month-old infants were able to match phonetic information with faces and voices when presented with vowels (Kuhl & Meltzoff, 1982, 1988; Patterson & Werker, 1999; Walton & Bower, 1993) or consonant-vowel disyllable speech (MacKain et al., 1983). Other studies evidenced indirect sensitivity of 4.5-month-old infants to articulatory movements and vowel sounds by revealing greater imitation abilities for the coherent entity rather than the audiovisual mismatch (Kuhl & Meltzoff, 1988; 1996; Legerstee, 1990; Patterson & Werker, 1999). A recent study demonstrated, however, that infants acquire the ability to match faces and voices for fluent speech only between 12 and 14 months of life, with a later emergence for non-native language affected by perceptual narrowing (Lewkowicz et al., 2015). This latter example illustrates that although the ability to associate audible speech with visual information emerges early in infancy (Lewkowicz, 2000; 2010), it seems to be, at an initial stage, restricted to the understanding of whether articulatory movements and sounds occur together.

Multisensory perception of gender. The audio–visual matching skills associated with distinct and more complex properties, such as native or non-native language (Lewkowicz et al., 2015), emotional expressiveness (Walker-Andrews, 1986), identity (Lewkowicz & Pons, 2013), or gender, appear later (Patterson & Werker, 2002; Poulin-Dubois, Serbin, & Derbyshire, 1998; Poulin-Dubois et al., 1994; for a review, see Ramsey, Langlois, & Marti, 2005, pp. 219–229; Walker-Andrews et al., 1991). Previous studies have demonstrated that infants’ ability to match audible and visible attributes of gender emerges between 6 and 8 months of age (e.g., Hillairet de Boisferon et al., 2015; Patterson & Werker, 2002; Poulin-Dubois et al., 1994). For example, Patterson and Werker (2002) evidenced that 8-month-old infants could match faces and voices when presented with infant-directed vowels (Figure 1.11). In contrast, Poulin-Dubois et al. (1994) revealed that the specific skills mandatory to effectively match faces and voices when presented with gender information emerged only between 9 and 12 months of age for female faces and not before 18 months for male faces (Poulin-Dubois et al., 1998). Consistent with these findings, a very recent study by Hillairet de Boisferon et al. (2015) revealed that infants were not able to extract gender information from faces and voices before the age of 9 months, with the relevant skills being restricted to female faces and voices (Figure 1.12). As reported by Ramsey et al. (2005), a large part of the inconsistency prevalent across studies regarding the developmental time course of the multisensory perception of gender might stem from methodological factors, such as differences in the visual stimuli (static vs. dynamic) or auditory stimuli (isolated vowels vs. fluent speech) used in tasks or their complexity. Interestingly, Kubicek et al. (2014) recently reported that the integration of multisensory information was also influenced by the manner of speech. These authors revealed that infant-directed speech (see **CHAPTER 2** and **CHAPTER 3**) facilitated inter-sensory matching of audible and visible native-language attributes. In contrast, adult-directed speech did not facilitate the integration of these attributes. To the best of our knowledge, no study has evaluated the influence of the speech manner on the integration of multisensory gender information yet. The first experimental contribution of this thesis aims at elucidating this question, also considering the aforementioned methodological considerations (**CHAPTER 3**).

In this first section, I provided a brief overview of the development of face processing. Subsequently, I will focus on the literature related to emotion, presenting the six basic expressions, the way their perception develops throughout lifespan, and whether naturally evolving dynamic expressions are better recognized than their static versions.

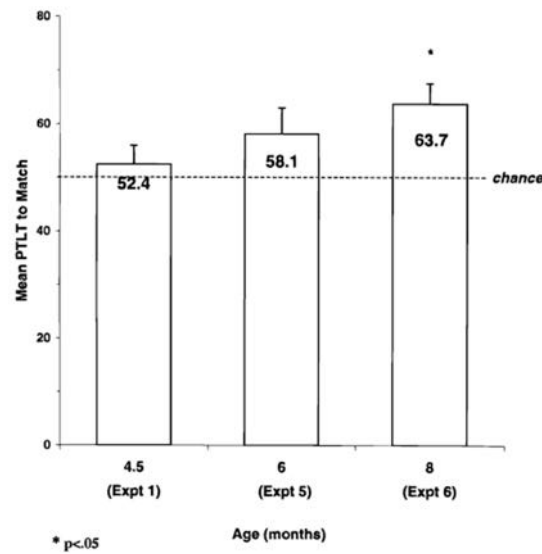
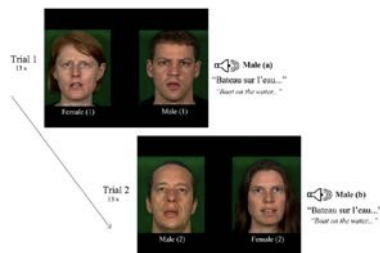


Figure 1.11. Multisensory integration of faces and voices. By the age of 8 months, infants can match faces and voices when presented with infant-directed vowels. Reprinted from Patterson and Werker (2002).

A



B

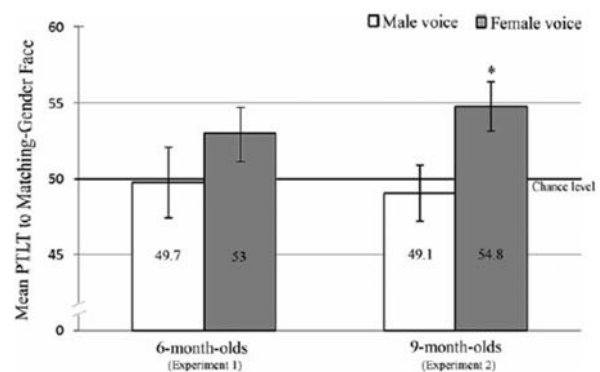


Figure 1.12. Multisensory integration of faces and voices is restricted to female faces. (A) Schematic representation of the procedure used by Hillairet de Boisferon et al. (2015) to investigate 6- and 9-month-old infants' abilities to match audible and visible attributes of gender when exposed to dynamic faces reciting infant-directed nursery rhymes. Infants viewed two side-by-side video clips of a man and a woman reciting a nursery rhyme and heard a synchronous male or female soundtrack. (B) Infants did not associate audible and visible gender attributes until the age of 9 months, and even then, it was only for female faces. Adapted from Hillairet de Boisferon et al. (2015).

1.2 FACIAL EXPRESSIONS OF EMOTION

1.2.1 Origins and categories of emotions

From an evolutionary perspective, physiological drives such as hunger, thirst, lust, avoidance of pain, and need for sleep are critical for our survival, hence the need to possess them. What about emotions? Are they also critical for our survival? Do we all have emotions? But first and foremost, what are emotions?

Aristotle (384–322 B.C.) defined emotions “as that which leads one’s condition become so transformed that his judgement is affected and which is accompanied by pleasure, and pain. Examples of emotion include anger, fear, pity and the like, as well as the opposites of these” (Solomon, 2008, p. 5). In Ancient Greek, emotions were initially described as *pathèma*, translating to passion in English and French, before Descartes (1596–1650) first introduced the term emotion in “*Les émotions de l’âme*” (Frijda, 2008). In his seminal work “*The Expression of the Emotions in Man and Animals*”, Darwin (Darwin, 1999/1872) emphasized the phylogenetic nature of facial expressions, paving the way for decades of research in the domain of emotional expressions. His evolutionary approach of considering emotions as biologically relevant signals, shaped by selection and critical for survival and adaptation, has become the common foundation for several psychological theories pertaining to emotions (Plutchik, 2003). From a psychological perspective, emotions are commonly defined as episodic changes in different subsystems of the organism, including neural responses, physiological reactions, internal feelings and muscular activations. These changes are triggered by external or internal influences that are of critical importance for the organism, enhancing fitness and reactions in specific situations (Scherer, 2000). Fear signals potential danger encouraging escape, disgust leads to the avoidance of harmful substances, anger provokes counterattack, sadness leads people to seek help or encourages them to overcome a loss, whereas joy encourages people to repeat actions that result in pleasure or happiness (Nesse & Ellsworth, 2009; Plutchik, 2003).

How many different emotion categories are there? This question has been at the core of ensuing debates over the past decades (Ekman, 1992a; 1992b; Ekman & Davidson, 1994; Nesse & Ellsworth, 2009; Ortony & Turner, 1990, see Table 1.1). Although some theorists have postulated the existence of only two distinct affective states (i.e., positive vs. negative) (Nesse & Ellsworth, 2009), Ekman and Friesen (1976b; 1978), founders of the most prominent categorical approach and proponents of the Darwinism approach, have posited the existence of a limited number of six basic emotions: anger, disgust, fear, joy, sadness, and surprise. These so-called pure emotions are defined by specific facial activations, autonomic responses, and subjective

feelings (Ekman, 1992a; Ekman & Friesen, 1971; 1976b; 1978; Ekman, Levenson, & Friesen, 1983).

By virtue of their evolutionary and biological relevance, the six basic expressions have long been considered as universal signals, critical for communicating internal states (e.g., Ekman & Friesen, 1978; Ekman, Sorenson, & Friesen, 1969). This universal hypothesis has, however, been recently challenged by findings from the cultural neuroscience literature. Jack et al. (2009), for instance, have observed that individuals from East Asia commonly rely on the region around the eyes to achieve expression discrimination tasks, whereas Western adults sample information from all facial features (Jack et al., 2009). By using a novel 4D method, coupled with a reverse correlation technique, the authors have further demonstrated that the diagnostic information used to decode and represent basic expressions differed across cultures (Jack, Caldara, & Schyns, 2012a; Jack et al., 2012b) with the eyes being more informative for Easterners, whereas the region around the mouth provided more diagnostic information to Westerners. Interestingly, these cultural differences were already observed in 7-month-old infants (Geangu et al., 2016), suggesting cultural influences very early on in the development of facial expression recognition (Caldara, 2017).

Table 1.1. The basic emotions. Adapted from Ortony and Turner (1990).

Reference	Fundamental emotion	Basis for inclusion
Arnold (1960)	Anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, sadness	Relation to action, tendencies
Ekman, Friesen and Ellsworth (1982)	Anger, disgust, fear, joy, sadness, surprise	Universal facial expressions
Gray (1982)	Rage and terror, anxiety, joy	Hardwired
Izard (1971)	Anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise	Hardwired
James (1884)	Fear, grief, love, rage	Bodily involvement
McDougall (1926)	Anger, disgust, elation, fear, subjection, tender-emotion, wonder	Relation to instincts
Mowrer (1960)	Pain, pleasure	Unlearned emotional states
Oatley and Johnson-Laird (1987)	Anger, disgust, anxiety, happiness, sadness	Do not require propositional content
Panksepp (1982)	Expectancy, fear, rage, panic	Hardwired
Plutchik (1980b)	Acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise	Density of neural firing
Watson (1930)	Fear, love, rage	Hardwired
Weiner and Graham (1984)	Happiness, sadness	Attribution independent

Although other theorists (e.g., Gray, 1982; Izard, 1971; Panksepp, 1982; Plutchik, 1980b) have postulated the existence of additional basic emotions (Table 1.1) such as pain, pleasure, anticipation, contempt, panic, or distress, studies investigating the recognition of facial expressions commonly rely on the six basic emotions originally defined by Ekman and Friesen (1976). This thesis follows this common direction, as the experimental contributions presented in **CHAPTER 4** and **CHAPTER 5** restrict their investigations to the expressions of “anger”, “disgust”, “fear”, “happiness”, “sadness”, and “surprise”.

1.2.2 Transmitting and decoding facial expressions

Understanding the manner in which emotions are transmitted and decoded is of fundamental importance for achieving adapted social interactions. The dynamic interaction between signaling and decoding facial information is inherent to all communication (e.g., emotions, gait, vocal prosody, body postures and gestures). As illustrated in Figure 1.13, a face, as a sender, transmits a signal (e.g., an expression) to the brain that decodes and extracts the perceptually-relevant information to achieve an adequate interpretation of the incoming inputs, based on prior knowledge (Jack, 2013; Jack & Schyns, 2015). Both perceptual signals transmitted by facial expressions and top-down mechanisms are necessary for an effective categorization of a perceived expression.

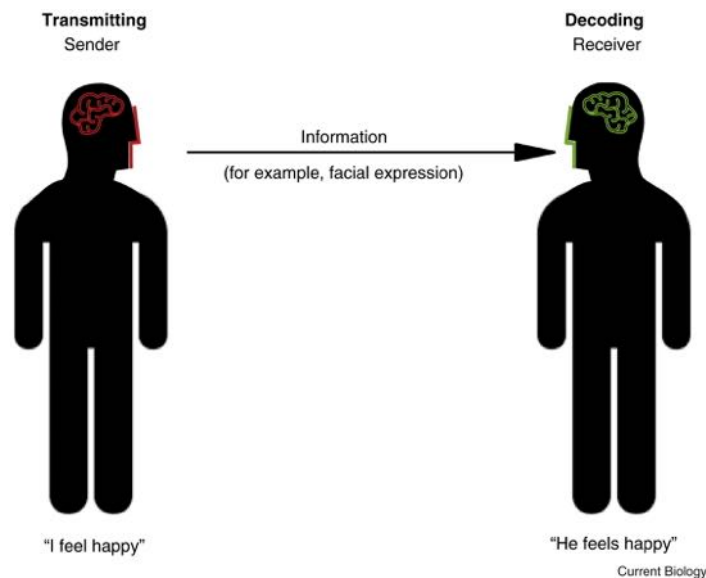


Figure 1.13. Transmitting and decoding facial expressions of emotion. Reprinted from Jack and Schyns (2015).

Smith, Cottrell, Gosselin, and Schyns (2005) examined the perceptual characteristics of the facial signals transmitted by the six basic expressions (plus neutral) by adopting an ideal observer model approach with a response classification technique named Bubbles, which samples information in a 3D space (Gosselin & Schyns, 2001; Schyns, Bonnar, & Gosselin, 2002) (Box 6, Figure 1.14). Their findings revealed that the face, as a facial expression transmitter, maximizes orthogonal signals with low overlap to reduce ambiguities and enhance categorization performance. As shown in Figure 1.15, the characteristics of the six basic expression signals were highly distinguishable from each other, as low Pearson correlations were observed between the different facial features used for transmitting the expressions.

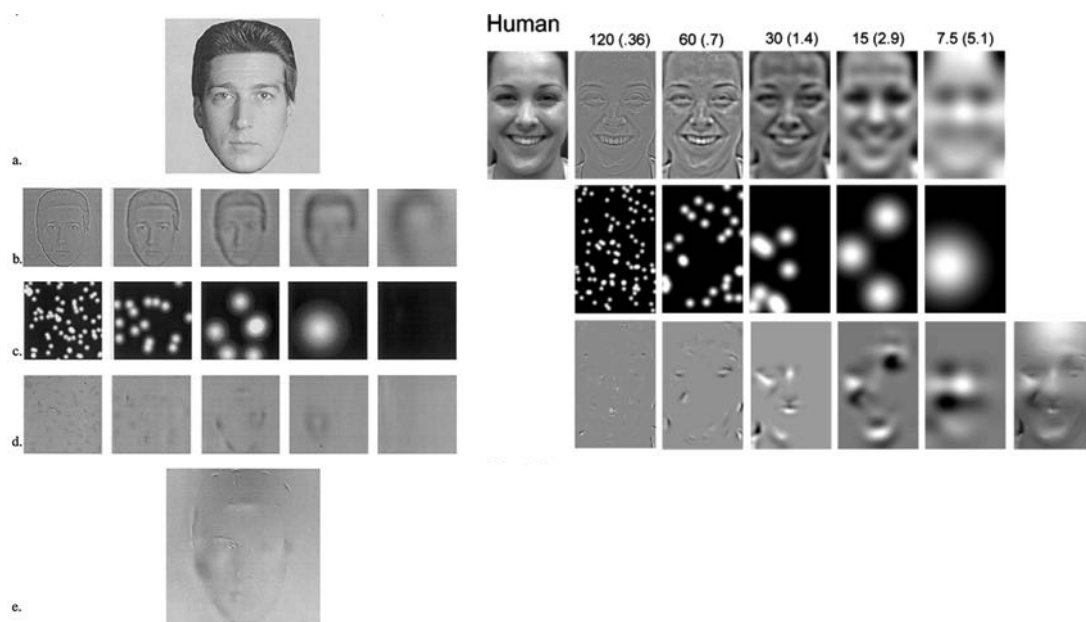


Figure 1.14. Bubbles. Bubbles is a response classification technique that samples the information in a 3D space (2D face and spatial frequency). *Left:* The sampling procedure used originally by Gosselin and Schyns (2001) to create bubbled stimuli. *Right:* The procedure used by Smith et al. (2005). The upper part of the figure provides a schematic illustration of how the stimuli were created. An original face (first column) was decomposed into five different bandwidths of spatial frequencies (first row, 120 to 7.5 cycles per face). Each bandwidth was then sampled with randomly positioned bubbles (second row). The third row illustrates how the bubbles were integrated within the decomposed pictures and summed to generate an experimental stimulus (for more details of this technique, see Gosselin & Schyns, 2001; Schyns et al., 2002). Reprinted from Smith et al. (2005).

Box 6 – Bubbles

Bubbles is a response classification technique that isolates the facial information used by observers to categorize different facial features. This general technique, elaborated by Gosselin and Schyns in 2001, has been used in different tasks (e.g., expression, identity, and gender categorization) to segregate the location of the facial cues responsible for adequate categorization. Although the Bubbles technique has been extensively used for face recognition, its principles have also been extended to the recognition of objects (Wasserman et al., 2004) and scenes (McCotter et al., 2005) as well as to the investigation of visual perception in brain-damaged patients (Adolphs et al., 2005; Caldara et al., 2005; Fiset et al., 2017). In the Bubbles technique, a mask consisting of randomly located bubbles (Gaussian apertures) is generated and applied on an original image decomposed into five bandwidths of spatial frequency. The resulting stimuli sample a subset of the visual information presented in the original image (Figure 1.14). Observers categorize the stimuli generated throughout this sampling process, a procedure that allows to assign credits to the bubbles leading to correct categorization. Over the testing procedure, the number of bubbles is adjusted in order to obtain a criterion level of performance of 75% of correct categorization (Gosselin & Schyns, 2001).

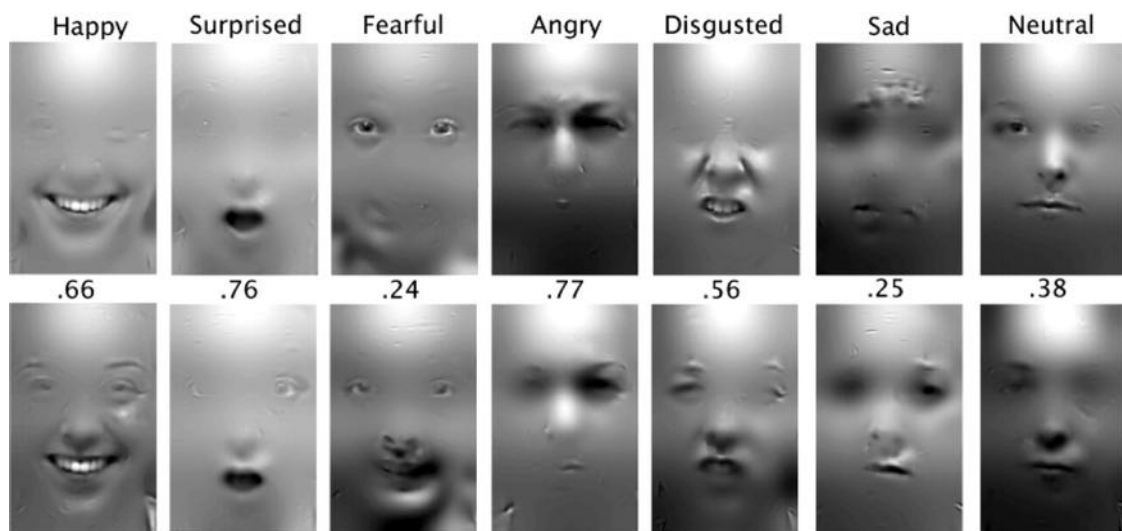


Figure 1.15. Facial expression signals are weakly correlated with one another. The upper row represents the diagnostic filtering function for the human observers whereas the lower row the diagnostic function of the model observer. A diagnostic filtering function was generated for each expression at each scale, by locating the pixels leading to a criterion level of performance of 75% correct. The resulting filters, specific to each scale, were then smoothened and multiplied by a selected stimulus sample to generate the images shown in the figure. The numbers reported indicate the values of the Pearson correlations between the diagnostic filtering functions of humans and observer model. High correlation values indicate high adaptation to image information statistics (for more details, see Smith et al., 2005). Reprinted from Smith et al. (2005).

Besides the response classification technique adopted by Smith et al. (2005) in characterizing the information transmitted by the six basic expressions, other systems, such as the Facial Action Coding System (FACS, Ekman & Friesen, 1976b; 1978), were previously developed in order to examine which facial signals are involved in the production of specific emotions.

1.2.3 Communicating emotions: the activation of different facial muscles

Over the last decades, most studies investigating facial expression recognition have been influenced by Ekman and Friesen's Facial Action Coding System (1976b; 1978), providing an anatomical taxonomy of the muscles activated during the generation of natural expressions. According to this system, facial expressions can be described in terms of facial muscle activations, the so-called action-units. An action unit corresponds to a small movement of a single muscle or a group of muscles, which is easily distinguishable from a different movement (for an illustration of the different muscles for facial expressions, see Figure 1.17). The activation of an action unit leads to a specific change in the appearance of the face, such as a smile, the lift of an eyebrow, a grin, or an upper or lower nose wrinkle. The FACS consists of 46 actions units (AUs) corresponding to all the independent different facial motions. As illustrated in Figure 1.16, an expression can be decomposed into its specific AUs. The FACS is a very sensitive system, which is able to distinguish subtle changes in facial appearance. It has been used to detect genuine, suppressed, or fake painful expressions in children (e.g., Larochette, Chambers, & Craig, 2006), older adults with or without Alzheimer (e.g., Lints-Martindale et al., 2007), patients suffering from chronic back pain (e.g., Craig, Hyde, & Patrick, 1991), and patients with depression (e.g., Reed, Sayette, & Cohn, 2007). It has also proven to be a useful tool for detecting subtle momentary changes in genuine or fake smiles (e.g., Del Giudice & Colle, 2007), for detecting inappropriate affects in schizophrenic patients (e.g., Kohler et al., 2008), and to examine how specific the activations of individual actions units are in healthy adults (e.g., Kohler et al., 2004a).

Ekman and Friesen's facial action coding system has also been the starting point for the development of state-of-the-art new technologies created to investigate facial expression recognition. For instance, Hamm et al. (2011) developed an automated FACS to examine dynamic changes of facial movements in neuropsychiatric disorders. Roesch et al. (2011) created FACSGen, an innovative tool based on the FACS, to develop static and dynamic realistic 3D face stimuli by controlling the activation of specific AUs. Furthermore, Jack, Caldara, and Schyns (2012a) reconstructed the internal representations of the six basic expressions in Easterner and Westerner adults using a FACS-based generator of random facial muscle

activations with a reverse correlation technique. This study revealed cultural specificity in the set of facial movements used by Easterners and Westerners to represent basic emotions (Jack et al., 2012, for more details on this technique, see **CHAPTER 2** and **CHAPTER 4**).

In the first part of this second section, I provided a theoretical overview of some aspects of emotions, their origins, how they are transmitted and decoded. The next part proposes a brief review of their development, from childhood to old age.



Figure 1.16. AUs for disgust. The expression of disgust leads to the activation of both AUs 9 (upper nose wrinkle) and 10 (lower nose wrinkle and lift of the upper lip). Reprinted from Pochedly, Widen and Russell (2012).

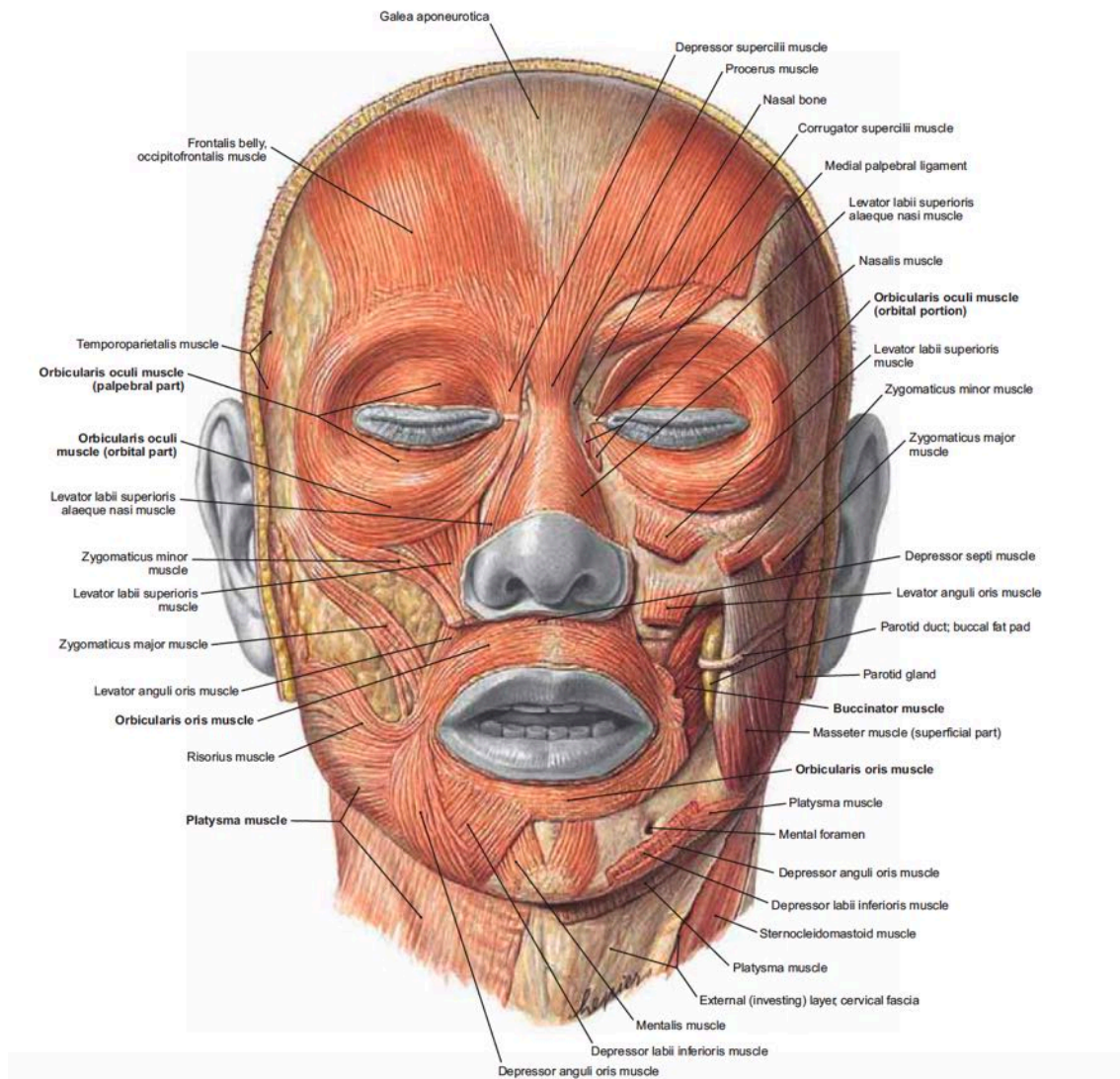


Figure 1.16. Muscles for facial expressions. Reprinted from Clemente (1997).

1.2.4 The development of expression recognition: a lifespan perspective

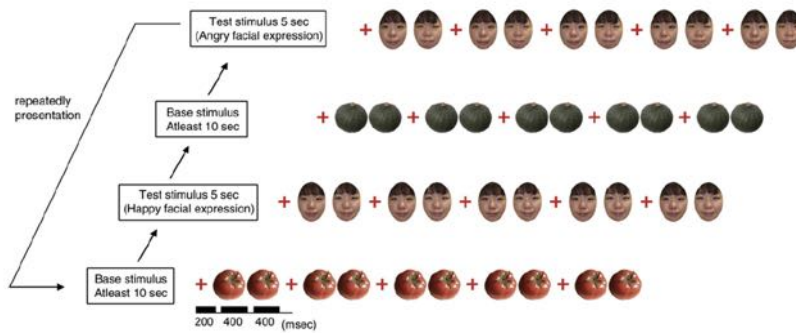
Infancy. The development of emotional processing has been thoroughly investigated in infants (for reviews, see Bayet, Pascalis, & Gentaz, 2014; Leppänen & Nelson, 2009). Several studies have evidenced a gradual improvement in the recognition of individual expressions, between 4 months and one year of age. For instance, performing a “peek-a-boo” game and measuring infants’ looking behavior, Montague and Walker-Andrews (2001) found that 4-month-old infants displayed varied patterns of responsiveness to the different emotions. These findings are consistent with other developmental studies showing early expression imitation abilities in neonates (Field et al., 1983) as well as research demonstrating that 4-month-old infants prefer smiling to neutral faces (LaBarbera et al., 1976; Rochat, Striano, & Blatt, 2002), smiling to sad faces (Montague & Walker-Andrews, 2002), and smiling to fearful faces (Farroni et al., 2007; but see, Peltola et al., 2009). Other studies have shown that infants at a very young age are able to determine the emotional intensities of subtle versus intense happy faces (e.g., Nelson & De Haan, 1996).

A first important developmental shift has been evidenced in 6- to 7-month-old infants with the emergence of categorization and discrimination abilities (Bayet et al., 2014; Leppänen & Nelson, 2009). Several studies have demonstrated that infants were able to perform categorical discrimination between happiness and surprise (Caron, Caron, & Myers, 1982), sadness (Leppänen & Nelson, 2009), and fearful faces (Kotsoni, de Haan, & Johnson, 2001). Using functional near-infrared spectroscopy (fNIRS), Nakato et al. (2011) further demonstrated that 6- to 7-month-old infants were able to discriminate happy from angry expressions and that their hemodynamic responses to these two expressions were distinct. More specifically, in their experiment, infants’ looking behavior was measured by video-recordings and their brain activity using fNIRS while they viewed passively angry and happy faces (Figure 1.18). As shown in Figure 1.19A, the experimental procedure comprised baseline and test phases. Vegetables were displayed during the baseline, while expressions were presented during the test phase. Changes in hemodynamic responses between baseline and test phases were computed to identify the hemodynamic activations during emotional face processing. Results revealed distinct patterns of hemodynamic responses to both expressions, with gradually increasing left temporal activations to happy faces, while angry faces elicited hemodynamic responses in the right-sided temporal areas (Figure 1.19 B and C).

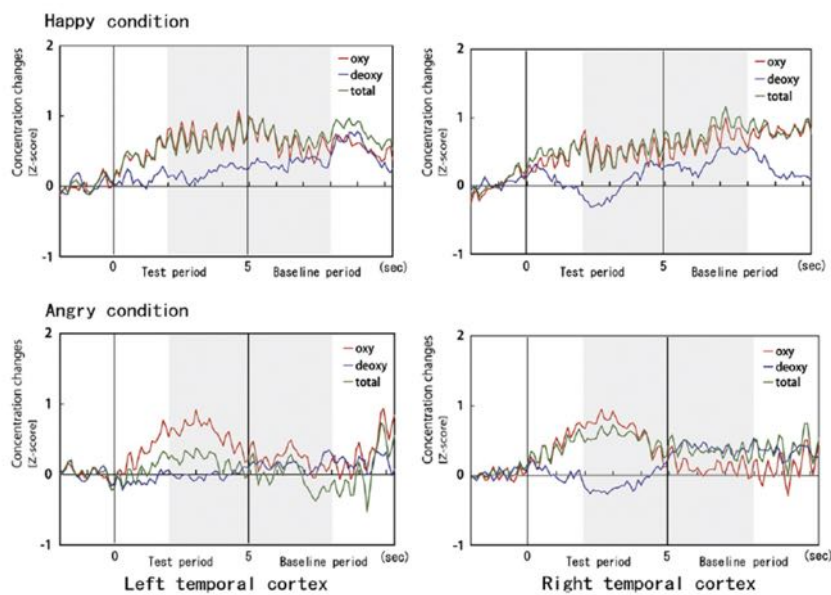
Figure 1.18. fNIRS procedure in developmental studies. The fNIRS cap is placed over the right- and left-sided temporal regions in 6- to 7-month-old infants. Reprinted from Nakato et al. (2011).



A



B



C

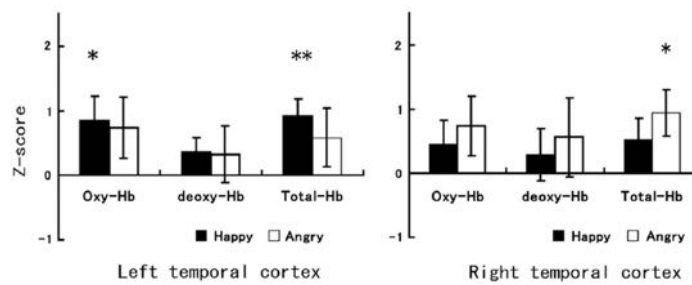


Figure 1.19. Differences in hemodynamic responses to happy and angry faces evidenced by a near-infrared spectroscopic study. (A) An example of the experimental procedure used in the study. **(B)** Time-course of oxy-, deoxy-, and total-hemoglobin concentrations during the happy and the fearful face conditions. The left column reports concentration changes in the left temporal cortex, whereas the right column indicates changes in the right-sided temporal cortex. **(C)** The figure displays the mean z-scores for oxy-, deoxy-, and total-hemoglobin concentrations for happy and angry faces during the last 3s of the test phase and the 3s following the test phase in left and right temporal cortices. Adapted from Nakato et al. (2011).

Although young infants already demonstrate a few expression differentiation abilities, it is not until the end of the first year of life that their emotion decoding system becomes sensitive enough to detect more subtle differences in emotional facial signals. For instance, 8- to 10-month-old infants expect specific emotional reactions (e.g., happy faces) in response to the outcomes of particular actions (e.g., success) and detect expressions (e.g., sad faces) that are incongruent with the outcomes elicited by specific conditions (e.g., success) (Skerry & Spelke, 2014).

Childhood and adolescence. While the behavioral and fNIRS data presented above demonstrate that 6- to 7-month-old infants display sensitivity to emotional faces and can discern certain emotions, their ability to reliably label and categorize expressions emerges later and gradually increases from childhood to early adulthood (Herba & Phillips, 2004; Montiroso et al., 2010; Rodger et al., 2015; Vicari et al., 2000). Although it is evident that emotion recognition steadily improves with increasing age, little consistency is found across studies regarding the exact developmental trajectory of this ability. Differences in the methodologies applied, in the stimuli (static versus dynamic), and in the emotions tested make it particularly difficult to compare findings across studies. For instance, recognition rates have been proven to be task dependent, varying even within the same study when the type of task is modified (Vicari et al., 2000). Furthermore, the responses required from the participants also varied across studies – free labeling tasks involve lexico-semantic abilities, matching tasks rely on perceptual faculties, whereas forced-choice tasks require categorization abilities (e.g., Bruce et al., 2000; Mancini et al., 2013). As different stimuli as well as cognitive faculties are tested across studies, drawing comparisons between the results is difficult.

Despite the heterogeneous results, there is a general consensus among researchers that the developmental trajectory of expression recognition varies among emotions. A substantial amount of data indicates that the expression of happiness is the first one to be accurately recognized, with 5-year-old children reaching identical recognition scores as adults (Durand et al., 2007; Gao & Maurer, 2010; Gross & Ballif, 1991; Herba & Phillips, 2004), even when the emotional faces are presented as fast as 500ms (Rodger et al., 2015). After happiness, studies have reported the emergence of recognition abilities for anger and sadness (Durand et al., 2007; Mancini et al., 2013; Montiroso et al., 2010) followed by the expressions of surprise, fear, and disgust (e.g., Mancini et al., 2013). In a recent study investigating facial expression recognition in late childhood, Mancini et al. (2013) observed steep improvements for neutral and sadness, whereas more gradual increases were found for anger, fear, and disgust. Using morphed faces varying in emotional content, Thomas et al. (2007) observed steep improvements from childhood to adolescence in the recognition of fear to neutral and fear to anger morphed faces, while an improvement for the recognition of the anger to neutral sequence was observed only later, from

adolescence onwards. These results imply that the developmental trajectories for fear and anger rely on distinct mechanisms.

In another study, using dynamic faces, Montirosso et al. (2010) examined the effects of emotional intensity on the categorization performance of the expressions of anger, sadness, fear, happiness, and disgust in 4- to 18-year-old observers. In line with previous studies (Herba et al., 2006), results have proven that labeling accuracy improved with the emotional intensity, although the effects of intensity varied across facial expressions. For instance, happy faces were equally well recognized at a low or high intensity. Among negative expressions, recognition rates were nearly identical for intense expressions, with fear however being better recognized than anger or sadness when presented at low intensity. Since the expression of fear signals potential dangers, the brain might be tuned to detect it more rapidly, even when presented in its subtle form (Montirosso et al., 2010; Plutchik, 1980a). Moreover, although fear was frequently confused with surprise, results have displayed that it was rarely confused with other negative emotions, probably because the action units commonly activated by fear (i.e., raised eyebrows) are perceptually more distinctive than those initially activated by angry or sad faces (Ekman & Friesen, 1978; Montirosso et al., 2010). Overall, this study revealed age-related increases in recognition performance for all expressions, except for disgust. Improvement in patterns were not uniform across emotions, accuracy rates for happiness and fear increasing faster than for anger and sadness. Recognition rates were also dependent on the intensity of the emotional expressions presented. Interestingly, although this study used dynamic expressions, its results were generally consistent with other research featuring static faces (Vicari et al., 2000; but see, Mancini et al., 2013, for disgust; Rodger et al., 2015).

In another emotional intensity study, Gao and Maurer (2010) determined children's thresholds for recognizing the six basic expressions presented with varying emotional intensity ranging from 0% (neutral) to 100% (apex). The results reported an early sensitivity to happy faces, consistent with previous findings (Camras & Allison, 1985; Durand et al., 2007; Vicari et al., 2000), while thresholds for negative expressions developed at different rates up to adulthood. Thresholds for fear, sadness, and disgust achieved adult-like levels earlier (10 years) than anger, which was the latest (after 10 years).

In line with these findings, a study by Rodger et al. (2015) very recently mapped the development of facial expression recognition from childhood up to adulthood by adopting an innovative psychophysical approach. Observers' perceptual thresholds for effectively detecting and labeling emotional faces were identified by using a threshold-seeking algorithm, enabling the manipulation of the quantity of signals provided by the faces (Figure 1.20A). The results revealed that the six basic expressions could be grouped into three distinct developmental

trajectories: the first included disgust, neutral, and anger exhibiting a steep increase with age; the second group consisted of sadness and surprise, two expressions that displayed more moderate improvements. The last group included expressions that were already well recognized at a very young age (happiness) or remained stable from childhood to adulthood (fear) (Figure 1.20B). The findings obtained from this study also outlined two important stages during the development of expression recognition, the first one spanning from 5 to 12 years and the second one from 13 years to early adulthood. These developmental stages mirror the results of previous studies, revealing noticeable improvements between 9 and 10 years as well as marked increases in recognition accuracy between 13 and 14 years (Kolb, Wilson, & Taylor, 1992; Somerville, Fani, & McClure-Tone, 2011).

Developmental stages as well as prominent improvements may be driven by structural and functional changes in the brain. For instance, substantial evidence has supported that the brain structures underlying emotional processing develop from childhood till late adolescence (e.g., Giedd, 2008; Giedd et al., 1999; Giedd et al., 1996; Shaw et al., 2008) with substantial increases in left and right amygdala volume between 7.5 and 18.5 years of age (Figure 1.21) (Giedd et al., 1996; Schumann et al., 2004). Other factors such as environmental and affective influences (Figure 1.22) (Bornstein et al., 2011; Bornstein et al., 2012; Pinel et al., 2014; Pollak & Kistler, 2002; Pollak & Sinha, 2002), cultural (Geangu et al., 2016), visual, or perceptual experiences (Bate et al., 2013; Calder et al., 2000; Gao, Maurer, & Nishimura, 2013) may also impact the development of facial expression recognition.

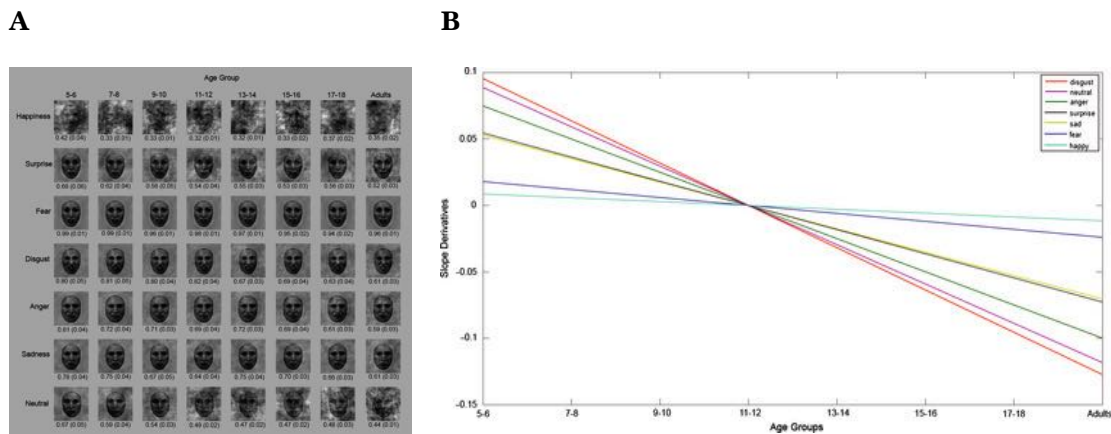


Figure 1.20. Fine-grained developmental mapping of emotional face processing. (A) Mean thresholds for each age group and each emotion category. Very low recognition thresholds can be observed for happiness, while recognition thresholds for fear remain high throughout development. Recognition thresholds for anger and neutral decrease with increasing age. (B) Fitted slopes of the General Linear Model indicating the steepness of decrease of recognition thresholds across ages. Disgust, neutral, and anger displayed a steep increase with age; sadness and surprise exhibited more moderate improvements, whereas happiness and fear remained stable from childhood to adulthood. Reprinted from Rodger et al. (2015).

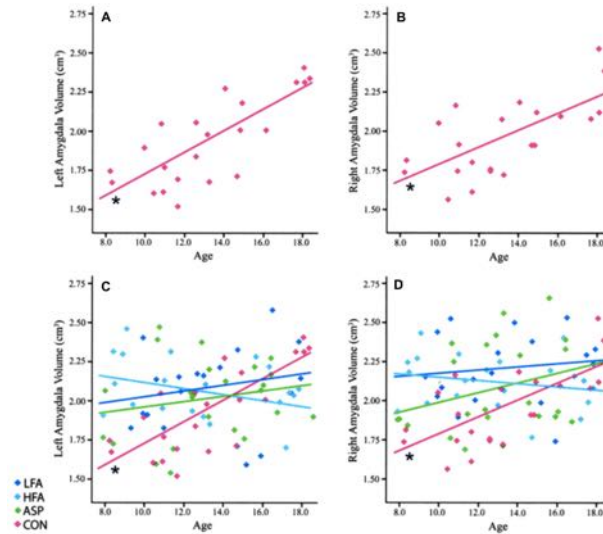


Figure 1.21. Neurobiological changes in the amygdala throughout development. The figure displays the linear regression scatterplot for amygdala volume as a function of age. The mean volumetric data revealed a positive correlation between (A) right- and (B) left-sided amygdala volume and age. This gradual increase in volume was however not observed in the (C) right- and (D) left-amygdala of autistic children. Adapted from Schumann et al. (2004).

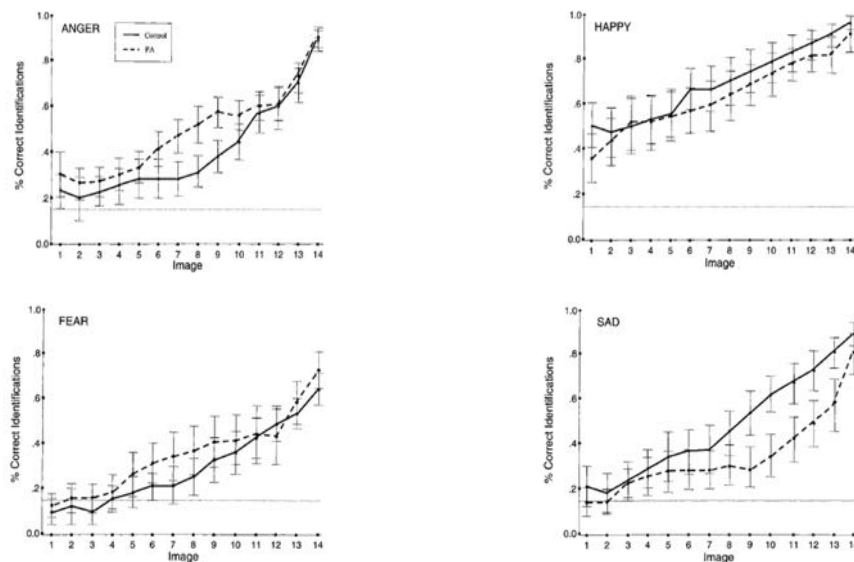


Figure 1.22. Affective influences in shaping expression recognition. Percentage of accurate identification rates for the expressions of anger, happiness, fear, and sadness in abused (dashed lines) and non-abused (solid lines) children as a function of the amount of information provided by the images (1: low; 10: high). Abused children required less information to accurately recognize the expressions of anger and fear compared to non-abused children. The opposite pattern was observed for sadness and happiness. Reprinted from Pollak and Sinha (2002).

Lifespan approaches. Studies relying on lifespan approaches to investigate the perception of emotional expression evidenced inverted u-shaped recognition trajectories, with younger and middle-aged adults performing the best, while children and elderly people were the least accurate (Horning, Cornwell, & Davis, 2012; Williams et al., 2009). Importantly, these lifespan studies revealed another fundamental shift in recognition performance occurring in late adulthood with normal aging.

Normal aging. Age-related gradual declines in the recognition of facial expressions have broad implications for the physical and mental well-being of elderly people as well as for their use of non-verbal communicative information. Although some previous reports have demonstrated an increase in the recognition of certain expressions with increasing age (Calder et al., 2003; Suzuki et al., 2006; but see, Sullivan & Ruffman, 2004a; 2004b), most studies have revealed that normal aging is followed by a gradual decline in emotion recognition. However, defining the exact emotion category that is affected is more challenging, because there is little consistency across studies in the methodologies used and the expressions tested – some relying only on a subset of expressions (Table 1.2). Using the Ekman and Friesen (1976) static posed faces, Calder et al. (2003) for instance observed a reduction in the recognition of fear and anger, and to a lesser extent sadness, in adults above 60, whereas the recognition of disgust further improved. By using dynamic faces providing richer and ecologically more valid depictions of the manner in which faces are perceived in daily life, Horning et al. (2012) revealed preserved recognition abilities for anger, surprise, and disgust, but marked deficiencies for happiness, fear, and sadness in the older age group spanning from 65 to 89 years of age. In a meta-analytic review of the literature of emotion recognition and aging, Ruffman et al. (2008) observed severe impairments in the recognition of anger and sadness across several modalities, such as voices, bodies, faces, or face-voice matching in normally aging people (mean age 70). Deficits were also observed for fear, surprise, and happiness, within some modalities, but not others. In contrast, the recognition of disgust was preserved (Ruffman et al., 2008, see also, Borod et al., 2004; MacPherson, Phillips, & Della Sala, 2002; Moreno et al., 1993; Phillips, MacLean, & Allen, 2002). However, this latter finding was neither confirmed by a more recent study evidencing a gradual decline in the recognition of disgust in very old people, aged 81 to 90 years (Williams et al., 2009), nor by previous reports revealing a deficit for this expression prevalent from 60 years onwards (Sullivan & Ruffman, 2004a; 2004b).

As shown in Table 1.2, the general pattern that arises from these different studies is that the recognition of some expressions remains relatively stable over the years, whereas the recognition of others drastically decreases with increasing age.

Table 1.2. Main studies investigating facial expression recognition across life span. Adapted from Isaacowitz et al. (2007).

Study	N	Age ranges by group	Stimuli	Answer format	A	D	F	H	Sa	Su	N
Borod et al. (2004)	36	25-35, 45-55, 65-75	450 posed faces (Moreno et al., 1993)	MC, verbal labels	NT	=	NT	=	=	=	=
Calder et al. (2003, Study 1)	48	18-30, 58-70	60 posed faces (Ekman & Friesen, 1976b)	MC, verbal labels	=	↑	↓	=	↓	=	NT
Calder et al. (2003, Study 2a)	227	17-30, 31-40, 41-50, 51-60, 61-70	60 posed faces (Ekman & Friesen, 1976b)	MC, verbal labels	↓	↑	↓	=	=	=	NT
Calder et al. (2003, Study 2b)	125	18-30, 31-40, 41-50, 51-60, 61-75	30 morphs of posed faces (Calder, 1996)	MC, verbal labels	↓	↑	↓	=	=	=	NT
Horning et al. (2012)	732	5-11, 12-17, 18-39, 40-64, 65-89	36 dynamic posed faces (Ekman & Friesen, 1976b)	MC, verbal labels	=	=	↓	↓	↓	=	NT
MacPherson et al. (2002)	90	20-38, 40-59, 61-80	49 posed faces (Matsumoto & Ekman, 1988)	MC, verbal labels	=	=	=	=	↓	=	NT
McDowell et al. (1994)	60	17-22, 65-90	50 posed faces, (Ekman & Friesen, 1976b)	MC, verbal labels	↓	NT	↓	=	↓	NT	↓
Moreno et al. (1993)	90	21-39, 40-59, 60-81	20 posed faces (Ekman & Friesen, 1976b)	MC, verbal labels	NT	=	NT	↑	↓	=	NT
Phillips et al. (2002)	60	20-40, 60-80	24 posed faces (Ekman & Friesen, 1976b)	MC, verbal labels	↓	=	=	=	↓	=	NT
Sullivan & Ruffman (2004a, Study 1)	61	20-38, 60-84	Eight animated morphs of posed faces (Ekman & Friesen, 1976b)	MC, verbal labels	↓	NT	=	=	↓	NT	NT
Sullivan & Ruffman (2004a, Study 2a)	56	18-29, 63-79	24 pairs of morphs of posed faces (Ekman & Friesen, 1976b)	MC, emotion matching task	↓	=	↓	=	↓	=	NT
Sullivan & Ruffman (2004a, Study 2b)	56	18-29, 63-79	24 emotional voices (verbal and nonverbal, Hobson, Ouston, & Lee, 1988)	MC, posed faces;	↓	↓	=	=	↓	=	NT
Sullivan & Ruffman (2004b)	48	20-46, 60-82	24 posed faces (Ekman & Friesen, 1976b)	MC, verbal labels	↓	↓	=	=	=	=	NT
Williams et al. (2009)	1000	6-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-91	72 posed faces (Gur et al., 2002)	MC, verbal label	↓	↓	↓	↓	↓	↓	↓
Zhao et al. (2016)	60	19-27, 68-78	56 posed faces (Tottenham et al., 2009)	MC, verbal label	↓	↓	↓	=	↓	↓	NT

Note. A = Anger; D = Disgust; F = Fear; H = Happiness; Sa = Sadness; Su = Surprise; N = Neutral; NT = Not tested; MC = Multiple choice.

1.2.5 Theoretical explanations for age-related expression recognition deficits

Age-related deficits in the recognition of facial expressions are commonly discussed in the context of three theoretical approaches: the socio-emotional selectivity theory, the cognitive aging view, and the neuropsychological approach of aging. These three perspectives will be now presented in more details.

According to the **socio-emotional selectivity theory**, aging emphasizes positivity effects (Carstensen, Fung, & Charles, 2003; Carstensen & Mikels, 2005; Mill et al., 2009; Ruffman et al., 2008), leading elderly people to preferentially attend and process positive emotional information as compared to negative one. These positivity effects have been attributed to an enhanced ability to avoid social conflicts with age, adapt to negative situations, and control and regulate emotions (Carstensen et al., 2003; Charles, Mather, & Carstensen, 2003). This theory posits that increasing age leads people to gain awareness of their own mortality, encouraging them to change their goals and motivation and direct their attention towards the emotionally positive aspects of their life. In fact, when pairs of photographs of negative (sadness and anger) and positive (happy) emotional faces were shown to young, middle, and old adults, the older participants exhibited an attentional bias against negative images and towards positive ones (Figure 1.23) (Charles et al., 2003).

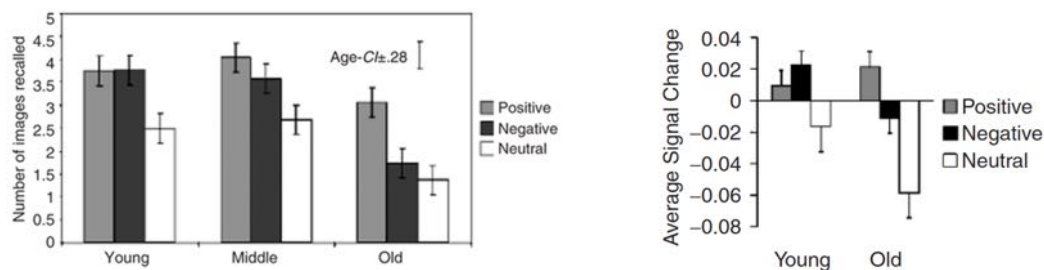


Figure 1.23. Positivity theory. *Left:* Number of positive, negative, and neutral expressions recalled by young, middle, and normal aging people. *Right:* Amygdala activation to positive, negative, and neutral expressions in young and elderly people. The average signal change indicates amygdala activation while participants were viewing each kind of expression. Adapted from Carstensen and Mikels (2005).

Consistent with these findings, Carstensen and Mikels (2005) have revealed a significantly greater activation of the amygdala in elderly people in response to positive as compared to negative images in relation to younger adults, although the amygdala had been previously presented as being central to the processing of both positive and negative stimuli

(Hamann et al., 2002). While the socio-emotional selectivity theory has been proposed as a potential explanation for older adults' perseveration in the recognition of positive emotional information, this approach has been limited by findings revealing that the recognition of certain positive expressions is not spared in elderly people. For instance, recognition deficits for happiness have been discovered in some previous studies (Brosigle & Weisman, 1995; Horning et al., 2012). Furthermore, this theory is not consistent with previous findings exhibiting increased recognition performance for the negative expression of disgust in elderly people (Calder et al., 2003; Suzuki et al., 2006). In light of these findings, Calder et al. (2003) have suggested that the attentional bias towards positive expressions should be reinterpreted as a reduced ability to process certain negative expressions, rather than an enhanced ability to regulate emotions with age. Additionally, Sullivan et al. (2007) have suggested that age differences in the recognition of positive and negative expressions can be explained by the distinct eye movement patterns employed by younger and older adults to process facial expressions. As shown in Figure 1.24, these authors have noted that elderly people sub-optimally used the diagnostic information provided by the eye region, focusing more on the mouth. This suboptimal use of the eyes is of particular interest, as negative expressions provide overall more information from the eye region (Sullivan et al., 2007; Wong, Cronin-Golomb, & Neargarder, 2005).

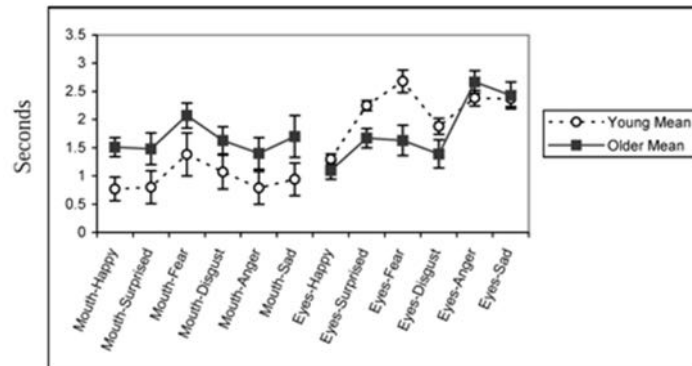


Figure 1.24. Eye movement patterns in young and elderly people. Time spent (in seconds) by young and older adults on the mouth and eye region for each emotion category. Reprinted from Sullivan et al. (2007).

According to **the cognitive aging theory**, a series of cognitive faculties, such as fluid intelligence (e.g., Horning et al., 2012; Sullivan & Ruffman, 2004a), processing speed (e.g., Orgeta & Phillips, 2007), verbal memory (e.g., MacPherson et al., 2002), or discrimination of visual information (e.g., Mill et al., 2009), are critical for the recognition of human facial expressions. These different cognitive abilities are known to decrease with age, independently or in combination with each other (Mill et al., 2009; Salthouse, 2004). Although age-related cognitive decline becomes greater and broader from the age of 50 onwards, previous evidence

has proven that the decline is already significant in adults between 18 and 50 years of age (Verhaeghen & Salthouse, 1997). For instance, relevant empirical evidence has revealed that processing speed as well as fluid intelligence gradually deteriorate in adults over 30 years, whereas the decline for memory linearly decreases from the age of 40 onwards (e.g., Salthouse, Atkinson, & Berish, 2003). Decreases in processing speed may affect expression recognition by reducing the number of facial features extracted from the emotional faces, if the faces are presented for a limited amount of time (e.g., Orgeta & Phillips, 2007; Suzuki & Akiyama, 2013; West et al., 2012; Zhao et al., 2016). Other faculties, however, such as verbal memory, vocabulary, or lexical knowledge have been noted to reach their performance peak later, between 40 and 60 years of age (e.g., Salthouse, 2004).

Although there is a general consensus among researchers that aging leads to a decline in facial expression recognition, previous reports have not provided compelling evidence supporting that these age-related impairments result from general cognitive decline (e.g., Horning et al., 2012; MacPherson et al., 2002; Ruffman et al., 2008). For instance, Horning et al. (2012) revealed that different cognitive faculties, such as processing speed, fluid intelligence, or memory contributed to the recognition of facial expressions in older adults, especially above the age of 45, but did not fully account for the gradual decline observed with increasing age. Furthermore, research by Keightley et al. (2007) and Vicari et al. (2000) indicated a relative independence between cognitive functions and emotion recognition, as participants' performance on various cognitive tests did not predict younger and older adults' accuracy scores on facial expression recognition tasks. Additionally, the general cognitive decline does not seem to be solely responsible for older adults' impairments in facial expression recognition, as these impairments would otherwise be expected to emerge earlier, as gradual declines are already observed beginning from the age of 30. Aside from the general cognitive decline, other factors, such as neuropsychological changes within the social brain may provide alternative, stronger, and more compelling explanations for the observed deficits.

According to a third theory, **the neuropsychological approach**, substantial evidence from patients and neuroimaging studies suggests that distinct brain regions underlie the processing of different emotions. For instance, the amygdala plays a critical role for the processing of fear (e.g., Adolphs & Tranel, 2004; Adolphs et al., 1994; Adolphs et al., 1995; Adolphs et al., 1999; Calder, Lawrence, & Young, 2001b; Phan et al., 2002; Phillips et al., 2003; Vuilleumier & Pourtois, 2007; Zhao et al., 2017), whereas the basal ganglia and insula underlie the processing of disgust (e.g., Calder et al., 2003). Other studies indicate that the cingulate cortex and the lateral orbitofrontal cortex (OFC) in combination with the amygdala are responsible for the decoding of anger (e.g., Blair & Cipolotti, 2000; Murphy, Nimmo-Smith, & Lawrence, 2003; Ruffman et al., 2008), while the recognition of happiness involves the amygdala

(e.g., Killgore & Yurgelun-Todd, 2004; Yang et al., 2002), frontal areas (e.g., Ruffman et al., 2008; Salloum et al., 2007), as well as the fusiform gyrus (e.g., Surguladze et al., 2005; Surguladze et al., 2003). Identifying sad faces increases and decreases neural responses in different brain regions within the limbic system, such as the amygdala (e.g., Adolphs & Tranel, 2004; Lennox et al., 2004; Ruffman et al., 2008), medial prefrontal cortices (Murphy et al., 2003), as well as the anterior cingulate cortex (e.g., Lennox et al., 2004; Murphy et al., 2003; Salloum et al., 2007). Finally, viewing surprised faces elicits activation in the amygdala and para-hippocampal gyrus as well as in the insula and post-central cortices (e.g., Zhao et al., 2017).

There is a general consensus among researchers that normal aging leads to structural changes in the brain. Some have argued that the pattern of expression recognition deficits observed in elderly people may be related to the pattern of age-related structural changes occurring in the brain areas subserving the processing of individual emotions (e.g., Calder et al., 2003; Ruffman et al., 2008). For instance, several studies have demonstrated that brain volume losses occur earlier in frontal areas, with evidence suggesting that among these frontal regions, the orbitofrontal cortex (OFC) might be even more vulnerable to aging, with more rapid and severe changes in this particular region. As reported by Ruffman et al. (2008), impairments of older adults in the recognition of anger may be primarily due to a degradation of the OFC, considering that this region is critical for the processing of angry faces (see above). Similarly, the marked deficits for the recognition of fear in the older adults (e.g., Calder et al., 2003; Horning et al., 2012) may be related to structural changes in the amygdala. Research has indeed revealed that this brain region undergoes atrophy with age (Mather et al., 2004; Ruffman et al., 2008) and becomes less reactive to negative stimuli with increasing age (Mather et al., 2004). In contrast, it is frequently argued that other brain regions, such as the insula or basal ganglia, critical for the processing of disgust, are less vulnerable to age-related degeneration. As reported by previous studies, preservation of these brain regions may be partially responsible for the intact ability of older adults in recognizing disgust (Calder et al., 2000; Horning et al., 2012; Ruffman et al., 2008; Suzuki et al., 2006).

Finally, previous fMRI studies comparing younger and older adults' abilities to categorize expressions did not only find differences in recognition performance but also differential brain activation patterns during the processing of emotional faces (e.g., Fischer, Nyberg, & Bäckman, 2010; Fischer et al., 2005; Gunning-Dixon et al., 2003; Tessitore et al., 2005). For instance, Gunning-Dixon et al. (2003) observed that the temporo-limbic areas, including the amygdala and surrounding regions, were activated in younger adults, while the anterior cingulate cortex, bilateral prefrontal, and parietal regions were selectively recruited in older adults for processing emotional faces. As suggested by the authors, reduced activity in the limbic areas in older adults may mirror age-related structural and functional changes within

these regions, an explanation that is consistent with the aforementioned behavioral and neuropsychological studies (e.g., Calder et al., 2003; Ruffman et al., 2008), with other fMRI studies (Fischer et al., 2010; Fischer et al., 2005; Raz, 2000), as well as with post-mortem reports evidencing age-related losses of brain volume and neuronal atrophy in the amygdala and surroundings regions (e.g., Geinisman et al., 1995; West, 1993). Importantly here, the neural responses observed in older adults to emotional faces (i.e., bilateral prefrontal and parietal activations) may imply an age-related reorganization of the neural substrates subtending emotional processing. Along this line, more recent research revealed decreased amygdala activation, but increased cortical activity (frontal and insular) in response to fearful faces in older compared to younger adults (Figure 1.25, Table 1.3) (Fischer et al., 2010). This age-related medial-temporal to neocortical shift may imply that negative emotional information (i.e., a fearful face) is processed more consciously with increasing age, as neocortical regions underpin the processing of regulatory and attentional mechanisms (e.g., Fischer et al., 2010; Ochsner et al., 2004).

Table 1.3. Differences in brain activation between young and old adults during the processing of fearful faces. Reprinted from Fischer et al. (2010).

Brain region	Hemisphere	Brodmann area	X	Y	Z	Cluster size in voxels	z-Value
<i>Hits > misses</i>							
Hippocampus	R		28	-10	-20	5	2.02
Amygdala	R		26	-6	-20	15	1.82
Amygdala	L		-16	-7	-12	1	1.81
<i>Young > old</i>							
Hippocampus	R		30	-26	-9	46	3.31
Hippocampus	L		-28	-26	-10	5	1.84
Amygdala	R		30	-3	-15	5	2.14
<i>Old > young</i>							
Insular cortex	L	13	-36	-28	18	95	3.58
Prefrontal cortex	R	9	26	44	30	33	3.37

Note: L = left; R = right. Spatial coordinates are for the maximum voxel activation(s) in a cluster of suprathreshold voxels that covers one or several brain regions and Brodmann areas ($p < .05$ uncorrected for the amygdala and the hippocampus, and $p < .001$ uncorrected for cortical areas). Coordinates are in millimetres and correspond to the stereotactic atlas of Talairach and Tournoux (1988).

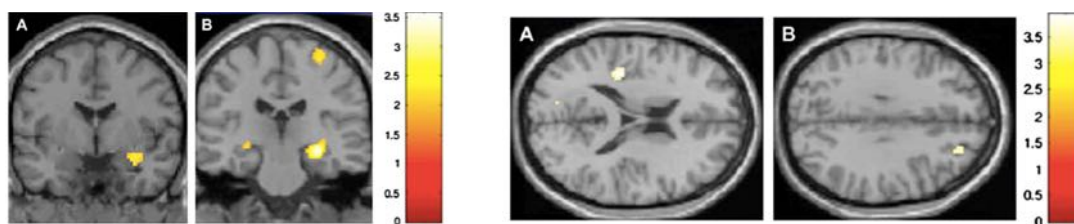


Figure 1.25. Brain activation to fearful faces in young and old adults. *Left:* Increased activation in (A) the right amygdala and (B) the bilateral hippocampus in young compared to old adults during the processing of fearful faces. *Right:* Increased activation in (A) the left insula and (B) the right prefrontal cortex in old compared to young adults during the same task. Adapted from Fischer et al. (2010).

After this developmental overview, we will subsequently note that putting motion into emotion can, in some specific situations, facilitate their recognition.

1.2.6 Adding motion to emotion: the role of dynamic cues in facial expression recognition

Our ability to rapidly detect changes in the facial composure of our conspecifics is key to the regulation of even the most basic social interactions. Most studies investigating facial expression recognition have relied on static images displaying the apex or highest state of an emotional expression. In everyday life however, natural expressions are dynamic events that evolve over time in some particular ways and convey critical information for adapted behaviors. As our visual system is steadily stimulated by dynamic cues, it appears reasonable to assume that we are adapted to more readily recognize expressions from moving faces. Although pioneering works in the field of emotion recognition have clearly demonstrated that we can infer other people's emotional state from static images (Ekman & Friesen, 1976b; 1978), the evidence is less consistent when it comes to determine whether we gain any perceptual advantage while viewing dynamic expressions.

This section will provide an overview of previous studies designed to evaluate the influence of facial motion on behavioral responses, such as recognition accuracy, response time, or perceived emotional intensity. Studies from the neuropsychological and clinical literature will be also briefly presented. The processing of static and dynamic faces in the brain will be approached in section 1.3.4.

Suboptimal situations. Over the last two decades, several studies have furnished evidence for a dynamic advantage in the recognition of facial expressions, particularly in suboptimal situations, when the information presented is limited, such as when the faces are presented as point-light displays, blurred, schematized, or degraded (for a recent review, see Krumhuber, Kappas, & Manstead, 2013). In his pioneering work, Bassili (1978) investigated observers' ability to recognize expressions from point-light moving faces (Figure 1.26). His findings revealed that participants were more effective in recognizing facial affects from moving than static displays. These results were further corroborated by Bruce and Valentine (1988) in a similar study investigating gender, identity, and emotion categorization using biological motion. Results revealed that the recognition of facial expression was much more accurate with biologically moving faces as compared to static displays.

More recently, relying on computer-animated faces, Wallraven et al. (2008) revealed that motion cues enhanced the recognition of facial affects when texture or shape information were systematically degraded or blurred. If dynamic cues were not provided, degrading face information significantly affected expression recognition, suggesting that temporal properties compensate or even eliminate the deleterious effects of degrading or diminishing shape and

texture information. A dynamic advantage for facial expression recognition was also found when synthetic images drawing depictions of facial muscle movements, were used (Wehrle et al., 2000). Interestingly, by comparing expression recognition using schematized and natural faces, Kätsyri and Sams (2008) and Ehrlich, Schiano, and Sheridan (2000) discovered a dynamic advantage only when synthetic expressions were presented; no significant differences in recognition accuracy when natural expressions were shown. Correspondingly, Cunningham and Wallraven (2009b) investigated participants' ability to recognize facial expressions using images with varying spatial resolutions (point-light faces, wireframe frames, and animated full-surface faces). Recognition performance was overall higher with dynamic expressions as compared to static ones, particularly for point-light faces, with the lowest spatial resolution. Moreover, reducing geometry or connectivity information did not affect the recognition of naturally evolving dynamic events; whereas this procedure strongly impacted the recognition of static expressions.

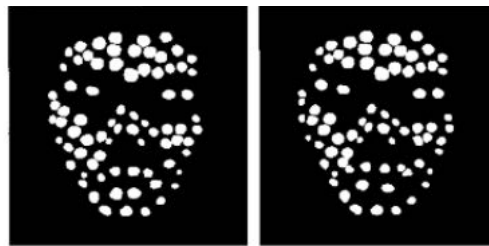


Figure 1.26. Point-light static displays. The left panel represents a neutral expression, whereas the right panel shows a happy face. As shown in early studies (Bassili, 1978; Bruce & Valentine, 1988), identifying facial expressions from static point-light displays is particularly difficult. Reprinted from Anderson and Fisher Anderson (2007).

Dynamic cues do not only mitigate the negative effects of degraded information, but also provide critical information when subtle expressions are presented (Ambadar et al., 2005; Bould et al., 2008). For instance, Ambadar et al. (2005) investigated the role of motion in the recognition of subtle facial expressions, by comparing the identification rates of healthy observers in four different experimental conditions (dynamic, static, multi-static, and first-last frame). The main aim of the study was to examine whether differences in recognition rates between static and dynamic conditions were due to motion itself or other properties such as temporal information (perception of change), additional information carried by dynamic events (multi-images), or motion-based processing strategies (configural vs. featural). The four experimental conditions started all with a mask presented for 200 seconds and then followed by either a static expression (static condition); a sequence of three to six frames at a rate of 30ms

per frame (dynamic condition); dynamic frames presented for 500ms, but with each frame interspersed by a 200ms noise mask to disrupt the perception of motion (multi-static condition); or the first and last image of the sequence (first-last condition) (Figure 1.27). The results revealed a robust effect of motion on the recognition of subtle facial expressions, for all expressions except for happiness. By designing a multi-static condition, presenting the same frames as the ones displayed in the dynamic condition, but interspersed with 200ms noise masks to attenuate the perception of motion, researchers ruled out the possibility that the major role played by dynamic stimuli was due to additional information provided by multiple images. Rather, the authors suggested that the beneficial effects of dynamic expressions stemmed from the properties of the moving faces themselves. This was further evidenced in the fourth experimental condition of the study (first-last frame condition), presenting only the first and last image of the sequence, a procedure that induced the perception of motion in participants. The recognition rates were significantly higher in both dynamic conditions (dynamic and first-last frame conditions) relative to the static ones (static and multi-static conditions). These findings enabled the researchers to conclude that the temporal characteristics inherent to the properties of the dynamic stimuli themselves explain the beneficial effects of motion, by facilitating the observers to perceive the direction of change.

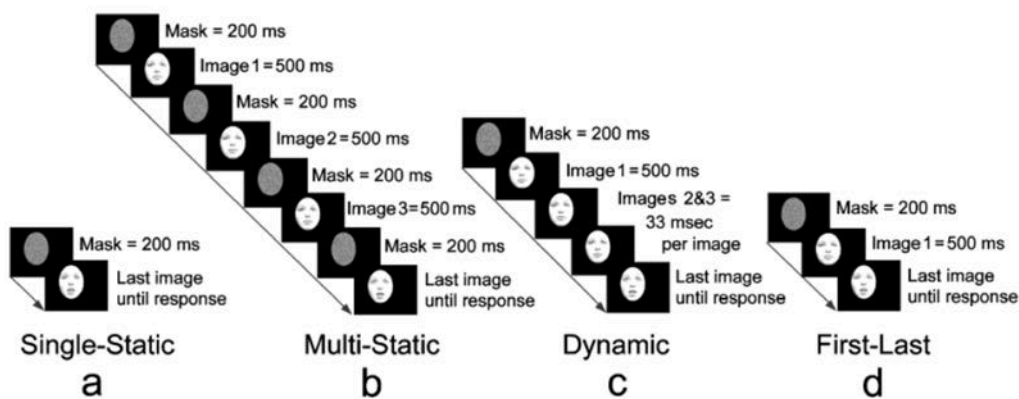


Figure 1.27. Investigating the recognition of subtle expressions with static and dynamic stimuli. Schematic representation of the stimuli used in the four experimental conditions. Reprinted from Ambadar et al. (2005).

Along the same line, Bould, Morris, and Wink (2008) further investigated the role played by dynamic temporal information in recognizing subtle emotional expressions. Three conditions were tested. In the first condition, dynamic sequences beginning from a neutral state and evolving into a subtle emotion were presented. In the second condition, nine static frames

taken from the dynamic sequences were displayed. The third condition presented only the first neutral and the last fully articulated frames. Results revealed that expression recognition rates were significantly higher in the first condition, presenting naturally evolving subtle expressions, than in both the second and third conditions. The authors then further examined the role of temporal dynamics by disrupting the naturally unfolding speed of the expressions (e.g., speeding up, slowing down). These methodological manipulations had an impact on performance, lowering significantly the recognition rates. Overall, this research revealed that both the perception of change as well as the speed of activation of the different facial muscles involved in the production of natural dynamic expressions enhance the recognition of subtle emotions. Contrary to Ambadar et al.'s (2005) assumption, Bould, Morris, and Wink (2008) revealed that the dynamic advantage may not only be due to an enhancement of the perception of directional change (first-last < dynamic) but may also stem from the perception of the temporal progression of information (i.e., the progressive activation of the specific muscular movements associated with the unfolding of a specific expression). This study, therefore, suggests that the temporal information may be more important than what was initially assumed by Ambadar et al. (2005).

Interestingly, a number of other studies have evidenced that the specific temporal properties inherent to the unfolding speed (Bould & Morris, 2008; Bould et al., 2008; Kamachi et al., 2001) or rise time of an expression (Recio, Schacht, & Sommer, 2013) provide some crucial information for its recognition. For instance, Kamachi et al. (2001, Experiment 2) and Recio et al. (2013) evidenced that some expressions were more accurately identified when they developed rapidly, whereas others (e.g., sadness) were better recognized from slow sequences or from their static versions (Recio et al., 2013). In a recent study, Recio and colleagues (2013) investigated whether the speed of onset and evolution of an expression had an influence on recognition performance by using artificially generated computer-morphs unfolding at three different speeds: 200, 500, and 900ms. The results revealed an overall beneficial effect of dynamic expressions, particularly for disgust, better recognized in fast and moderate sequences, and happiness, being better classified in all three sequences. In contrast, recognition rates for sadness were higher in the static condition. The latter finding supports previous evidence revealing that this expression is better recognized in slow sequences or static presentations (Bould et al., 2008; Kamachi et al., 2001). All other expressions (anger, fear, and surprise) were well recognized across all three velocities. Unfolding speeds between 200 and 500ms appear to be therefore optimal for the recognition of all expressions, excepting sadness.

Finally, by distorting the temporal direction of an expression unfolding over time (i.e., presenting the expression in the random or reverse order), Cunningham and Wallraven (2009a) revealed that the mere presence of motion signals did not trigger the dynamic advantage. This evidence suggests that human observers are sensitive to the temporal information embedded in

the contiguous progression of an expression unfolding over time (but see, Gold et al., 2013) and that the dynamic advantage may not only be due to the perception of directional change as suggested by Ambadar et al. (2005), but also to the temporal characteristics conveyed by naturally unfolding expressions (Bould et al., 2008; Recio et al., 2013).

Clinical and neuropsychological conditions. Dynamic cues not only offer processing benefits in suboptimal situations, but also lead to a noticeable improvement in clinical and neuropsychological conditions. For instance, previous studies have demonstrated that dynamic presentations significantly facilitated expression recognition in depressive and schizophrenic patients (e.g., Atkinson et al., 2004; Kan et al., 2004; Schaefer et al., 2010), patients suffering from pervasive developmental disorder (Uono, Sato, & Toichi, 2010), children with mental retardation (Harwood, Hall, & Shinkfield, 1999), autistic patients (Back, Ropar, & Mitchell, 2007; Gepner, Deruelle, & Grynfeldt, 2001; Tardif et al., 2007), as well as brain-damaged patients (Adolphs, Tranel, & Damasio, 2003; Humphreys, Donnelly, & Riddoch, 1993; Richoz et al., 2015).

In an early neuropsychological study, Humphreys, Donnelly, and Riddoch (1993), reported the case of an agnosic patient, HJA, with severe lesions to the ventral, occipital, and temporal lobes. HJA was considerably impaired when asked to judge identities and expressions from still images. In contrast, his recognition performance was within the normal range when exposed to a subset of facial expressions (i.e., smiling, frowning, or surprise) depicted as moving light-dots. Correspondingly, Adolphs, Tranel, and Damasio (2003) investigated the ability of a single case of acquired agnosia, patient B – with multiple and extensive bilateral brain lesions – to categorize static faces displaying one of the six basic expressions. The results revealed that he was greatly impaired in categorizing all the static expressions, with the exception of happiness (Figure 1.28A). When asked to categorize the dynamic expressions executed by an experimenter, the patient correctly labelled with maximum accuracy, every emotion excepted disgust (Figure 1.28B). As suggested by the authors, this latter impairment may be accounted for by his set of lesions, including the insula, a brain structure that has been proven to be critical for the recognition of disgust (Calder et al., 2000). Critically, this neuropsychological evidence suggests a functional dissociation between static and dynamic face discrimination tasks and supports the assumption that different neural substrates underpin the processing of static and dynamic stimuli (Duchaine & Yovel, 2015; Pitcher, Duchaine, & Walsh, 2014; Tomlinson et al., 2006) (for more details on the different brain regions involved in the processing of static and dynamic faces, see section 1.3.4).

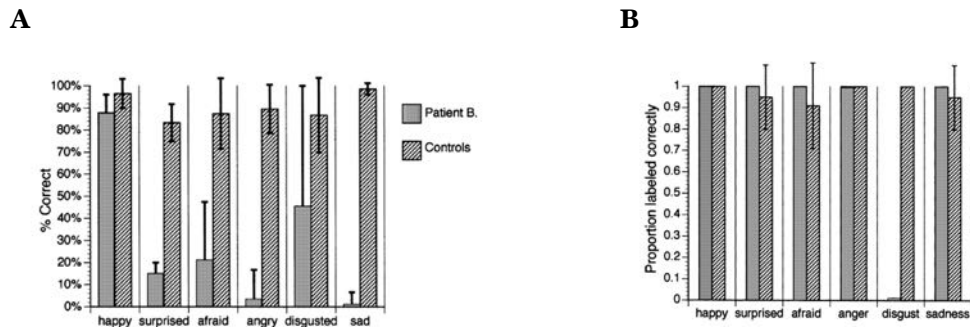


Figure 1.28. Neuropsychological evidence for a dynamic advantage. Patient B and healthy controls' recognition performance for the static and dynamic expressions of "happy", "surprise", "afraid", "anger", "disgust", and "sad" (Adolphs et al., 2003). **(A)** Mean accuracies for the static images. **(B)** Proportion of dynamic expressions correctly labelled by the patient and healthy controls. Adapted from Adolphs et al. (2003).

Optimal situations. While several studies have revealed a dynamic advantage for the recognition of facial expressions in suboptimal situations and clinical conditions, others have suggested that the beneficial effects of motion are minimal or non-existent when static information is readily available (Fiorentini & Viviani, 2011; Gold et al., 2013, for a summary, see Table 1.4) or when the expressions presented are of higher intensity (Bould & Morris, 2008; Kamachi et al., 2001, Experiment 2).

For instance, Fiorentini and Viviani (2011) investigated the dynamic advantage hypothesis by using linearly morphed graded blends of pairs of dynamic expressions (e.g., anger-fear, fear-sadness, happiness-disgust) unfolding from a neutral state to the apex. An additional set of static stimuli was generated by taking the apex of the dynamic movies. In a two-alternative forced-choice task, participants were asked to determine which expression the blended stimulus corresponded more to (i.e., which expression was predominant). The results revealed that the response accuracy was similar in both conditions, whereas reaction times were shorter in the static condition, although most responses in the dynamic condition were given before the blended expressions reached their highest state. According to the authors, this latter finding may suggest that the temporal information conveyed by dynamic stimuli may compensate for their incomplete presentations.

In another study, Gold et al. (2013) relied on a Bayesian integrator, an ideal observer model, to measure the amount of information carried by static, dynamic, and shuffled (temporally randomized) stimuli over time. Their results revealed that moving faces did not facilitate expression recognition when the amount of discriminative information conveyed by the expressions was considered. In other words, dynamic stimuli did not offer processing

advantages to the observers as compared to static images (Gold et al., 2013, for more details, see [CHAPTER 2](#) and [CHAPTER 5](#)).

The non-rigid muscular movements associated with the temporal unfolding of an expression provide additional cues that are unavailable from static faces, allowing observers to perceive the direction of change and disentangle the ambiguity triggered by insufficient or subtle information. In optimal situations and with expressions of high intensity, the visual system seems to be powerful enough to efficiently categorize static emotional expressions, leaving only little scope for improvement when dynamic movies are shown. However, dynamic cues are beneficial *only* when static information is insufficient or lost, compensating for the deleterious consequences of degraded or subtle information. In clinical conditions, suboptimal (and sometimes slower) processing takes place and dynamic expressions may force patients to shift their attention to different facial features, helping them to constantly reassess and modify their original hypothesis as supplementary information is added over the course of time. As a consequence, an advantage for dynamic faces is observed in those populations.

Table 1.4. Main studies comparing the recognition of static and dynamic expressions.
Adapted from Alves et al. (2013).

Study	Participants	Method	Stimuli	Main findings
Adolphs et al. (2003)	Brain damage (acquired agnosia)	Behavioral	36 posed faces (Ekman & Friesen, 1976b), 6 basic expressions; the dynamic expressions were performed by an experimenter	Dynamic > Static (with the exception of disgust)
Ambadar et al. (2005)	Healthy participants	Behavioral	6 basic expressions in their subtle form; expressions were taken from the Cohn-Kanade facial expression database (Kanade, Cohn, & Tian, 2000)	Dynamic > Static (the important role of motion in the perception of change)
Biele and Grabowska (2006)	Healthy participants	Behavioral	Anger and happiness, 4 posed faces from the Montreal set of facial expression of emotion (Beaupré & Hess, 2005)	Dynamic > Static (higher intensity ratings for dynamic vs. static expressions and anger vs. happy expressions)
Fiorentini and Viviani (2011)	Healthy participants	Behavioral/Threshold model	6 basic expressions from a database of stimuli created by the authors	Dynamic = Static
Gold et al. (2013)	Healthy participants	Behavioral/Ideal observer approach	6 basic expressions, 48 posed faces; database of stimuli created by Gold et al. (2013)	Dynamic = Static (for efficiency and thresholds)
Humphreys et al. (1993)	Brain damage (visual object agnosia)	Behavioral	Static condition: 30 posed faces (Ekman & Friesen, 1976b) displaying the expressions of happiness, worried and angry. Dynamic condition: moving light-dots displaying the expressions of frowning, smiling and surprise	Dynamic > Static
Kamachi et al. (2001)	Healthy participants	Behavioral	Morphed expressions (7 different expressions) from the ATR face database (Kamachi et al., 2001)	Dynamic ≠ Static (influence of the unfolding speed on expression recognition. Happiness and surprise were better recognized with dynamic stimuli while sadness was better recognized from slow sequences or static images).
Kätsyri et al. (2008)	Asperger Syndrome	Behavioral	16 posed faces, 4 expressions (anger, disgust, fear, happiness) from a database created by (Kätsyri, 2006)	Dynamic = Static
Krendl & Ambady (2010)	Elderly people	Behavioral	Static condition: 4 expressions (happiness, sadness, anger, fear) Dynamic condition: 2 s video-clips of positive/negative dominant expressions	Dynamic > Static
Recio et al. (2011)	Healthy participants	Behavioral/EEG	50 posed faces, neutral, anger and happiness	Dynamic > Static (dynamic vs. static expressions were recognized faster and more accurately)
Richoz et al. (2015)	Brain damage (acquired prosopagnosia)	Behavioral	Static and dynamic reconstructed mental models of the six basic expressions of the prosopagnosic patient (Yu, Garrod, & Schyns, 2012)	Dynamic > Static (for all the expressions with the exception of fear)
Uono et al. (2010)	Pervasive Development Disorder	Behavioral	Neutral, fearful and happy faces from Ekman and Friesen (1976) standard set of facial expressions	Dynamic > Static (dynamic expressions were perceived as being more intense than static images).

1.3 FACE PROCESSING IN THE BRAIN

In the last section, I gave an overview of some aspects of facial expressions of emotion, the way we perceive them throughout lifespan, and whether the dynamic properties of human faces offer significant processing advantages for their recognition. Here onwards, I will present electrophysiological, neuropsychological, and functional neuroimaging studies elucidating the neural bases of human face processing. I will also describe some of the most influential models of face processing that have guided more than two decades of research and discuss recent work (Duchaine & Yovel, 2015; Rossion, 2015), suggesting modifications to these prominent models in light of more recent findings.

1.3.1 A distributed neural network

Human face processing is an extremely complex ability that involves a vast number of cortical regions in the brain. Electrophysiological studies in monkeys as well as neuropsychological reports from brain-damaged patients have provided critical knowledge and important insights into the distributed neural network involved in face perception.

Electrophysiological and fMRI studies in non-human primates. In 1969, pioneering work based on single-cell recordings have evidenced face-selective neurons in the inferior temporal visual cortex (IT cortex) of rhesus macaques (Gross, Bender, & Rocha-Miranda, 1969). Commonly known as IT cells, these neurons, displaying stronger responses to faces than to any other objects, have been subsequently identified in many other studies providing compelling evidence in support of the existence of face-selective areas in the non-human primate brain (Baylis, Rolls, & Leonard, 1987; Desimone et al., 1984; Gross, Rocha-Miranda, & Bender, 1972; Perrett, Rolls, & Caan, 1982; Rolls, 1984). Coupling optical imaging (for more details on this technique, see [Box 1](#)) with single-unit recordings, Wang et al. (1996) further revealed columnar organization and regional clustering of these face-selective IT cells in the infero-temporal cortex of monkeys. In addition, findings from fMRI studies have evidenced three highly-selective face-patches in the ventral prefrontal cortex of macaque monkeys (Tsao et al., 2008b), with a face-selective cluster in the orbito-frontal cortex being particularly tuned to emotional faces as compared to non-face objects (Figure 1.29). Other scattered clusters with proportions of face-selective neurons varying between 20 and 70% have been reported in the upper and lower banks of the superior temporal sulcus (STS) of monkey brains (Bell et al., 2011; Tsao et al., 2006), with the largest face-selective patch containing up to 97% face-selective neurons (Tsao et al., 2006). However, although face-selective clusters have been described in the non-human primate

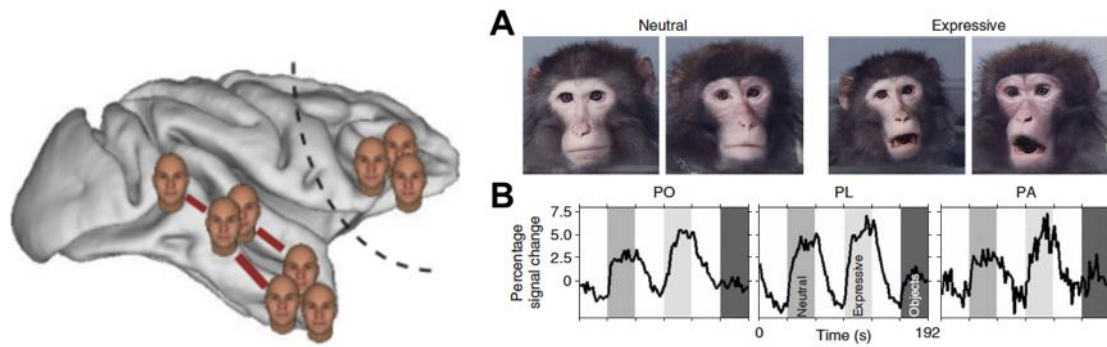


Figure 1.29. Three face-selective patches in the ventral prefrontal and temporal cortex of macaque monkeys. *Left:* fMRI findings evidenced three face-selective clusters in the ventral prefrontal cortex of four macaque monkeys (Tsao et al., 2008b). Reprinted from Backer (2008). *Right:* (A) Monkey faces displaying either neutral or expressive faces were presented to the macaques. (B) Duration of signal change to objects (dark gray); expressive (light gray) or neutral faces (medium gray) in the PO (Prefrontal Orbital Cortex), PL (Prefrontal Lateral Cortex), and PA (Prefrontal Arcuate) of macaque monkeys. Stronger responses to expressive than to neutral faces were observed in all three face-selective patches, with a highly significant difference observed in the PO. Adapted from Tsao, Schweers, Moeller, and Freiwald (2008b)

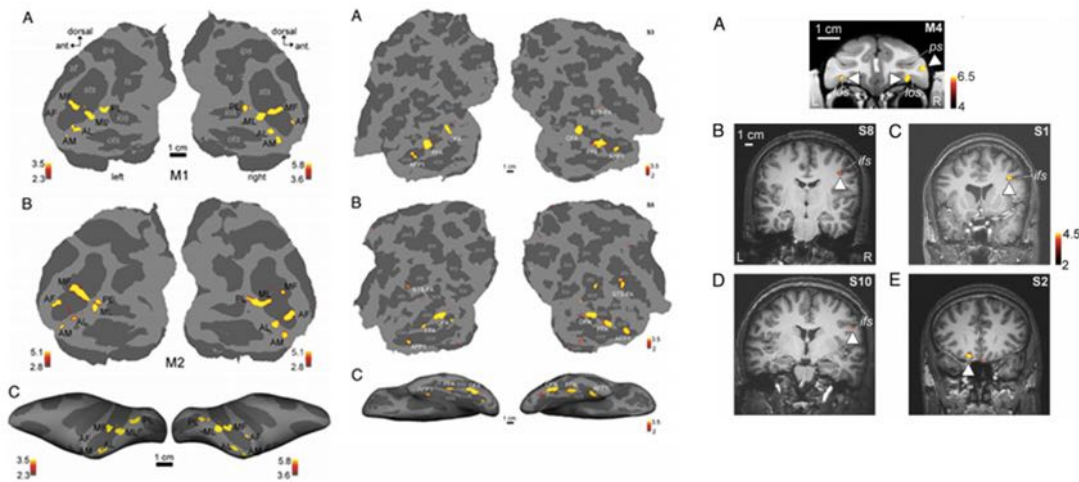


Figure 1.30. A comparison of face-selective clusters in macaque and human brains. *Left:* (A and B) Face-selective patches in two macaques. (C) A lateral view presenting both hemispheres of the macaque brain. *Center:* (A and B) Face-selective areas in the left- and right-sided temporal lobes of two human brains. (C) A ventral view presenting both hemispheres of the human brain. *Right:* Pre-frontal face-selective regions were activated bilaterally in the macaque brain (M4, top) while a strong right-hemispheric dominance was observed in human brains (S8, S1, S10). Adapted from Tsao, Moeller, and Freiwald (2008a).

brain, their correspondences to human face-selective areas remain unclear (Figure 1.30) (Rossion, 2015; Tsao, Moeller, & Freiwald, 2008a; Yovel & Freiwald, 2013). According to Duchaine and Yovel (2015) future research identifying homologies between face-selective patches in macaques (i.e., with single cell-recordings) and human face-selective regions are needed in order to better understand the neural basis of face perception. The extent to which single-cell findings in monkeys are informative for the understanding of human face processing remains however debatable.

Prosopagnosia. In human brains, one of the most striking evidence that some cortical areas are selectively involved in face perception has been demonstrated by prosopagnosic patients. These patients suffer from an inability to recognize familiar faces, despite a relatively preserved ability to identify other objects, normal intellectual functioning, and no other perceptual impairments (e.g., Farah, 1990; Haxby, Hoffman, & Gobbini, 2000; Rossion et al., 2003; Sergent & Signoret, 1992).

Although the first clinical observations had already been reported during the 19th century (Quaglino, 1867; Wigan, 1844), the term *Prosopagnosia*, describing this unique clinical condition, was introduced by Bodamer in 1947 (Bodamer, 1947). This face deficit is a very rare and spectacular condition that can be found in less than 1% of the clinical conditions following brain damages (Sergent & Villemure, 1989). Despite its rarity, acquired prosopagnosia has attracted great attention in the neuropsychological literature, and an impressive body of work has been dedicated to its understanding (Bukach et al., 2006; Bukach et al., 2008; Busigny et al., 2010a; 2010b; Busigny & Rossion, 2010; Caldara et al., 2005; Orban de Xivry et al., 2008; Ramon, Busigny, & Rossion, 2010; Rossion et al., 2003). Prosopagnosic patients are severely impaired in recognizing familiar faces, including faces of relatives, friends, famous people, or even their own (Damasio, 1985), despite no other obvious memory or visual deficits and unimpaired abilities to discriminate faces from non-face objects (for some case reports, see [Box 7](#)). This spectacular disorder is usually caused by bilateral lesions in the occipito-temporal areas (e.g., Damasio, Damasio, & Van Hoesen, 1982; Farah, 1990; Landis et al., 1988; Sergent & Signoret, 1992), although unilateral lesions in the right hemisphere are sufficient to trigger the syndrome (Figure 1.31) (Landis et al., 1988; Sergent & Signoret, 1992).

The clinical and anatomical conditions of prosopagnosic patients offer new and interesting insights into the normal functioning of human face processing. For example, the occurrence of a double dissociation between the ability to recognize familiar and unfamiliar faces (e.g., Malone et al., 1982) as well as the neurological segregation between face recognition and lip reading (Campbell, Landis, & Regard, 1986) clarify the sub-functions and independence of key areas within the face processing architecture.

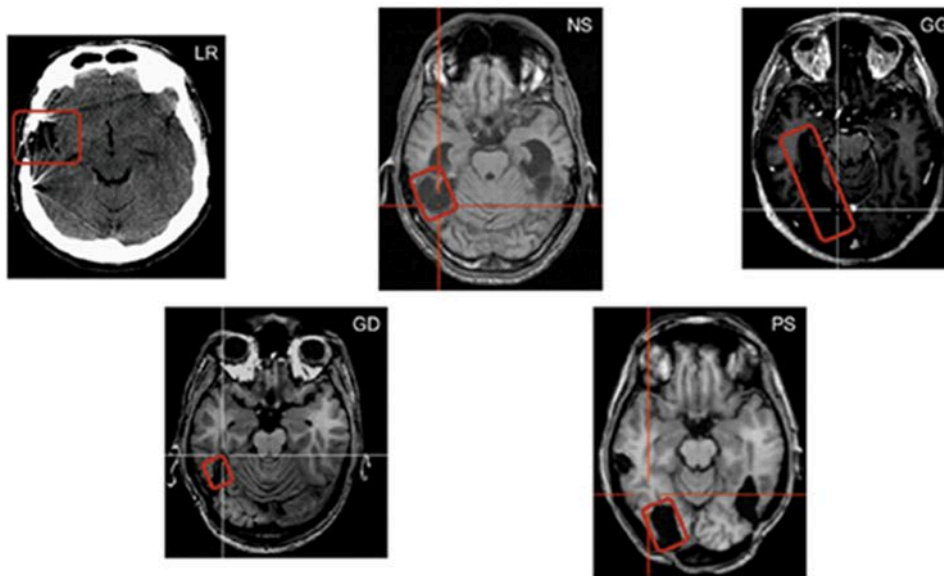


Figure 1.31. Five single-cases of acquired prosopagnosia, LR, NS, GG, GD, and PS and their corresponding brain lesions. Following a car accident, LR suffers from lesions to the inferior and anterior right temporal lobe (Bukach et al., 2006). While on his bicycle, NS got hit by a car leaving him with extensive lesions in the occipito-temporal junction bilaterally (Delvenne et al., 2004). GG suffers from right unilateral damage to the occipital lobe, fusiform, and para-hippocampal gyrus as a consequence of a cerebral vascular accident (Busigny et al., 2010b). GD developed unilateral left prosopo-metamorphopsia following the birth of her child. This clinical condition leads to severe distortions of the left-half of previously seen or well-known faces (Trojano et al., 2009). PS underwent a closed-head injury leading to multiple hemorrhages in the bilateral occipito-temporal areas (for a case description, see [Box 7](#)) (Rossion et al., 2003). Reprinted from Rossion (2015).

Box 7 – Three Cases of Acquired Prosopagnosia – Case Descriptions

GG. GG is a right-handed retired engineer who suffered from an ischemic infarct in 2002, leading to right-hemispheric damages in the occipital lobe, fusiform, and parahippocampal gyrus (Figure 1.32A, Box 8). Following his lesions, GG had difficulties in recognizing previously seen familiar faces (e.g., neighbors, friends, colleagues), famous faces (e.g., actors, athletes), and suffered also from left lateral homonymous hemianopsia and topographical disorientation. His neuropsychological assessment revealed no perceptual deficits. His ability to recognize objects was also preserved, even when holistic-processing or fine-grained analyses were required (Busigny et al., 2010b).

NS. NS got hit by a car in 1991 while cycling, causing major lesions in parietal, temporal, and occipital areas in the left hemisphere and temporo-occipital areas in the right hemisphere (Figure 1.32B). After 23 days in a coma, NS's clinical description evidenced severe deficits in recognizing faces and objects. In addition, he also displayed transcortical aphasia, severe dyslexia, apraxia, anosognosia, and an anterograde amnesia. Neuropsychological examinations two years after his accident revealed important improvements thanks to clinical rehabilitation, although most of his performance remained under a healthy control average (Pesenti et al., 2000). More specifically, NS was within the normal range for primary visual perception tasks and exhibited good performance in calculation and number processing. In contrast, he showed marked deficits when asked to name objects based on their verbal description and had striking difficulties in face recognition tasks (Delvenne et al., 2004; Pesenti et al., 2000).

PS. PS is a *pure* case of acquired prosopagnosia, who underwent a closed-head injury 25 years ago, leaving her with major lesions in the left mid-ventral and right inferior occipital cortex and minor damage to the right middle temporal gyrus and left posterior cerebellum (Rossion, 2008; for a complete case report, see Rossion et al., 2003; and for an exhaustive anatomical description, see Sorger et al., 2007) (Figure 1.32C). She recovered from cognitive deficits following psychological rehabilitation and medical treatment, apart from a severe deficit in recognizing familiar faces, including her own and those of her family members. Moreover, as revealed by a response classification technique (Bubbles) (for more details, see Caldara et al., 2005; Gosselin & Schyns, 2001), PS does not use optimal information from the eye region to identify familiar faces, but relies instead on the lower part of the face, including the mouth and external contours (Caldara et al., 2005). PS's low-level vision is well preserved with a good visual acuity in both eyes, except from a small right paracentral scotoma. She can categorize a face as a face, distinguishes faces from other classes of stimuli, and has normal perception and recognition of objects (Rossion et al., 2003). However, as a consequence of her lesions, she is not as good as controls in categorizing gender (Rossion et al., 2003) and static facial expressions (Richoz et al., 2015). As reported by neuropsychological assessments, she is severely impaired on the Benton Face Matching Test (Benton & Van Allen, 1972) as well as on the Short Recognition Memory Test for Faces (Rossion et al., 2003; Warrington, 1984). PS is a very cooperative patient, without cognitive impairments or attention deficits. She therefore represents an exemplary case to investigate the underlying architecture of normal face processing. The second experimental contribution of this thesis aims to investigate her ability to recognize static and dynamic facial expressions of emotion (CHAPTER 4).

Box 8 – Three Cases of Acquired Prosopagnosia – Lesion Localizations

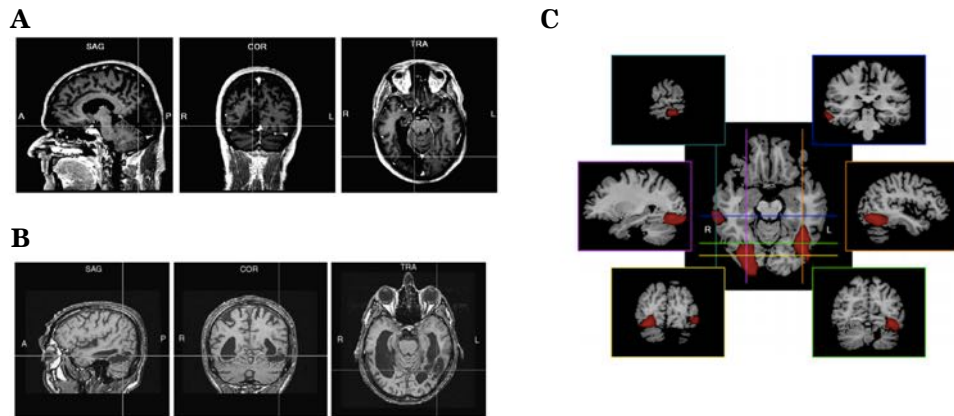


Figure 1.32. Lesion localizations. (A) GG's MRI images reveal brain damages to the occipital lobe, fusiform and parahippocampal gyrus in the right hemisphere. Reprinted from Busigny et al. (2010b). (B) NS's structural MRI showing lesions in the occipitotemporal junction. Reprinted from Delvenne et al. (2004). (C) T1-weighted MRI based reconstructions of PS's lesions. Reprinted from Sorger et al. (2007).

Face-selective brain regions. The first face-selective brain region located in the middle, fusiform gyrus and commonly named Fusiform Face Area (FFA), has been identified by fMRI in 1997 using a face localizer approach (Kanwisher et al., 1997). Nevertheless, it is worth emphasizing that early PET studies had already identified this region at the functional level in 1992 (Sergent et al., 1992), even if its activations were broader compared to those isolated with fMRI. This area, representing a key module for efficient face processing (Grill-Spector, Knouf, & Kanwisher, 2004), has been noted to elicit stronger response to faces than houses, hands, or scrambled depictions of faces (Kanwisher et al., 1997; McCarthy et al., 1997). Substantial evidence has also suggested that this region is involved in the processing of objects of expertise (i.e., objects for which one has an extensive expertise) (Bilalic et al., 2011; Gauthier et al., 2000a; Gauthier et al., 1999; Grill-Spector et al., 2004). Another region located in the lateral inferior occipital gyrus, referred to as the Occipital Face Area (OFA), also represents a key area for the processing of face information (Gauthier et al., 2000b; Haxby et al., 1999). As reported by neuropsychological evidence, both the right FFA, and the right OFA appear to be necessary for normal face processing (Gauthier et al., 2000b; Rossion et al., 2003). Finally, the posterior part of the Superior Temporal Sulcus (pSTS) has been further defined as another face-selective brain area, playing a crucial role in human face processing (Hoffman & Haxby, 2000; Kanwisher et al., 1997; Puce et al., 1998). According to Haxby et al. (2000), these core regions are connected to an extended system, including the amygdala, the auditory cortex, and the intraparietal sulcus,

contributing to the processing of different aspects of faces. As stated by Rossion et al. (2012), there is a large inter-individual variability in the localization of face-selective brain regions. In fact, single-subject analyses have evidenced numerous face-selective clusters in the inferior occipital lobe, fusiform gyrus, and superior temporal sulcus of individual brains (Figure 1.33).

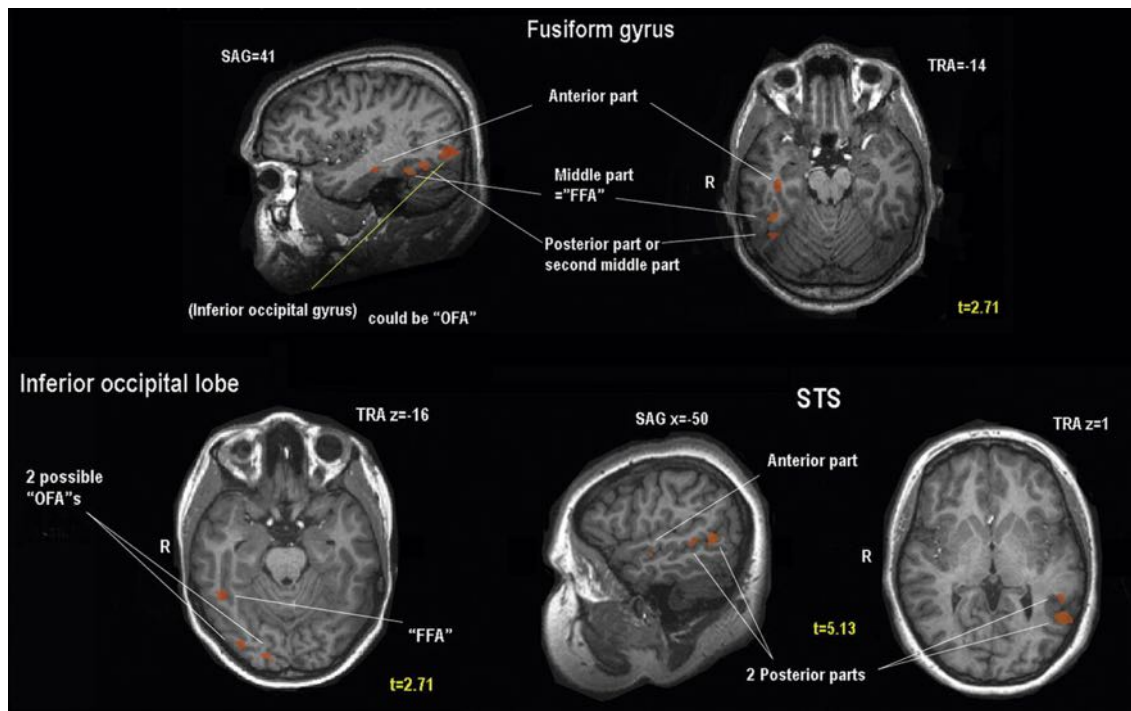


Figure 1.33. Numerous face-selective clusters in individual brains. *Top:* Four face-selective clusters were identified in the ventral stream of one participant. *Bottom left:* Two face-selective clusters were observed in the inferior occipital lobe of another participant. *Bottom right:* Three face-selective clusters were identified along the left dorsal face-selective pathway of an additional subject. Reprinted from Rossion et al. (2012).

Additional face-selective brain regions have also more recently been evidenced by electrophysiological studies in monkeys (Tsao et al., 2008a; 2008b) and fMRI studies using dynamic localizers (Fox, Iaria, & Barton, 2009a; Pitcher et al., 2011). As reported by Duchaine and Yovel in 2015 (see also, Bernstein & Yovel, 2015), a few of these supplementary areas include the Anterior Temporal Lobe (ATL), the anterior Superior Temporal Sulcus (aSTS), and the Inferior Frontal Gyrus (IFG) (Figure 1.34). Another face-selective area in the posterior fusiform gyrus (pFus), distinct from the face-selective region located in the middle fusiform gyrus, was recently identified by Weiner and Grill-Spector (2012) as well as a region in the anterior part of the fusiform gyrus, referred to as the antFus was reported by Rossion et al.

(2012) (Figure 1.33). Intracranial stimulation of the latter region in an epileptic patient implanted with cerebral electrodes led to distorted face perception as well as face recognition impairments (Jonas et al., 2015).



Figure 1.34. Face-selective brain regions. *Left:* The dorsal face-selective stream includes the posterior part of the Superior Temporal Sulcus Face Area (pSTS-FA), the anterior Superior Temporal Sulcus Face Area (aSTS-FA), and the Inferior Frontal Gyrus Face Area (IFG-FA). *Right:* The ventral stream involves the Occipital Face Area (OFA), the Fusiform Face Area (FFA), and the Anterior Temporal Lobe Face Area (ATL-FA). Reprinted from Duchaine and Yovel (2015).

Right-hemisphere dominance. Early lesion studies and positron emission tomography (PET) scans on face processing have reported a strong right-hemispheric dominance for face perception (e.g., Meadows, 1974; Sergent et al., 1992). These findings were supported by later fMRI studies in the field (e.g., Kanwisher et al., 1997; McCarthy et al., 1997; Puce et al., 1996) and by behavioral studies that showed human observers being faster when instructed to process faces presented in the left visual field (Hilliard, 1973). A significant body of literature on prosopagnosia has also evidenced a right-hemisphere superiority for face processing. Although lesion locations in prosopagnosic patients vary across individual brains (Figure 1.31), unilateral damages to the right-sided occipital and temporal lobes have been extensively observed in single-case studies (e.g., Barton et al., 2002; Bouvier & Engel, 2004; Busigny et al., 2010b; Rossion, 2014) with the first reports by Hecaen and Angerlergues dating back to 1962 (Hecaen & Angerlergues, 1962; in Rossion, 2014). In contrast, unilateral left-hemispheric posterior damages were reported in only five prosopagnosic patients. As mentioned by Rossion (2014), as four of them were left-handed, atypical left-hemispheric cerebral lateralization could be responsible for the face perception deficits observed (see also, Bukowski et al., 2013). In the healthy brain, more recent fMRI studies have evidenced face-selective regions in both hemispheres, although the right-hemispheric regions have been reported as being more critical for face processing and broader compared to the left-hemispheric regions (Bukowski et al., 2013; Rossion, 2014). Event-

Related Potential (ERP) studies also reported larger early face-sensitive ERPs over the right-hemisphere as compared to the amplitude observed in the left-hemisphere in response to faces (Bentin et al., 1996). In addition, Rangarajan et al. (2014) recently evidenced that electrical brain stimulations of the right fusiform gyrus, one of the most critical face-selective area, affected face perception, while brain stimulations of the left fusiform gyrus led only to visual changes that were unrelated to face perception (Figure 1.35). These findings together provide a strong support for a right-hemispheric lateralization of the human face processing network. Previous research in macaques have yielded more equivocal results, with some studies reporting right-hemisphere dominance for face perception in non-human primate brains (e.g., Hauser, 1993; Pinsk et al., 2009; Vermeire, Hamilton, & Erdmann, 1998), while others have not (Tsao et al., 2003; Tsao et al., 2008a). Interestingly, a very recent study implementing methods from optogenetics and pharmacology revealed that macaque monkeys were impaired in discriminating face-gender when neural suppression was applied to face-selective areas in both the right and the left-hemispheres (Afraz, Boyden, & DiCarlo, 2015).

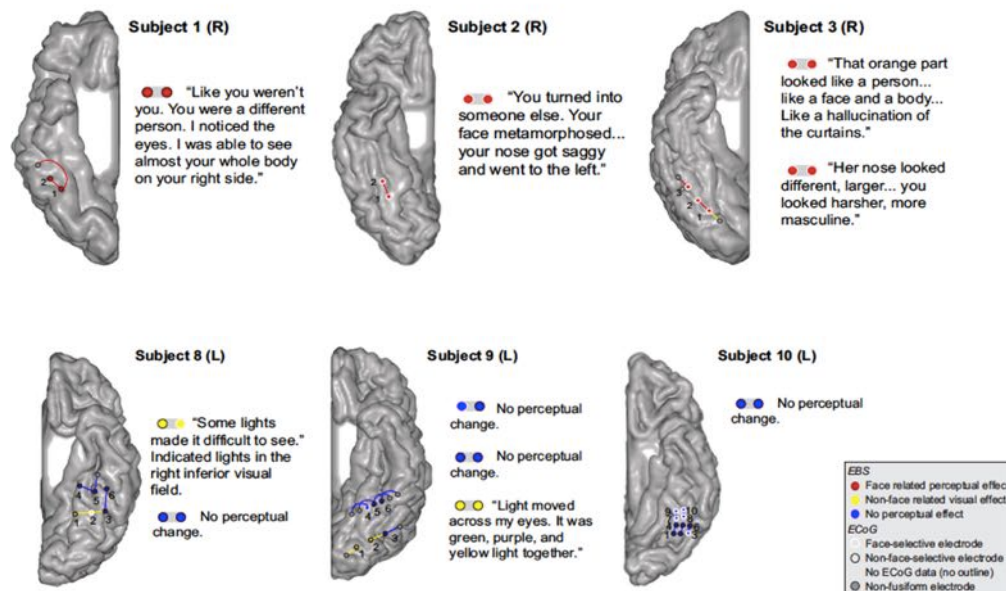


Figure 1.35. Electrical brain stimulations applied to face-selective regions in the right- and left-sided fusiform gyrus. Stimulation of the fusiform gyrus in the right-hemisphere (R) affected face perception (Subject 1, 2, 3), while stimulation of face-selective fusiform regions in the left hemisphere (L) only caused visual effects that were not related to faces (Subject 8, 9, 10). Adapted from Rangarajan et al. (2014).

1.3.2 Prominent models of face processing

The Bruce and Young model of face recognition. Based on neuropsychological and behavioral evidence, Bruce and Young (1986) have formulated a cognitive functional model of face processing taking into account identity, expression, and semantic information. As presented in Figure 1.36, during the first processing step, a view-centered description of the face is generated, which serves as a basis for separating the processing of more advanced features, such as expression or facial speech analysis. Face recognition involves the comparison of the generated face description with stored face recognition units (FRUs). A match between the description generated during the structural encoding and face recognition units results in the activation of identity-specific semantic nodes and ultimately, the retrieval of name codes. The directed visual processing pathway, for its part, directs attention towards a certain face.

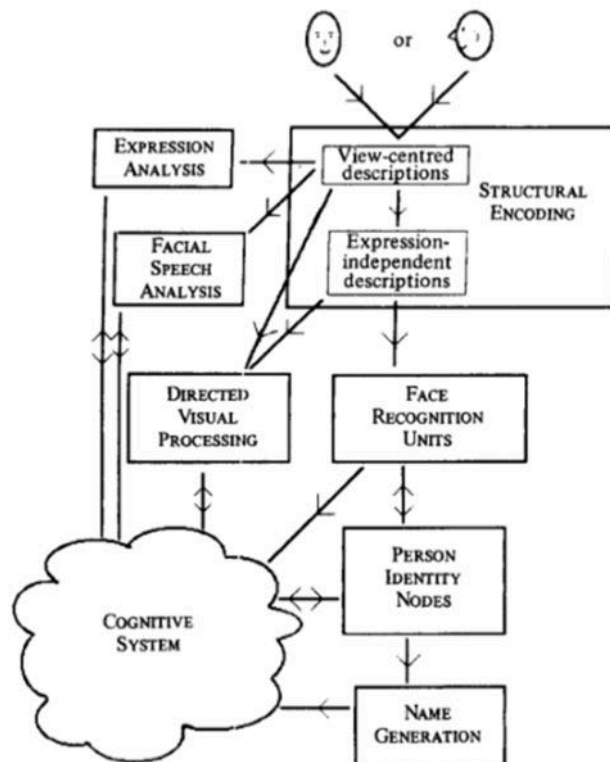


Figure 1.36. Bruce and Young's cognitive model of face processing (see text for description). Reprinted from Bruce and Young (1986).

As illustrated in Figure 1.36, this cognitive model involves independent parallel routes for the processing of identity and expression. Neuropsychological findings displaying spared facial expression recognition abilities in prosopagnosic patients despite severe deficits in

recognizing face identities (for a recent review, see Bate & Bennetts, 2015; Bruyer et al., 1983; Tranel, Damasio, & Damasio, 1988; Young et al., 1986), cognitive studies (e.g., Young et al., 1986), as well as studies in non-human primates (e.g., George et al., 1993) were taken as evidence for this independence. However, in the light of more recent behavioral (e.g., Campbell & Burke, 2009; Ganel, Goshen-Gottstein, & Ganel, 2004), computational (Calder et al., 2001a), and neuroimaging evidence (e.g., Fox & Barton, 2007; Fox et al., 2009b), some authors have cast doubts on the segregation between identity and expression processing, suggesting that the division might rather be partial than absolute (Calder, 2011; Calder & Young, 2005; Lander & Butcher, 2015, see also **CHAPTER 2** and **CHAPTER 4** of this thesis).

Because this cognitive model did not propose a comprehensive view of the neural substrates involved in face perception, other models based on fMRI and neuropsychological evidence were subsequently developed. One of the most prominent neuroanatomical model was proposed by Haxby and colleagues in 2000, guiding the research on face processing over the last decades.

The distributed neural model proposed by Haxby et al. (2000). The neuroanatomical model proposed by Haxby and colleagues in 2000 was motivated by findings from electrophysiological research with macaques, neuropsychological reports from brain-damaged patients, and human functional neuroimaging studies. According to this model, three core regions, divided into two functionally distinct cortical pathways, underlie the processing of different facial aspects (Figure 1.37). The ventral pathway, subtending the processing of invariant facial aspects, such as individual face identity, involves the lateral fusiform gyrus (Fusiform Face Area – FFA), whereas the dorsal pathway, preferentially responding to changeable aspects of faces such as emotional expression or eye-gaze direction, includes the superior temporal sulcus (STS) (for a review see, Calder & Young, 2005; Pessoa, 2008). Critically, the inferior occipital gyrus (Occipital Face Area – OFA) is considered to be the entrance to the face processing system for both cortical pathways, despite their functional disparities (Figure 1.38).

Haxby et al. (2000) further suggest that an extended face network, comprising areas connected to the core regions, is specialized in the processing of distinct facial features. Three areas are connected to the STS – the intraparietal sulcus, which directs attentional resources towards specific spatial locations according to eye-gaze direction; the amygdala and limbic system, involved in the processing of the emotional aspects of faces; and the auditory cortex specialized in the processing of prelexical speech (Figure 1.38). An additional area in the anterior temporal cortex, connected to the FFA underlies the processing of semantic information about a person, such as identity or name (Haxby et al., 2000).

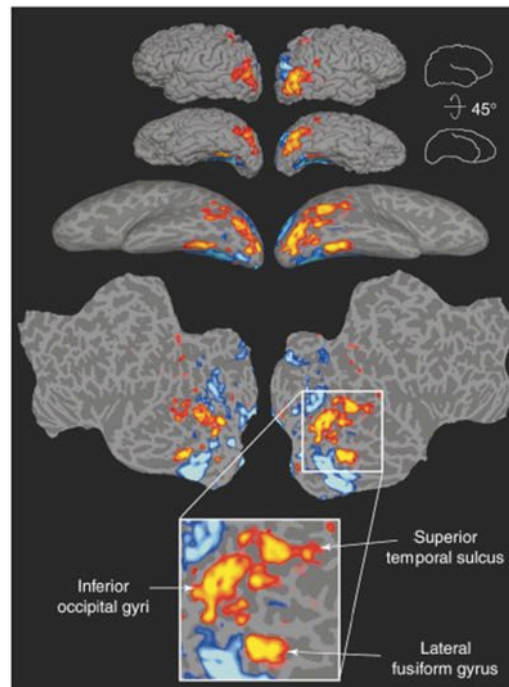


Figure 1.37. The three core regions of the face processing system. Single-subject fMRI data displaying the inferior occipital gyri, the superior temporal sulcus, and the lateral fusiform gyrus. Reprinted from Haxby et al. (2000).

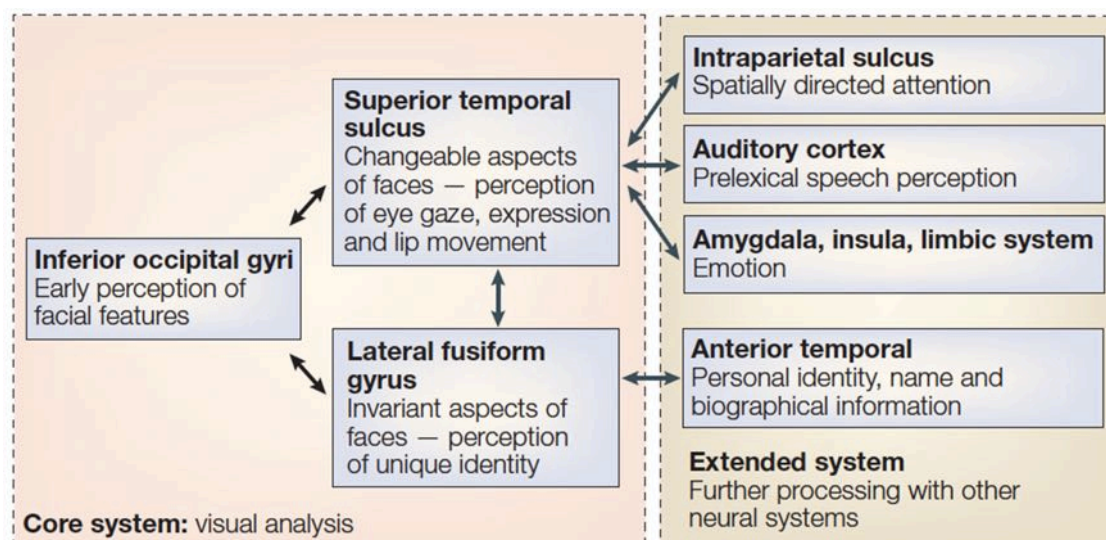


Figure 1.38. The influential neural model of human face processing proposed by Haxby and colleagues in 2000. The inferior occipital gyri, the superior temporal sulcus, and the lateral fusiform gyrus are the three face-selective regions involved in the core system. These areas are connected to an extended system. Reprinted from Calder and Young (2005).

1.3.3 A critical review of early influential models of face perception

The dissociation between invariant vs. changeable aspects of faces is less clear than initially presumed. According to Haxby et al. (2000), two functionally distinct cortical pathways, a ventral and a dorsal stream, are responsible for the processing of invariant and changeable aspects of faces. Bernstein and colleagues (2017) very recently challenged this prevalent view by conducting an fMRI study in which the neural responses to invariant (gender) and changeable (expression) facial features were measured for both static and dynamic stimuli. The aim of this study was to assess the nature of the segregation between dorsal and ventral pathways in order to establish whether the division between both streams impacts the processing of changeable vs. invariant facial aspects or dynamic vs. static face information. The results revealed that the pSTS in the dorsal stream exhibited a stronger sensitivity to changeable aspects of faces (expression > gender) and face motion (dynamic > static), while the OFA and FFA in the ventral stream showed similar responses to static and dynamic information and changeable and invariant facial features (Figure 1.39). Critically, these findings do not support the clear-cut neuroanatomical functional division between invariant and changeable aspects of faces suggested by Haxby et al. (2000), as the ventral stream processes information from faces regardless of the type of feature extracted (invariant vs. changeable) and their temporal properties (static vs. dynamic).

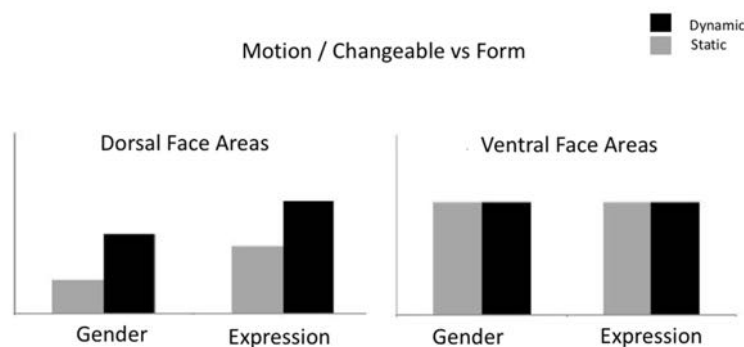
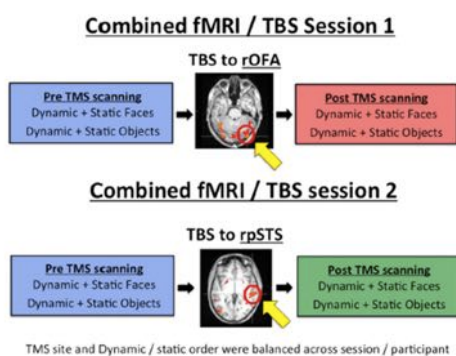


Figure 1.39. The processing of invariant vs. changeable aspects of faces in dorsal and ventral streams. Dorsal face areas are selectively more sensitive to motion and changeable aspects of faces, whereas ventral face areas extract similar information from dynamic and static faces for both invariant and changeable facial aspects. Adapted from Bernstein et al. (2017).

Several routes are connecting early visual areas to face-selective regions. Inconsistent with the neural model proposed by Haxby et al. (2000), recent findings have provided evidence supporting the idea that several routes connect the primary visual cortex to the core regions involved in the face processing network. For instance, by combining thetaburst transcranial

magnetic stimulation (TBS) with fMRI, Pitcher and colleagues (2014) recently observed a causal engagement of the right inferior occipital gyrus (rOFA) in the processing of static face images, whereas the processing of dynamic faces (Pitcher et al., 2014) was supported by the posterior superior temporal sulcus (pSTS). Delivering TBS over the rOFA affected the BOLD responses to static but not dynamic faces in the pSTS, while TBS disruptions of the pSTS affected the neural responses to dynamic but not static faces in the pSTS (Figure 1.40A, Figure 1.41). These findings suggest that static and dynamic properties of human faces are processed via distinct cortical pathways, originating in the early visual cortex and not in the rOFA, as predicted by Haxby and colleagues (2000) (Figure 1.40B). An fMRI study using a diffusion tensor probabilistic tractography to identify white matter connections between different face-selective brain regions has recently provided further evidence for multiple routes connecting early occipital areas to the core regions of the face processing system (Gschwind et al., 2012). The authors revealed that the OFA and FFA were tightly connected, with white-matter connections being stronger in the right-hemisphere. Conversely, no connections were found between the OFA and pSTS or FFA and pSTS (but see Pitcher et al., 2014), implying that the neural computations occurring in the pSTS are distinct and functionally independent from those occurring in the FFA and OFA. Similar findings were reported in two subsequent studies (Bernstein et al., 2017; Pyles et al., 2013). Finally, neuropsychological evidence revealing brain activations in the pSTS or FFA of prosopagnosic individuals with lesioned OFA underpins the notion that cortical projections bypassing the inferior occipital gyrus directly connect early visual areas to the fusiform face area or the posterior section of the superior temporal sulcus (Dalrymple et al., 2011; Sorger et al., 2007; Yang, Susilo, & Duchaine, 2015). These latter findings have led Rossion (2008; 2014) to propose a reverse OFA/FFA hierarchical framework of the face processing system.

A



B

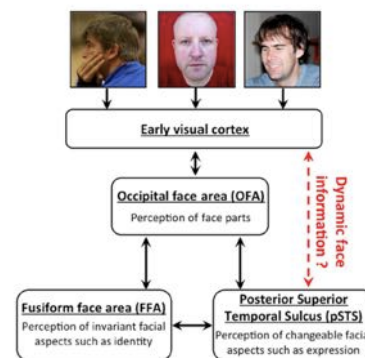


Figure 1.40. Distinct cortical pathways for static and dynamic face information. (A) Pitcher et al. (2014) delivered thetasturbst transcranial magnetic stimulations (TBS) over the rOFA and rpSTS while participants viewed dynamic and static faces and objects. (B) By adopting this virtual lesion approach, the authors probed the existence of a direct pathway connecting the early visual areas to the pSTS. Adapted from Pitcher et al. (2014).

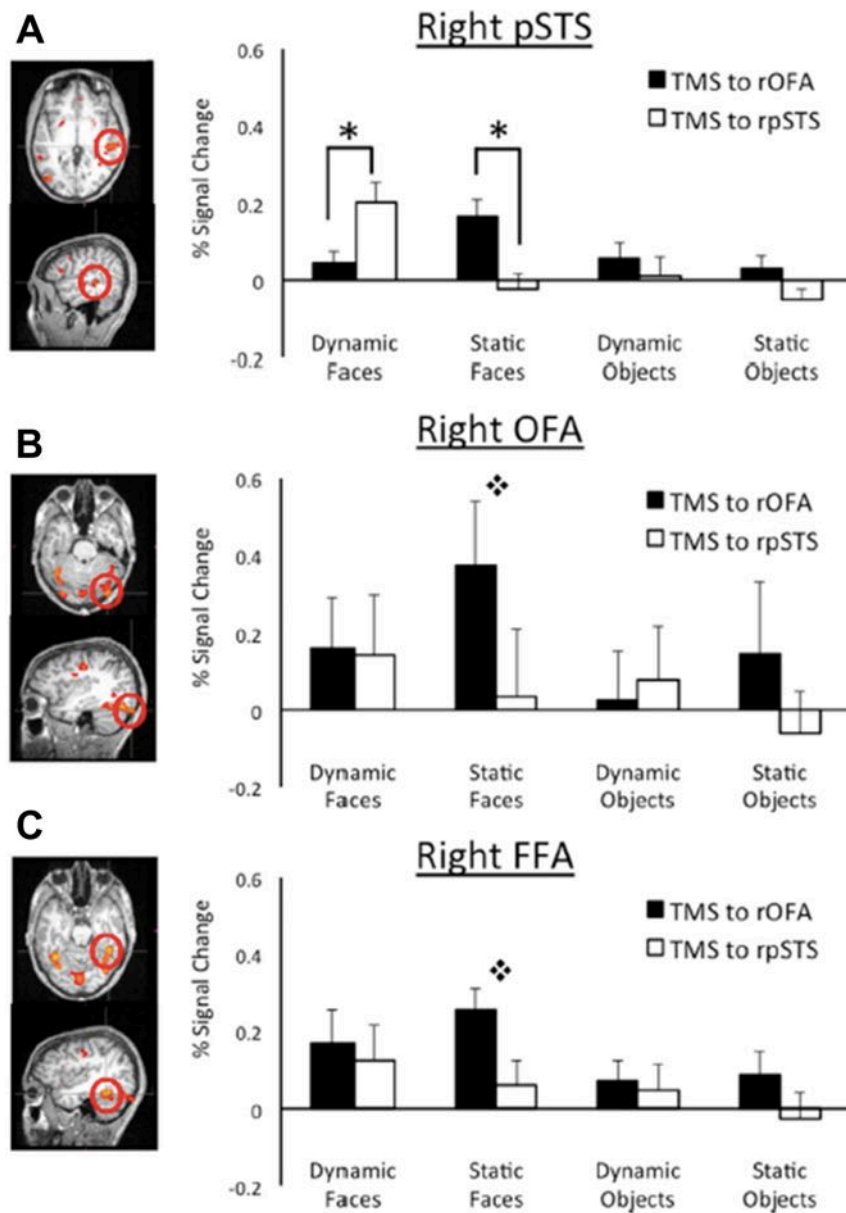


Figure 1.41. The size of the TBS disruption effect in the right pSTS, rOFA and rFFA for static and dynamic faces and objects. A positive value indicates a reduction in the signal caused by the TBS stimulation. (A) TBS applied over the right rOFA reduced the response to static but not dynamic faces in the right pSTS, while TBS applied over the right pSTS reduced the response to dynamic but not static faces in this region (asterisks indicate significant differences). (B) TBS applied over the rOFA reduced the response to static, but not dynamic faces in the rOFA. (C) Delivering TBS over the rOFA reduced the response to static, but not dynamic faces in the rFFA. (diamonds indicate a significant difference after Bonferroni adjustment). Delivering TBS over the right pSTS, the rOFA and rFFA did not affect neural responses to dynamic and static objects. Reprinted from Pitcher et al. (2014).

Rossion's reverse hierarchical view of face perception. According to the hierarchical view proposed by Haxby and colleagues (2000), the OFA is the first processing locus in the face cortical network, transmitting visual inputs to the FFA and pSTS, two areas that perform more advanced processes associated with invariant and changeable aspects of faces. Observations from brain-damaged patients, however, have called into question this serial staged hierarchy, by showing activations of the FFA or pSTS, despite lesions to the inferior occipital gyrus (OFA) (Figure 1.42). For instance, the acquired prosopagnosic patient PS with extensive brain lesions to the right inferior occipital cortex, encompassing the rOFA, exhibits face-selective activations in the rFFA and right pSTS comparable to those observed in controls (Rossion et al., 2003; Sorger et al., 2007). Further observations from two other brain-damaged patients revealed that lesions to the rOFA and rFFA did not prevent face-selectivity in the ipsilateral pSTS (Dalrymple et al., 2011; see also Yang et al., 2015). Moreover, the findings from patient DF suffering from visual form agnosia revealed that intact activations of the FFA and pSTS were not due to the transmission of visual inputs from the OFA contralateral to the lesion (i.e., the intact lOFA forwarding visual information to the rFFA). In fact, in patient DF, neuroimaging findings provided evidence that lesions in the lateral occipital regions enclosing both the rOFA and lOFA (Goodale & Milner, 1992) did not prevent bilateral activations of the FFA and pSTS (Milner & Goodale, 2008; Steeves et al., 2006). All these findings suggest that cortical routes independent from the OFA are connecting early visual areas to the FFA, findings that are in line with the results from diffusion tensor imaging studies (DTI, Kim et al., 2006).

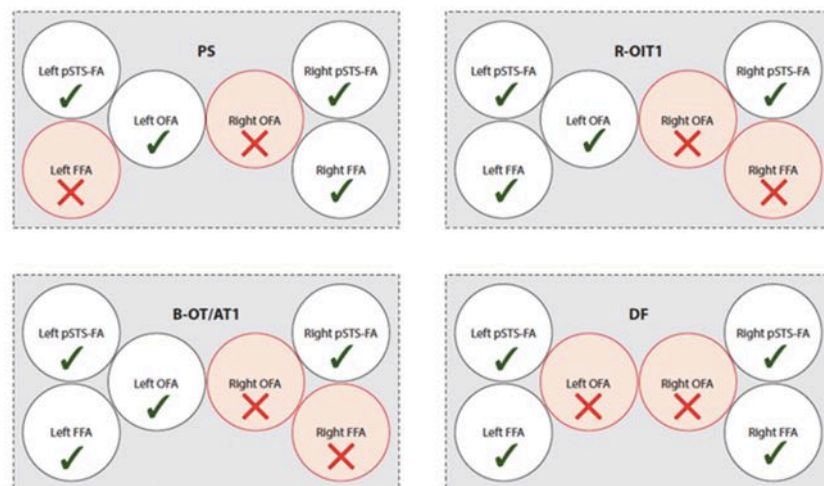


Figure 1.42. A schematic representation of lesioned (✗) and preserved (✓) face-selective brain regions in four patients. Despite lesioned rOFA, PS (Sorger et al., 2007), R-OIT1, and B-OT/AT1 (Dalrymple et al., 2011) exhibit normal activations of the rFFA (PS) and right pSTS (PS, R-OIT1, B-OT/AT1). In DF, damage to the rOFA and lOFA does not prevent bilateral activation of the FFA and pSTS (Steeves et al., 2006). Adapted from Duchaine and Yovel (2015).

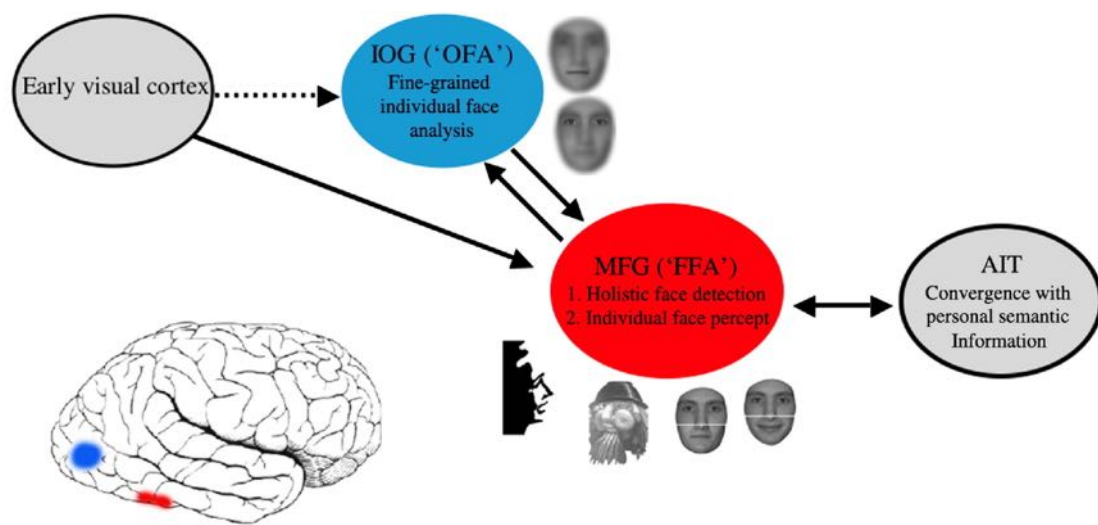


Figure 1.43. The reverse hierarchical view of face perception proposed by Rossion (2008; 2014). After an initial processing in the early visual cortex, visual information is transmitted to the middle fusiform gyrus (MFG) where face-selectivity emerges in the FFA if the transmitted information corresponds to the representation of a face. The FFA processes faces as wholes, before transmitting the visual information to the more posteriorly located inferior occipital gyrus, where the OFA computes finer-grained analyses of the individual facial features. The individual faces are finally holistically integrated in the FFA. Reprinted from Rossion (2008).

Given these observations, Rossion (2008) reformulated the neuro-functional organization of the face processing network, proposing a reverse OFA/FFA hierarchical view of human face perception. According to this model, visual information is initially processed in the early visual cortex before being transmitted to the middle fusiform gyrus (MFG) via two pathways – the first one bypassing the OFA; the second one traversing it (Figure 1.43). In the MFG, a face is represented as a whole (holistic face detection), activating the face selectivity of this region. In contrast to the neural model proposed by Haxby et al. (2000), visual information can reach the FFA independently of the OFA, a prediction that can explain intact activation of the FFA despite lesioned ipsilateral (Dalrymple et al., 2011; Rossion et al., 2003; Sorger et al., 2007) or bilateral OFA (Steeves et al., 2006). The global face representation generated in the MFG is then sharpened through re-entrant connections between the MFG (FFA) and IOG (OFA), triggering additionally the face preferential neural activation in this lower-level visual region (Rossion, 2008). As OFA face-cells have smaller receptive fields than FFA neurons, they are tuned to perform fine-grained visual analysis of distinct facial features, such as the eyes, nose, and mouth (Rossion, 2008). The refined representations are then sent back to the FFA, where a given face is finally holistically represented (individual face percept).

This reverse hierarchical neurofunctional model of face processing is based on fMRI evidence revealing that the rOFA is not activated by poorly-defined faces, such as two-tone Mooney faces or Arcimboldo's paintings (Figure 1.44). Because their features are ambiguous and difficult to identify individually, Mooney faces or Arcimboldo face-like paintings are perceived as faces only when integrated holistically. Interestingly however, these stimuli activate the rFFA, even when the face-selectivity in the rOFA is missing, supporting the view that the FFA subserves the categorization of faces as perceptual wholes, while the OFA underlies the processing of their individual features (Figure 1.44) (Rossion et al., 2011). Consistent with these observations, an event-related fMRI study further revealed that coarsely defined faces (low spatial frequency (LSF) faces) selectively activated the rFFA, but not the rOFA (Goffaux et al., 2011). In addition, in an fMRI study investigating the timing of activity in the rFFA and rOFA, Jiang and colleagues (2011) observed the earliest face-selective responses in the rFFA and not in the rOFA. As shown in Figure 1.45, differential activation to faces and cars emerged earlier in the rFFA as compared to the rOFA, where face-selectivity appeared significantly later. Finally, another striking evidence from the neuropsychological literature supporting this reverse hierarchical view comes with the patient, NS. This patient suffers from prosopagnosia following lesions to face-selective fusiform regions encompassing the rFFA. Despite being preserved, the rOFA of this patient does not exhibit face-selective neural activation in response to faces (Delvenne et al., 2004; Rossion, 2008).

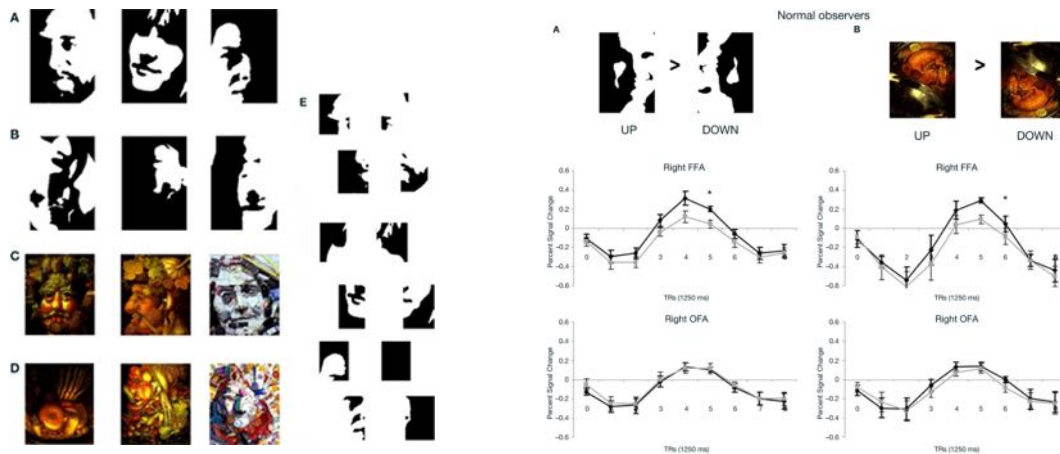


Figure 1.44. Face-selective neural activation in the rOFA and rFFA to upright and inverted Mooney faces and Arcimboldo face-like paintings. Left: (A) Upright and (B) inverted Mooney faces. (C) Upright and (D) inverted Arcimboldo's paintings. (E) Individual parts of the Mooney faces are particularly ambiguous and difficult to identify. Right: (A) Differential neural activation was observed in the rFFA to upright compared to inverted Mooney faces as indicated by the asterisk. No differential activation for upright and inverted Mooney faces was observed in the rOFA. The same pattern of result was observed for Arcimboldo's paintings. Reprinted from Rossion et al. (2011).

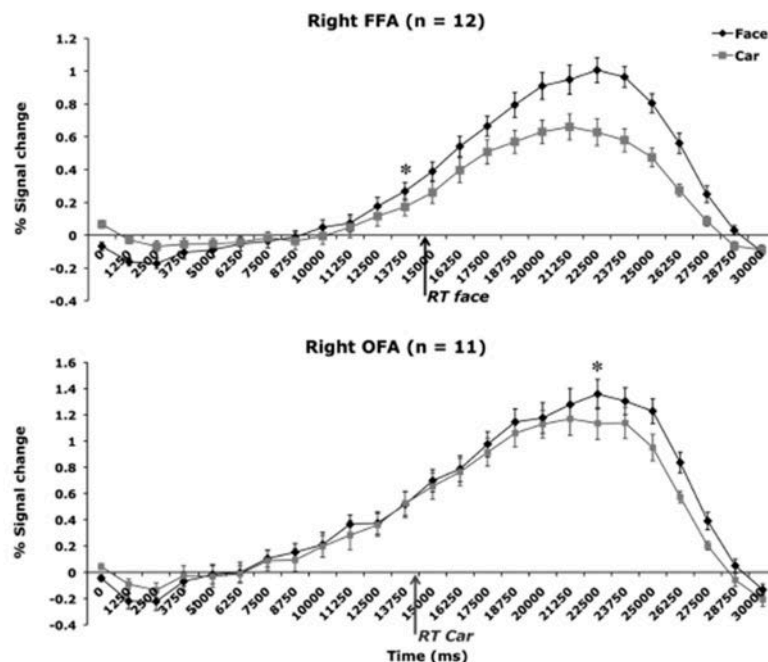


Figure 1.45. Differential neural activation to faces and cars in the rFFA and rOFA. Face-selectivity emerged earlier in the rFFA compared to the more posteriorly located rOFA. The asterisks indicate the onset time at which the face-selective brain regions significantly differ in their sensitivity to faces and cars. Reprinted from Jiang et al. (2011).

Duchaine and Yovel's revised neural framework for face perception. In a similar way that is also based on new findings regarding face-selective areas and the discovery of additional face-selective regions, Duchaine and Yovel (2015) recently proposed a revised neural framework of the face processing system. This revised framework suggests that the face processing network depends on two distinct pathways, a ventral stream, responding preferentially to form information, and a dorsal stream, mainly involved in the processing of dynamic and form information.

This framework is primarily based on the aforementioned evidence suggesting that several routes project from early visual cortex into the face processing system. Based on neuropsychological evidence (e.g., Delvenne et al., 2004; Rossion, 2008; 2014), diffusion tensor imaging findings (Gschwind et al., 2012; Pyles et al., 2013), and functional connectivity studies (Avidan et al., 2014), the authors suggest that the rOFA is not the only gateway to the face processing network.

Second, this framework is also inspired by findings from TMS and combined EEG–fMRI studies, providing insights into the timing of neural activity in face-selective regions (e.g., Pitcher, 2014; Pitcher et al., 2008; Pitcher et al., 2012; Pitcher et al., 2007; Sadeh et al., 2010). In a recent study, double-pulse TMS (dTMS) was delivered over the rOFA and over the pSTS during a facial expression recognition task, with latencies varying up to 210ms after the stimulus presentation. Delivering dTMS over the pSTS impaired expression recognition over a longer period (60 to 140ms) than disrupting the rOFA (60-100ms) (Pitcher, 2014). Importantly, dTMS delivered at the latencies of 60 to 100ms disrupted recognition performance in both face-selective regions (Figure 1.46). These findings suggest that the processing of face information in the rOFA and pSTS is not as hierarchized as initially presumed (Haxby et al., 2000), but also occurs simultaneously. These findings also support the perspective that both face-selective regions receive visual inputs from functionally distinct cortical pathways directly connecting them to the early visual cortex.

An additional modification to the Haxby model proposed by Duchaine and Yovel (2015) concerns the role played by the FFA in the face processing system. According to Haxby and colleagues (2000), the FFA preferentially responds to invariant aspects of faces, such as individual identity or gender, while the pSTS is involved in the processing of changeable aspects of faces, such as emotional expressions or lip movements. However, recent findings have suggested that the FFA may also be involved in the processing of changeable aspects of faces. As mentioned above, Bernstein et al. (2017), for instance, have demonstrated similar sensitivity of the OFA and FFA to both changeable and invariant aspects of faces, while the

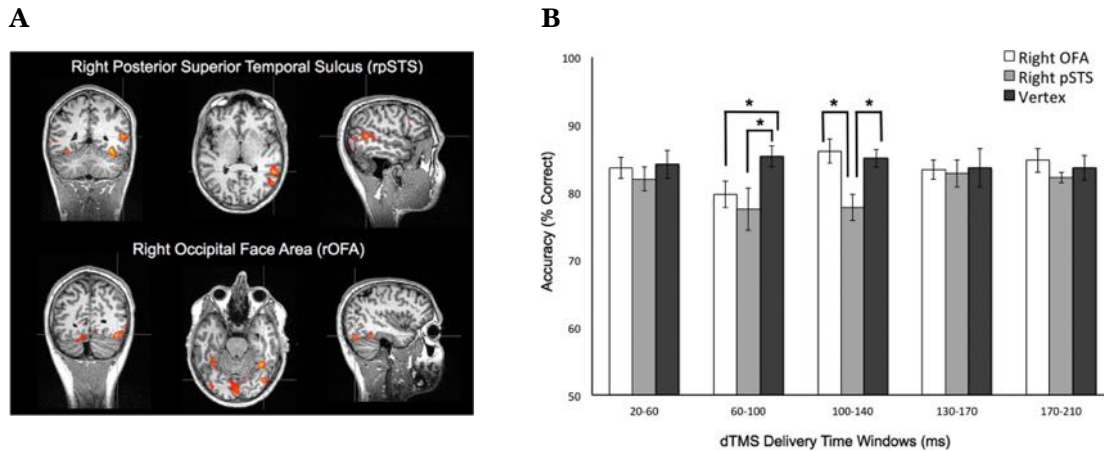


Figure 1.46. Facial expression recognition occurs simultaneously in the rOFA and pSTS as evidenced by a TMS study (Pitcher, 2014). (A) Stimulation sites (rOFA and pSTS) in one participant. (B) Disruption of the rOFA impaired recognition performance at the latencies of 60–100ms, while stimulation of the pSTS impaired recognition performance at both the latencies of 60–100ms and 100–140ms as compared to a control site (vertex). Adapted from Pitcher (2014).

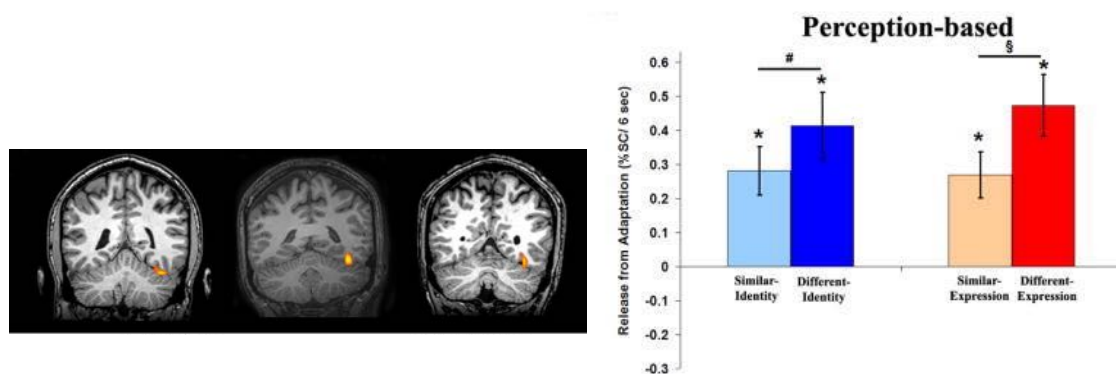


Figure 1.47. The FFA is involved in the processing of emotional faces: Evidence from an fMRI adaption study. *Left:* Coronal slices of the right fusiform face area (rFFA) of three participants. *Right:* Perception-based analysis revealed release from adaptation in the rFFA when observers noticed changes in the expression or identity of the faces presented. Adapted from Fox et al. (2009a).

pSTS was only sensitive to changeable aspects of faces. Additionally, an earlier fMRI-based study by Ganel et al. (2005) revealed an increased activation in the FFA when observers were asked to judge facial expressions. Moreover, these findings were observed even when participants were asked to explicitly direct their attention to the identity of the presented faces. Later, fMRI adaptation studies (fMRIa) also challenged the idea of a distinct separation between the processing of changeable and invariant aspects of faces by revealing that changes in facial expressions led to a release from adaptation in the FFA (Figure 1.47). These findings imply a sensitivity of this face-selective area for the processing of emotional faces and do not support anatomic division of emotion and identity processing (Cohen-Kadosh et al., 2010; Fox et al., 2009b; Xu & Biederman, 2010).

Finally, Duchaine and Yovel's revised framework also considers recent findings showing larger activation to dynamic faces as compared to static ones in the pSTS, the aSTS and IFG (Bernstein et al., 2017; Fox et al., 2009a; Pitcher et al., 2011)¹.

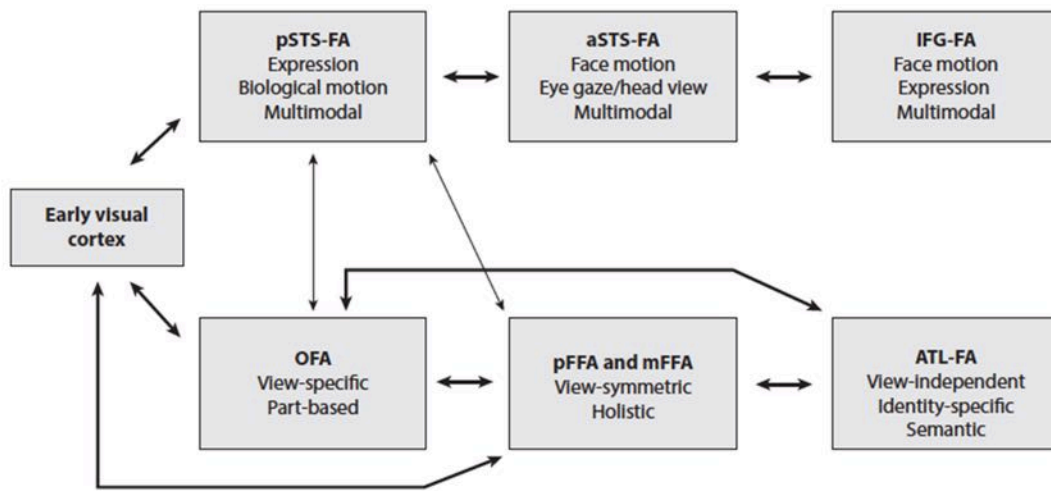


Figure 1.48. The revised neural framework of the face processing system proposed by Duchaine and Yovel (2015). The model involves the core areas suggested by Haxby et al. (2000), the OFA, pSTS, and FFA, connected with the aSTS, IFG, and ATL. This framework suggests that face processing relies on two separate streams, a ventral stream involving the OFA, FFA, and ATL and responding preferentially to form information; and a dorsal stream, involving the pSTS, aSTS, and IFG specifically tuned to the processing of dynamic information. Reprinted from Duchaine and Yovel (2015).

¹Because this thesis investigates the processing of static and dynamic faces, these recent findings revealing increased responses to dynamic faces as compared to static ones in dorsal face-selective brain areas will be approached in more details in the next section.

As shown in Figure 1.48, the revised framework proposed by Duchaine and Yovel (2015) comprises the three core regions suggested by Haxby et al. (2000), the OFA, the FFA, and the pSTS, connected with the recently discovered additional face-selective areas, the aSTS, the IFG, and the ATL. Based on evidence from TMS (Pitcher et al., 2014), fMRI (Fox et al., 2009a; Pitcher et al., 2011), and connectivity studies (Gschwind et al., 2012; Pyles et al., 2013), this framework is divided into two distinct, but functionally connected, cortical pathways.

As described by Duchaine and Yovel (2015), the ventral pathway, which involves the OFA, FFA, and ATL, preferentially responds to invariant form information, such as age, gender, or identity and also performs the processing of emotional expressions. Located in the most posterior part of the ventral area, the OFA face-selective cells, characterized by their very small receptive fields (Hemond, Kanwisher, & De Beeck, 2007; Rossion, 2008, 2015), process individual face parts in a view-specific manner. Importantly, although the OFA is linked to the pSTS, little is known regarding the nature of this functional connection (Duchaine & Yovel, 2015; see also, Pitcher et al., 2014; Pitcher et al., 2017). Positioned at a higher stage of the hierarchy, the FFA receives visual inputs from both the OFA and early visual areas (Rossion, 2008; 2014; 2015) and accomplishes the processing of face symmetry (e.g., Axelrod & Yovel, 2012; Caldara & Seghier, 2009; Caldara et al., 2006). As the FFA is involved in the processing of symmetry and form information, it also contributes to the processing of emotional faces (Cohen-Kadosh et al., 2010; Ganel et al., 2005; Xu & Biederman, 2010). Finally, the anterior temporal lobe face area (ATL), the face-selective area located in the most anterior part of the ventral pathway, receives visual inputs from both the OFA and FFA (Gschwind et al., 2012; Pyles et al., 2013). This area is considered to play a central role in the processing of invariant aspects of faces, such as individual identity (Anzellotti, Fairhall, & Caramazza, 2013; Yang et al., 2015) and the processing of semantic information. As mentioned by Duchaine and Yovel (2015), this region most probably corresponds to the neuroanatomical locus of the stored face recognition units proposed by Bruce and Young (1986) in their cognitive model of face processing.

The dorsal pathway, which involves the pSTS, the aSTS, and the IFG, contributes to the processing of form and dynamic information. As revealed in recent studies, this dorsal pathway elicits stronger responses to dynamic faces when compared to static ones (e.g., Bernstein et al., 2017; Fox et al., 2009a; Pitcher et al., 2011, see also section 1.3.4), which is consistent with the notion that it contributes to the processing of changeable aspects of faces, such as emotional expressions, lip movements, or eye-gaze direction. In an earlier study, O'Toole et al. (2002) also suggested that this dorsally-based stream is involved in the processing of facial identity and the recognition of familiar faces when dynamic cues are provided (see also, Lander & Butcher, 2015; O'Toole & Roark, 2010). In addition, Arsalidou et al. (2011) reported increased

sensitivity of the superior temporal sulcus to dynamic as compared to static emotional faces (see also, LaBar et al., 2003), an observation that is consistent with findings reported by a later study revealing an involvement of the right pSTS in the integration of social signals conveyed by both faces and voices (Watson et al., 2014; see also, Hasan et al., 2016; Yovel & O'Toole, 2016). Altogether, these results suggest that the dorsal pathway plays a dedicated role in the processing and interpretation of socially and emotionally relevant stimuli.

1.3.4 The processing of static and dynamic faces

Over the course of social interactions, our visual system is steadily stimulated by dynamic cues, yet most studies investigating face processing have relied on static images and only little evidence has been accumulated on the processing of dynamic faces. However, more recently, fMRI studies have addressed the question as to whether differential neuroanatomical bases underlie the processing of static and dynamic faces (Fox et al., 2009a; Pitcher et al., 2011; Pitcher et al., 2014). For instance, Pitcher et al. (2011) compared neural activation to static and dynamic stimuli in different face-selective regions including the rFFA, rOFA, and pSTS. They used a dynamic localizer and two independent sets of data, the first one to localize the face-selective brain regions and the second one to systematically examine whether static and dynamic faces, bodies, scenes, and objects differentially affect neural responses in these face-selective regions. While the neural response to static and dynamic stimuli did not differ in the rOFA and rFFA, the activation elicited by the pSTS to dynamic faces was nearly three times stronger than that to static faces. Moreover, as shown in Figure 1.49, the aSTS, a face-selective region located in the anterior superior temporal sulcus responded only to dynamic faces. It is worth noting that the activation to dynamic faces in the pSTS and aSTS was larger than the activation observed to moving bodies or objects, implying that these two regions are not only tuned to process motion, but particularly driven by dynamic face information (Pitcher et al., 2011).

In another study, Fox, Iaria, and Barton (2009a) investigated whether the presentation of dynamic stimuli was more effective in localizing face-selective brain areas. To this aim, the authors compared two functional localizers, the first one involving dynamic movies of faces and objects, the second one their static versions. Regions within the core face processing system (OFA, FFA, STS) were better identified with the functional localizer involving dynamic stimuli – the localization efficiency reaching up to 98%, while only 72% of the regions of interest were identified with the static localizer. Furthermore, their findings also revealed that the temporal properties of the stimuli influenced the cluster size at which the highest face-selectivity occurred. The dynamic localizer facilitated the localization of larger face-selective clusters, twice

as large in the rFFA and rOFA, nearly nine times larger in the right pSTS (Figure 1.50). Finally, neural responses to videos of moving faces or static photographs of faces were not different in the rOFA and rFFA – decreasing even when dynamic objects were presented (Figure 1.51). In contrast, the pSTS elicited neural responses that were nearly two times larger to dynamic as compared to static faces, findings that are consistent with the results reported by Pitcher et al. (2011). This study offers a novel insight into the use of dynamic stimuli in functional localizers, implying that they provide a sensitive, selective, and consistent approach to identify face-selective brain regions.

Converging evidence also suggests that dynamic emotional faces elicit larger activations in brain areas crucial for the processing of social and emotional information (e.g., Arsalidou et al., 2011; Liang et al., 2017). In contrast to static emotional faces, dynamic faces have been associated with greater responses in the fusiform gyrus (e.g., Kessler et al., 2011; Sato et al., 2004; Trautmann et al., 2013; Trautmann, Fehr, & Herrmann, 2009), the superior temporal sulcus (e.g., Kessler et al., 2011; Kilts et al., 2003; Trautmann et al., 2009), the inferior frontal gyrus (e.g., Trautmann et al., 2009), as well as in the visual motion area, V5, in the occipito-temporal lobe (Johnston et al., 2013; Kilts et al., 2003; Sato et al., 2004) (Table 1.5). Increased activity to dynamic information in the superior temporal sulcus and in the visual motion area supports the results of previous studies presenting an involvement of these regions in the processing of biological motion (Peelen, Wiggett, & Downing, 2006; Puce & Perrett, 2003; Saygin, 2007).

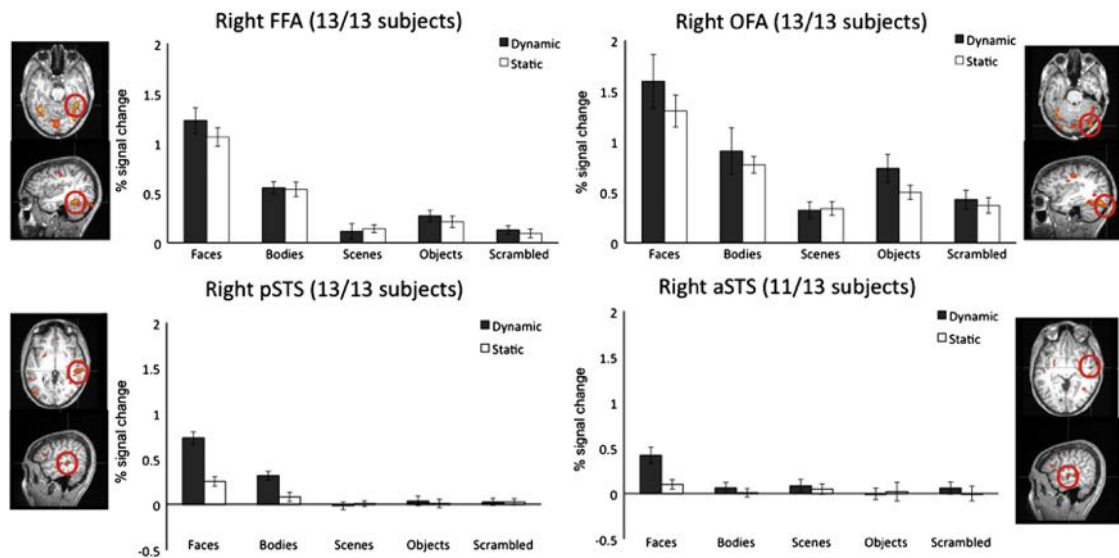


Figure 1.49. Percentage of signal change to dynamic and static faces, bodies, scenes, objects and scrambled faces in the rFFA, rOFA, right pSTS and right aSTS. While the rFFA and rOFA exhibited only a slight increase to dynamic as compared to static faces, the response in the pSTS to dynamic faces was nearly three times as strong as that to static faces. In a similar way, the aSTS exhibited significantly stronger response to dynamic as compared to static faces. Reprinted from Pitcher et al. (2011).

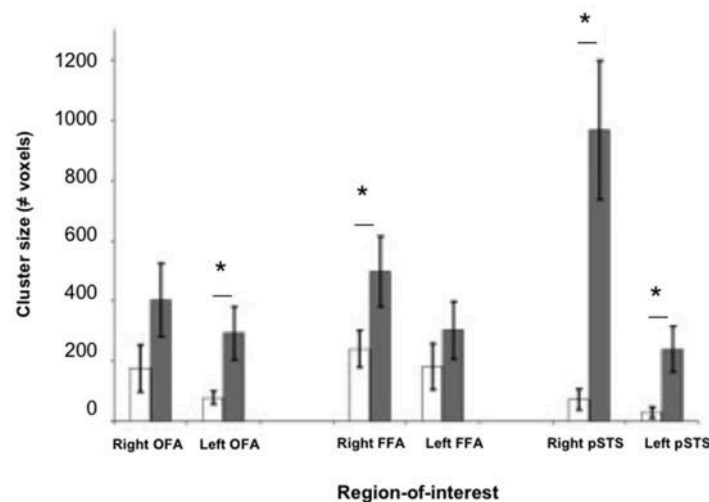


Figure 1.50. Cluster size with static and dynamic localizers. The presentation of dynamic images (gray bars) resulted in the activation of larger clusters in face-selective brain regions as compared to the presentation of static images (white bars). Adapted from Fox et al. (2009a).

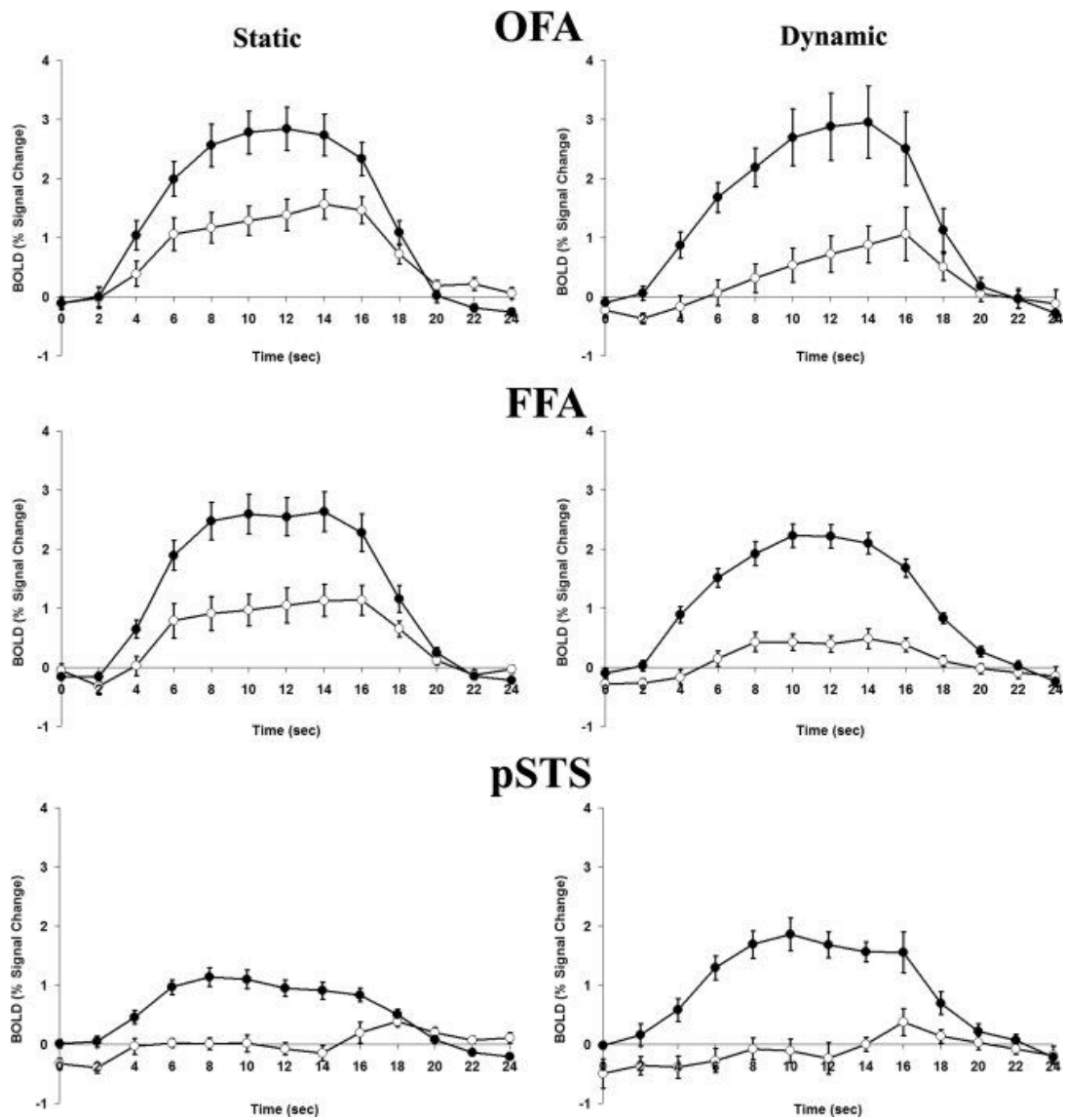


Figure 1.51. Static and dynamic functional localizers. Time course of signal intensity changes within the rOFA, rFFA, and right pSTS to faces (black circles) and objects (white circles) presented in their static (left column) and dynamic (right column) versions. Reprinted from Fox et al. (2009a).

Table 1.5. Neuroimaging studies comparing the processing of static and dynamic facial expressions. Adapted from Alves (2013).

Study	Participants	Brain recording technique	Main findings
Arsalidou et al. (2011)	24 healthy participants	Functional neuroimaging (fMRI)	Increased neural activity to dynamic expressions was observed in brain regions involved in the processing of emotional and social signals (STS, middle temporal gyri, and amygdala).
Johnston et al. (2013)	28 healthy participants	fMRI	Expression recognition was associated with a widespread network of regions in the occipito-temporal, parietal, and frontal cortex for both static and dynamic stimuli. Moreover, portions of the inferior frontal gyrus were more activated during the processing of static as compared to dynamic expressions suggesting a stronger involvement of this brain region in the processing of static face information.
Kessler et al. (2011)	30 healthy participants	fMRI	Regardless of the emotion presented, dynamic faces selectively activated the superior temporal sulcus, visual area V5, the fusiform gyrus, the thalamus, and other brain regions in frontal and parietal areas. Static expressions of happiness selectively elicited greater activity in the medial prefrontal cortex.
Kilts et al. (2003)	13 healthy participants	Positron emission tomography (PET)	Differential neural activation to dynamic as compared to static faces was observed in the visual area V5, the STS, the periamygdaloid cortex, and the cerebellum for the expression of anger. For the expression of happiness, dynamic faces evoked increased activation in the visual area V5, in the cuneus, lingual, middle, and temporal cortical regions.
LaBar et al. (2003)	12 healthy participants	fMRI	Stronger activation was observed in the fusiform gyrus, the ventromedial prefrontal cortex and the superior temporal sulcus for dynamic vs. static expressions.
Liang et al. (2017)	20 healthy participants	Neuroimaging with multi-voxel pattern analysis (MVPA)	Compared to static expressions, dynamic stimuli elicited higher neural activation in both face-selective and motion-selective areas.
Trautmann et al. (2013)	19 undergraduate students	Combined fMRI and EEG	The processing of static emotional expressions elicited higher neural activation in the fusiform gyrus for anger and disgust as compared to neutral. Dynamic stimuli activated a more widespread of regions in posterior brain regions and in anterior brain areas.
Trautmann et al. (2009)	16 healthy adults	fMRI	Dynamic expressions elicited enhanced activation in the parahippocampal gyrus (PG), including the amygdala (AMG), fusiform gyrus (FG), superior temporal gyrus (STS), inferior frontal gyrus (IFG), occipital, and orbito-frontal cortex (OFC).
Recio et al. (2011)	21 healthy adults	EEG	Early and late emotion-related ERPs were enhanced with the dynamic expressions. Moreover, dynamic stimuli increased processing in early visual brain areas.
Sato et al. (2004)	22 health participants	fMRI	Greater activation to dynamic vs. static expressions was observed in the inferior occipital gyri, middle temporal gyri, and fusiform gyri over the right hemisphere.
Sato et al. (2010)	21 healthy adults	fMRI	Higher activity in the left amygdala was observed in response to dynamic but not static emotional expressions and this activity was modulated by the intensity of the emotional expressions presented.
Schultz and Pilz (2009)	10 healthy students	fMRI	Dynamics expressions were associated to increased activation in lateral temporal areas, in the visual area V5 and the superior temporal sulcus. Brain regions usually associated with the processing of static face information (bilateral fusiform gyrus and left inferior occipital gyrus) also exhibited higher neural responses to dynamic faces.

CHAPTER 2

INTRODUCTION TO THE EXPERIMENTAL CONTRIBUTION

Although in everyday life we are continuously stimulated by dynamic cues, the use of dynamic stimuli is not prominent in the face processing literature. Most studies on face perception relied on static images and only a little is known about how moving faces are processed. In a paper published in 2009, Cunningham and Wallraven critically reported the following statement: “Any attempt to understand how humans use their eyes, face, and head to communicate that only uses static photographs – whether it is a single photograph or a series of photographs seen one after another – will never be able to explain the perception of expressions. Likewise, any system designed to describe facial expressions that does not explicitly allow for the description of dynamic information will prove ultimately to be inadequate” (2009a, pp. 13–14).

Moreover, as pointed out in very recent reviews by Duchaine and Yovel (2015) and Yovel and O’Toole (2016), there is an important need to investigate face processing with dynamic stimuli, as in our daily life, we are mainly exposed to dynamic faces evolving in natural and unconstrained environments.

Considering the above, the main goal of my thesis was to use dynamic faces to investigate different aspects of face processing in different populations and age groups. The experimental contributions will be briefly introduced in this chapter and then presented in the form of three papers in the following chapters. A general discussion, summarizing the main findings will follow.

2.1 A developmental study

The first experimental contribution of this thesis, presented in **CHAPTER 3**, aimed to investigate how 6-, 9-, and 12-month-old infants extract multisensory gender information from faces and voices. More specifically, the main goal was to address the following question: **Is the extraction of gender from *dynamic* faces and voices influenced by the nature of speech signal, namely by infant- or adult-directed speech?**

As mentioned in the developmental part of **CHAPTER 1**, studies exploring the multisensory perception of gender have yielded inconsistent findings regarding the age of emergence of this ability. Methodological considerations, such as the auditory or visual stimuli presented, the task complexity, or the speech manner, may be partly responsible for this lack of consistency across studies. In order to address a few of these methodological limitations, we decided to conduct a new study using dynamic stimuli, rather than static images of frozen faces (Poulin-Dubois et al., 1998; Poulin-Dubois et al., 1994), because they are richer and ecologically more valid signals of how infants perceive faces in their daily life. Moreover, in order to

investigate whether the extraction of gender from faces and voices is modulated by the very nature of speech signals, we decided to include the two different speech manners to which infants are daily exposed, namely infant- and adult-directed speech.

Infant-directed speech (IDS), in contrast with **adult-directed speech (ADS)**, is a speech manner that is commonly used by parents to communicate with their infants. This speech manner is characterized by both prosodic and linguistic variations (Cooper & Aslin, 1990) used to attract and hold infants' attention. Several studies have revealed that infants exhibit a spontaneous preference for infant-directed over adult-directed speech (Figure 2.1) (Cooper & Aslin, 1990). Slower rhythm, vowel hyper-articulation, frequent breaks, higher pitch, exaggerated facial affects, and the use of special words are all characteristics of IDS (e.g., Cooper & Aslin, 1990; Pegg, Werker, & McLeod, 1992; Trainor, Austin, & Desjardins, 2000). Previous studies have reported that this language has a positive influence on the cognitive (Zangl & Mills, 2007) and early socio-emotional development of infants (Figure 2.2) (Schachner & Hannon, 2011). IDS facilitates, for instance, language acquisition, segmentation of words in a sentence, as well as lexical comprehension (e.g., Aslin, 2000; Fernald et al., 1998; Thiessen, Hill, & Saffran, 2005).

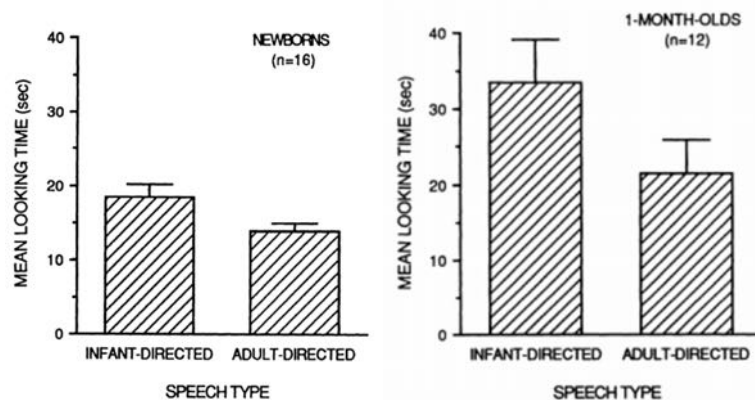


Figure 2.1. Infants preference for infant-directed speech (IDS). Cooper and Aslin (1990) demonstrated that both newborns, and 1-month-old infants exhibited auditory preference for IDS over ADS based on visual fixation duration. The authors concluded from their findings that infants' preference for IDS already present at birth might be influenced by prenatal experience. Reprinted from Cooper and Aslin (1990).

Interestingly, in an event-related potential study, Peter et al. (2016) very recently evidenced differential neural activation to ADS and IDS in both infant and adult brains. More specifically, their findings revealed that the neural responses to IDS, but not to ADS, were

mature in 9-month-old infants. When presented with an IDS sequence embedded in an ADS sequence, infants' brains detected the difference, while this was not the case in the reverse situation (i.e., an ADS sequence in an IDS sequence). Importantly, the differential neural responses observed were interpreted to primarily result from the acoustic salience of the IDS stimuli, such as the exaggerated articulation, the high-pitched intonation, or other phonetic characteristics. In another recent study using fNIRS, Naoi et al. (2012) also observed differential neural responses to IDS and ADS in pre-verbal infants. Higher neural responses to IDS as compared to ADS were found in left- and right-sided temporal regions, varying as a function of age.

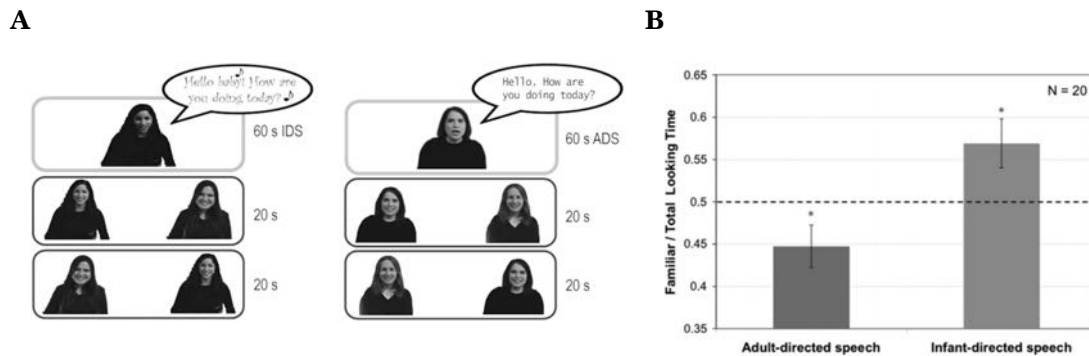


Figure 2.2. IDS conveys cues for social selection. After hearing an individual speak in IDS during a familiarization phase, infants attend more to that individual than to a new person. This social preference is not found in the situation where an individual speaks in ADS during the familiarization phase. Adapted from Schachner and Hannon (2011).

In this first study, we adopted **the preferential looking paradigm** (Fantz, 1961, 1964, 1965) to investigate infants' ability to integrate multisensory information from faces and voices. This procedure has been intensively used over the last decades to gain insights into the development of human face perception. In preferential looking tasks, two stimuli stemming from different categories (e.g., female vs. male faces; own vs. other race faces) are simultaneously presented to infants. When these two exemplars are exhibited together, they compete for infants' gaze and attention. The infants' preference for one category over the other is inferred by measuring looking behavior, the time spent on each stimulus directly reflecting the level of interest. We decided to apply this paradigm, rather than another method, because it has been proven to be a reliable and robust technique (e.g., Pascalis et al., 1995) that has been successfully used in the past to assess infants' preference for own vs. other race faces (Bar-Haim et al., 2006), for attractive vs. less-attractive faces (Langlois et al., 1987; Quinn et al., 2008a; Van Duuren, Kendell-Scott, & Stark, 2003), or female vs. male faces (Quinn et al., 2002).

2.2 A neuropsychological study

The second experimental contribution of this thesis presented in **CHAPTER 4** aimed to clarify the following question: **Are the *dynamic* internal representations of the six basic expressions in a prosopagnosic patient similar to those of healthy controls? Further, are these *dynamic* expressions better recognized by patient PS than *static* images?**

In their influential cognitive model of human face processing, Bruce and Young (1986) proposed that the processing of identity and expression occurs independently. This independence was motivated by findings from the prosopagnosia literature, indicating that patients with severe deficits in recognizing familiar faces demonstrated preserved abilities in categorizing facial expressions (Tranel et al., 1988), and by findings from cognitive studies, suggesting that emotion identification is independent of face familiarity (Young et al., 1986). However, more recent fMRI and neuropsychological studies have led some authors to question the degree of independence between these two processes. By using a principal-component-analysis (PCA) approach, Calder et al. (2001a) for instance evidenced that expression and identity were coded by separate dimensions within a single multidimensional system, implying that the independent perception of facial identity and expression may not necessarily involve separate visuo-perceptual pathways at all stages of face processing. In line with this suggestion and more recently, Palermo and colleagues (2013) have proposed a first common pathway for the processing of both identity and expression, with a separation occurring at a later stage.

In a review published in 2005, Calder and Young have also called into question the independence hypothesis by underlining clinical, statistical, and methodological limitations in patient-based evidence. As pointed out by the authors, more evidence from neuropsychological research is required to advance our understanding of the processing of these different facial aspects. The second experimental contribution of this thesis re-examined this debated issue with four behavioral experiments investigating the ability of a single-case of acquired prosopagnosia to recognize static and dynamic emotional faces. In Experiment 1, we assessed patient PS's ability to categorize static expressions. In Experiment 2, we used an exciting and innovative 4D technique developed by Yu, Garrod, and Schyns (2012, see below) to map out the facial features used by patient PS when instructed to categorize random facial muscle activations into the six basic emotions. Facial information used for identity has been previously investigated in this patient with the Bubbles technique (for more details on this technique, see [Box 6](#)). These findings have reported that the patient used facial features in a suboptimal way when asked to recognize the identity of previously seen faces, focusing on the mouth and the external contours (Caldara et al., 2005). An optimal use of all facial features for the categorization of dynamic emotional faces would sharply contrast with her suboptimal way of extracting information in

identity recognition tasks (Caldara et al., 2005), favoring the view of independence between both processes. Finally, in Experiment 3 and 4, we assessed patient PS's ability to categorize the static and dynamic random facial activations that she and healthy controls categorized as expressions in Experiment 2.

The technique applied in our second study, **an innovative 4D technique coupled with a reverse correlation technique**, is a perception-driven platform, developed by Yu, Garrod, and Schyns in 2012. This technique modulates the activation of actions units from the Facial Action Coding System (FACS) to flexibly synthesize 3D facial expressions. More specifically, a subset of action units with specific temporal parameters is selected by a FACS-based generator on each trial. In the example presented in Figure 2.3, AU9, AU10L, and AU19 are selected (with specific colored labels). The colored curves represent the amplitude of acceleration and deceleration of movement for each specific AU over time. The random facial muscle activations are then applied to a neutral-face identity.

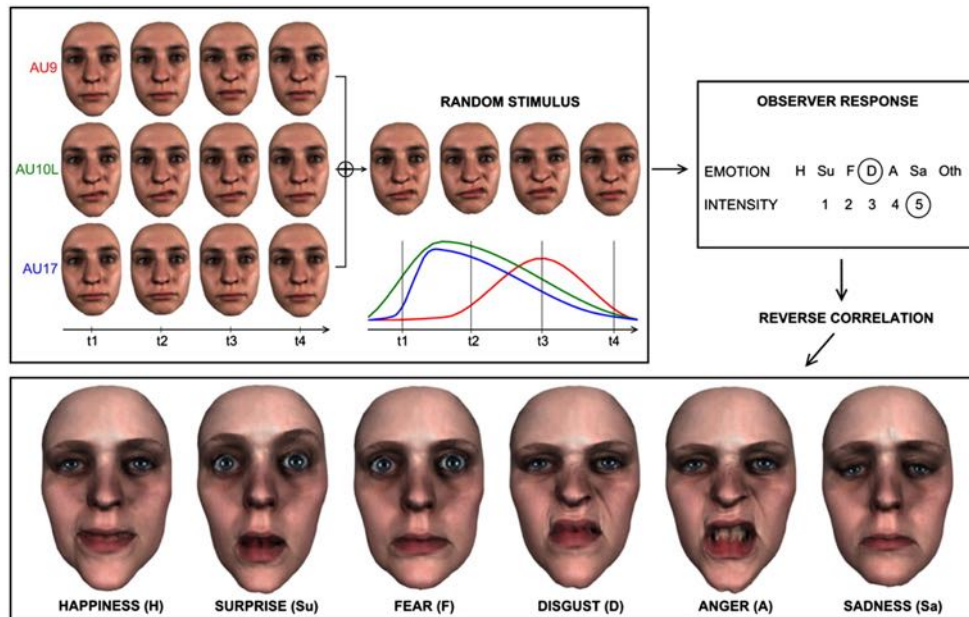


Figure 2.3. FACS synthesis. Random facial muscle activations were presented to human observers who were asked to categorize them according to the six basic expressions (H = Happiness, Su = Surprise, F = Fear, D = Disgust, A = Anger, Sa = Sadness, Oth = Other). By reverse correlating the randomly presented AUs with human observers' responses, Yu et al. (2012) reconstructed the internal representations of the six basic expressions for each observer. Reprinted from Yu et al. (2012).

Following this synthesis, human observers are instructed to categorize the parametrized facial movements in a 7-Alternative-Forced-Choice (7AFC) facial expression recognition task (6 basic expressions plus an “Other” response) and further rate the perceived emotional intensity on a five-point scale (“Very Weak”, “Weak”, “Medium”, “Strong”, and “Very Strong”). In the example displayed in Figure 2.3, observers interpreted the randomly presented facial animation as a meaningful expression of disgust at medium intensity, because the action units activated corresponded to their internal mental representation of that expression.

The following step uses methods from human psychophysics and computer graphic techniques to reverse correlate human observers’ answers with the perceptually relevant expression parameters (Figure 2.4). This procedure enables the reconstruction of 3D dynamic models of the internal representations of the six basic expressions for each observer. The six expressions presented at the lower part of Figure 2.3 correspond to the internal representations of one human observer (Yu et al., 2012).

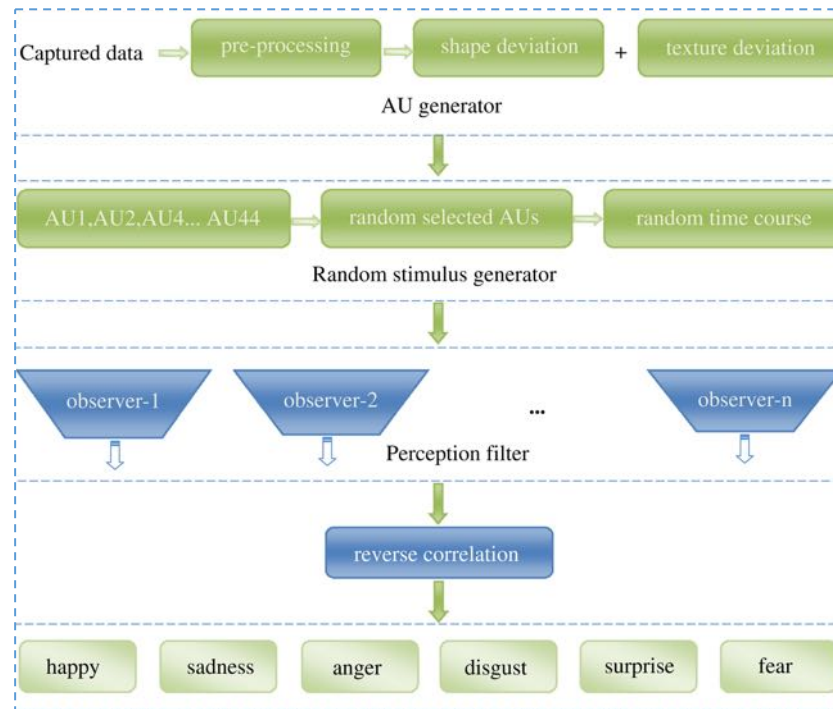


Figure 2.4. The procedure used by Yu et al. (2012) to reconstruct the internal representations of the six basic expressions in human observers. Reprinted from Yu et al. (2012).

2.3 A cross-sectional study

Dynamic faces transmit a wealth of visual signals that are critical for the optimal categorization of facial expressions of emotion. Unlike static snapshots that are used in most experiments, naturally evolving dynamic expressions provide observers with richer and ecologically more valid signals that should facilitate their processing. While results obtained from healthy observers are equivocal, an advantage for dynamic expressions has been previously reported in several clinical conditions (Back et al., 2007; Uono et al., 2010) as well as in brain-damaged patients (Adolphs et al., 2003; Humphreys et al., 1993). Consistent with these results, in the second experimental contribution of this thesis (**CHAPTER 4**), we found that our prosopagnosic patient PS performed on a comparable level to healthy controls on a facial expression recognition task when dynamic stimuli were presented, but exhibited impairments in recognition with static images.

Following on from this previous work, the last experimental contribution of this thesis presented in **CHAPTER 5** examined the following question: **Besides prosopagnosic individuals, do other populations with immature (i.e., young children) or fragile (i.e., elderly adults) face processing systems, also benefit from the presentation of *dynamic* emotional signals?**

To this aim, we conducted a large sample cross-sectional study ($N = 412$) in order to investigate whether the recognition of facial emotional expressions changes from childhood to old age and whether dynamic information facilitates their processing at any given point in development. We assessed participants' abilities to categorize expressions presented as dynamic, static, or shuffled (temporally randomized) events. In this third study, we relied on a database of stimuli created by Gold et al. (2013), who adopted an **ideal observer approach** to objectively measure whether the amount of information provided by static, dynamic, and shuffled stimuli was identical in all three conditions. We decided to rely on this database to rule out the possibility that differences in recognition abilities across ages are modulated by differences in the amount of discriminative information provided by the expressions. By using this approach, Gold et al. (2013) compared human expression recognition thresholds and efficiencies to the performance of a statistically optimal observer. The ideal observer model is based on an optimal Bayesian rule that takes into account all the information available to maximize the probability to get an accurate response. This model has been frequently used in psychophysics tasks to measure the physical or perceptual properties of the stimuli presented (Lai et al., 2014). This approach enabled Gold et al. (2013) to reveal that their dynamic stimuli did not provide additional information to human observers than their static images (Figure 2.5A). Moreover, when the information physically available was considered, the corresponding human

efficiencies for recognizing the expressions were nearly the same across all three conditions (dynamic = static = shuffled) (Figure 2.5B).

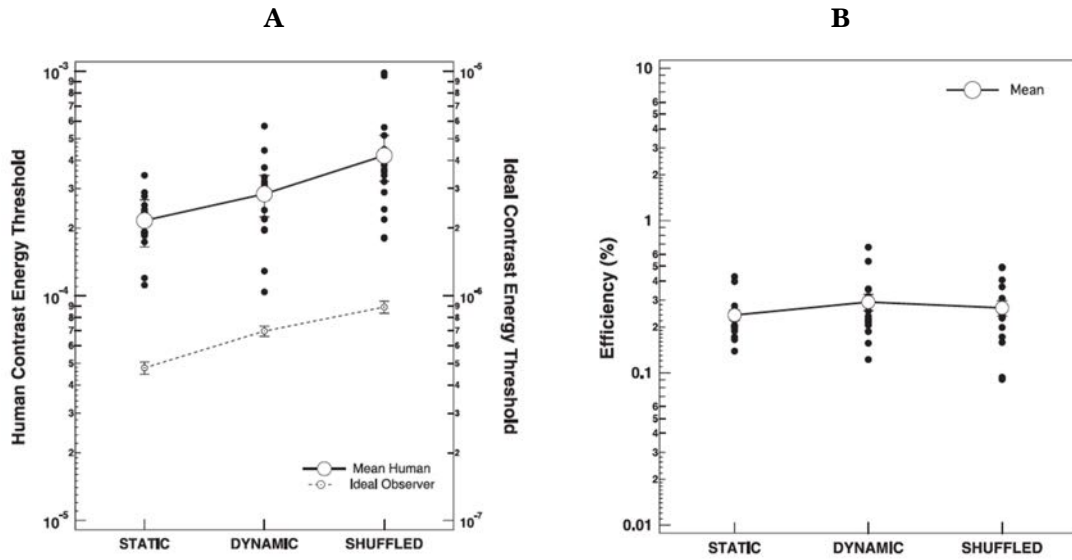


Figure 2.5. The recognition of static and dynamic expressions using an ideal observer approach. (A) Contrast energy thresholds for recognizing static, dynamic, and shuffled expressions in ideal and human observers. **(B)** The corresponding efficiencies were nearly identical across conditions. Adapted from Gold et al. (2013).

CHAPTER 3

AUDIO-VISUAL PERCEPTION OF GENDER BY INFANTS EMERGES EARLIER FOR ADULT-DIRECTED SPEECH

Richoz, A.R., Quinn, P.C., Hillairet de Boisferon, A., Berger, C., Loevenbruck, H., Lewkowicz, D.J., Lee, K., Dole, M., Caldara, R., & Pascalis, O. (2017). Audio-visual perception of gender by infants emerges earlier for adult-directed speech. *PLoS One*, 12(1), 1-15.

Research highlights

- ◇ Speech manner (IDS vs. ADS) affects multisensory perception of gender in infants
- ◇ Infants from 6 months of age matched female faces and voices with ADS
- ◇ Infants from 9 months of age matched female faces and voices with IDS
- ◇ Early multisensory abilities are shaped by the circumstances of social interaction

Keywords: Gender-speech matching, Infants, Multisensory, Speech manner, Infant-directed speech, Adult-directed speech

3.1 Abstract

Early multisensory perceptual experiences shape the abilities of infants to perform socially-relevant visual categorization, such as the extraction of gender, age, and emotion from faces. Here, we investigated whether multisensory perception of gender is influenced by infant-directed (IDS) or adult-directed (ADS) speech. Six-, 9-, and 12-month-old infants saw side-by-side silent video-clips of talking faces (a male and a female) and heard either a soundtrack of a female or a male voice telling a story in IDS or ADS. Infants participated in only one condition, either IDS or ADS. Consistent with earlier work, infants displayed advantages in matching female relative to male faces and voices. Moreover, the new finding that emerged in the current study was that extraction of gender from face and voice was stronger at 6 months with ADS than with IDS, whereas at 9 and 12 months, matching did not differ for IDS versus ADS. The results indicate that the ability to perceive gender in audiovisual speech is influenced by speech manner. Our data suggest that infants may extract multisensory gender information developmentally earlier when looking at adults engaged in conversation with other adults (i.e., ADS) than when adults are directly talking to them (i.e., IDS). Overall, our findings imply that the circumstances of social interaction may shape early multisensory abilities to perceive gender.

3.2 Introduction

Human faces provide multisensory inputs to infants, exposing them not only to visual information but also to the voice and language of their caregivers. This perceptual experience shapes early multisensory abilities that are critical for the development of social categories related to vocalizing and talking human faces.

One social category to which infants have extensive exposure is gender. The ability of humans to process male and female faces has been widely studied over the last two decades (e.g., Brown & Perrett, 1993; Bruce et al., 1993). Adults reliably and rapidly identify facial gender, even when relying only on individual facial features, such as eyebrows, jaw, chin, or mouth (Brown & Perrett, 1993; Bruce et al., 1993; O'Toole et al., 1998). In infants, the ability to categorize face gender develops between three months and one year of age (e.g., Cohen & Strauss, 1979; Cornell, 1974; Fagan & Singer, 1979; Leinbach & Fagot, 1993; Younger & Fearing, 1999). Even though some research suggests that infants form categories for both female and male faces (Cornell, 1974; Younger & Fearing, 1999), other studies point to a developmental asymmetry in the acquisition of gender categories: when infants are presented with a category of male faces, they subsequently prefer female over novel male faces, but when infants are presented with a category of female faces, they do not subsequently prefer male over novel female faces (Leinbach & Fagot, 1993). The asymmetry has been further shown to reflect a spontaneous preference for female over male faces when the primary caregiver is female (e.g., Liu et al., 2015; Quinn et al., 2002; for a review see, Ramsey, Langlois, & Marti, 2005). These studies also suggest that infants develop more structured representations for female faces than for male faces. The processing advantages for female over male faces are in turn believed to reflect experiential differences between female and male faces. Infants reared by a female primary caregiver experience over 70% female faces and less than 30% male faces in the first months of life (Liu et al., 2015; Rennels & Davis, 2008; Sugden, Mohamed-Ali, & Moulson, 2014).

Even though gender processing is mainly based on visual properties of the face, vocal cues also play an important role. Adults perceive faces and voices as coherent entities (e.g., McGurk & MacDonald, 1976; Rosenblum, 2008) and face-voice associations can be the basis for gender identification, e.g., long hair and thin and softly curved eyebrows associated with high-pitched voice (Barenholtz et al., 2014). Although infants perceive audiovisual coherence of speech syllables as early as 2 to 4 months of age (Kuhl & Meltzoff, 1988; Patterson & Werker, 1999, 2002), their ability to use audio-visual correspondences to respond to gender emerges in the second half of the first year of life, and consistent with the work on visual categories, is restricted to female faces (Hillairet de Boisferon et al., 2015; Poulin-Dubois et al., 1994). For

example, the work of Walker-Andrews et al. (1991) has revealed that 6-month-old infants were able to reliably match synchronous faces and voices when presented with gender information. However, to ensure perfect synchrony between faces and voices, Walker-Andrews et al. (1991) asked each actor to dub their voice onto the video-recordings of their own face. This procedure made it difficult to tease apart whether infants genuinely matched faces and voices based on gender information or whether they made the match based on idiosyncratic relations between faces and voices. This latter possibility is supported by a more recent study demonstrating that 6-month-old infants are able to link idiosyncratic cross-modal identity cues of unfamiliar faces (Trehub, Plantinga, & Brcic, 2009). Subsequent studies have reported that the emergence of audio-visual perception of gender can vary from 6 months to 8 months to even later (Hillairt de Boisferon et al., 2015; Patterson & Werker, 2002; Poulin-Dubois, Serbin, & Derbyshire, 1998; Poulin-Dubois et al., 1994; for a review see, Ramsey et al., 2005, pp. 219-229; Walker-Andrews et al., 1991).

The specific timing of the emergence of the ability to perceive the multisensory coherence of gender might depend on the different types of visual and auditory stimuli used across the different studies and the methodology of the tasks (Ramsey et al., 2005). Some studies investigating perception of multisensory gender coherence have presented dynamic faces (Hillairt de Boisferon et al., 2015; Patterson & Werker, 2002; Walker-Andrews et al., 1991), whereas others have relied on static images (Poulin-Dubois et al., 1994). Use of dynamic faces provides a more ecologically valid approach to investigate multisensory perception of gender, as our natural environment is surrounded with dynamic multisensory cues (Xiao et al., 2014). Dynamic facial cues also seem to play a critical role in the way faces are encoded (Xiao et al., 2015) and recognized (Otsuka et al., 2009). Moreover, dynamic faces seem to influence facial scanning, prompting infants to shift their fixations to different facial features (Xiao et al., 2015). Multiple fixation shifts on major facial areas are likely to benefit gender processing given that the visual cues for gender can be found throughout the face (Brown & Perrett, 1993; Bruce et al., 1993; O'Toole et al., 1998). The auditory stimuli used to investigate multisensory perception of gender have also varied across different studies. Some studies have presented isolated vowels (Patterson & Werker, 2002), whereas others have played recordings of fluent and continuous speech (Poulin-Dubois et al., 1998; Poulin-Dubois et al., 1994; Walker-Andrews et al., 1991). The latter may facilitate the detection of the gender of the speaker via additional cues such as intonation, stress, duration, respiratory patterns, and vocal breathiness (Klatt & Klatt, 1990; Van Borsel, Janssens, & De Bodt, 2009).

It is additionally possible that the manner of speech (infant- or adult-directed) influences the perception of multisensory gender coherence. In daily life, parents or siblings interact with infants using infant-directed speech (IDS), varying at both prosodic and linguistic

levels (Cooper & Aslin, 1990). Infants are particularly sensitive to prosody (Fernald, 1985; Kim & Johnson, 2014; Nakata & Trehub, 2004) and largely prefer IDS to adult-directed speech (ADS) (Cooper & Aslin, 1990). IDS is characterized by better articulation, higher pitch, slower rhythm, many breaks, and the use of special words (e.g., Cooper & Aslin, 1990; Pegg, Werker, & Mcleod, 1992; Trainor, Austin, & Desjardins, 2000).

Several studies have shown that the manner of speech has an influence on the cognitive (Zangl & Mills, 2007) and social development of infants (Schachner & Hannon, 2011) with IDS promoting, for example, language acquisition, inclusive of word segmentation and lexical comprehension (e.g., Aslin, 2000; Fernald et al., 1998; Thiessen, Hill, & Saffran, 2005). Based on this prior evidence, it could be reasoned that IDS might also facilitate the extraction of gender information from face and voice. However, even though IDS facilitates lexical learning and word segmentation (e.g., Aslin, 2000; Thiessen et al., 2005), its facilitating role in perceiving multisensory gender coherence might be minimal or non-existent. Indeed, the high pitch of IDS, which brings male voices closer to the daily frequencies of female voices, might even impede the ability of infants to perceive multisensory gender coherence. Consistent with this suggestion, Trainor and Desjardins (2002) demonstrated that the high pitch of IDS negatively affected vowel discrimination in 6- to 7-month-old infants. The authors concluded that rather than being a facilitator of vowel learning, the primary role of high pitch in IDS is to attract attention and communicate intention and emotion.

Using dynamic stimuli and exaggerated prosodic nursery rhymes characteristic of IDS, Hillairet de Boisferon et al. (2015) recently investigated the developmental emergence of the ability to perceive the coherence of auditory and visual attributes of gender in 6- and 9-month-old infants. Infants viewed two side-by-side video clips of a man and a woman reciting a nursery rhyme and heard a synchronous male or female soundtrack. Infants did not associate audible and visible gender attributes until the age of 9 months, and only for female faces. The authors interpreted these findings as evidence that a combination of different factors (e.g., stimulus and task complexity, and amount of perceptual experience) influence the multisensory responsiveness of infants.

Because Hillairet de Boisferon et al. (2015) only investigated IDS and not ADS, we conducted a new study to clarify whether and how the manner of speech may influence the perception of multisensory gender coherence. The experiment consisted of four 12-s trials, during which infants saw two side-by-side silent video-clips of actors (a male and a female) reciting a nursery rhyme; they were the same video-clips used by Hillairet de Boisferon et al. (2015). A soundtrack with a female or a male voice telling a story in IDS or ADS was played at the same time. Infants participated in only one condition, either IDS or ADS, and heard a male

voice on half of the trials and a female voice on the other half of the trials. If the infants could detect the gender of the audio-visual correspondence, we expected them to look longer at the face whose gender matched the gender of the accompanying voice than at the face that did not match it. We decided to use dynamic video recordings rather than static faces because dynamic stimuli provide observers with richer and ecologically more valid signals of the sorts of events that infants experience in their daily lives (e.g., Knight & Johnston, 1997; O'Toole, Roark, & Abdi, 2002; Xiao et al., 2014).

We reasoned that if IDS affects the extraction of gender from face and voice in the same way that it affects word segmentation and lexical comprehension, then the perception of multisensory gender coherence would emerge earlier with IDS than with ADS. However, if IDS draws attentional resources to the prosodic features and the linguistic content of the speech rather than to the gender of the talking face (Hillaiet de Boisferon et al., 2015), then the perception of multisensory gender coherence might emerge later with IDS than with ADS. Moreover, the high-pitched intonation characteristic of IDS might bring male voices closer to the usual daily life frequencies of female ADS, which might make it more difficult for infants to match an IDS utterance to either a male or female face.

An additional expectation was that the perception of multisensory gender coherence might be specific to female faces. This prediction is supported by results from previous intersensory matching studies (Hillaiet de Boisferon et al., 2015; Poulin-Dubois et al., 1994), and findings that infants exhibit a spontaneous preference for female over male faces (Liu et al., 2015; Quinn et al., 2002), as well as findings that infants possess a more advanced category representation for female than male faces (Quinn et al., 2002; Ramsey et al., 2005).

3.3 Materials and Method

Participants

6-month-olds (11 females) ($M_{age} = 195$ days, $SD = 4$ days), 9-month-olds (12 females) ($M_{age} = 283$ days, $SD = 5$ days), and 12-month-olds (13 females) ($M_{age} = 379$ days, $SD = 6$ days) were included in the analyses ($n = 31$ per age category). Infants were healthy, full-term participants, recruited from the maternity of the Centre Hospitalier Universitaire of Grenoble in France. They were all Caucasians and living in a French speaking environment. Eighteen additional infants were tested, but excluded from the analyses because they did not complete the procedure due to crying ($n = 6$) or fussiness ($n = 12$).

Stimuli

We used the same visual stimuli as those used by Hillaiet de Boisferon et al. (2015), although

we did not use the soundtracks presented in that study. Instead, we presented new soundtracks that are described below. The video recordings consisted of six Caucasian adult faces (3 females) reciting a nursery rhyme in French while maintaining a neutral facial expression. During the video recordings, actors were asked to look directly at the video camera and to recite at the same tempo as a model whose video was played as a template before the recordings. The videos in each pair were edited to ensure that they started with the same opening of the mouth (Hillaiet de Boisferon et al., 2015). The faces were recorded against a green background and neither of the actors wore gender-specific visual information, such as makeup or jewelry. The faces subtended a visual angle of 19.09° (vertical) and 14.32° (horizontal).

As in previous studies (e.g., Hillaiet de Boisferon et al., 2015; Patterson & Werker, 2002), two adults (1 male) who were different from those seen in the videos were selected to record the soundtracks. The content of the soundtracks differed from the nursery rhyme told by the actors in the videos (i.e., the speech did not correspond with the articulatory motions of the actors). This was done to ensure that infants were genuinely representing gender across face and voice (i.e., extracting the amodal invariance of gender *per se*) and not simply responding to the co-occurrence of the two streams of information or making the match based on speaker idiosyncrasies in visible and audible articulatory or respiratory patterns. We registered the audio-recordings with a high-quality audio microphone in a special soundproof recording room. We asked the two speakers to provide four recordings: each of two stories was recorded in both ADS and IDS. The short “stories” (each a 12-s long utterance) were as follows:

Story 1: “Moi je trouve que c’est une belle journée aujourd’hui. Je vois le soleil qui brille. C’est vraiment super que tu sois venu me voir”.
(English translation: “It’s a beautiful day today. I see the sun shining. It’s really nice that you came to see me.”)

Story 2: “Tu es venu de loin pour me voir aujourd’hui. C’est vraiment super de t’avoir rencontré. J’espère que c’est le cas aussi pour toi. Je trouve que tu es un très beau bébé.”
(English translation: “You came from far away to see me today. It’s really nice to meet you. I hope it is also the case for you. You are a beautiful baby.”)

The semantic content of both stories was very similar. In order to have 12-s length stories, the second story included four sentences instead of three as in the first story. The number of words in each sentence was, however, similar. Twenty French-speaking adults were asked to categorize the stories as adult or infant-directed. Results revealed strong agreement between observers as indicated by the Fleiss’ kappa of 0.73.

The utterances were acoustically analyzed using Praat (Boersma & Weenink, 2011). Pitch level was computed as the mean of the fundamental frequency (f0) values for each utterance as a whole. Pitch span was computed by subtracting the minimum from the maximum f0 value for each utterance as a further index of pitch variation. The acoustic analyses of the voices heard by the infants in our study revealed the following fundamental frequencies: For IDS, the male voice varied between 102 and 289 Hertz (Hz) and the female voice varied between 158 and 502 Hz. For ADS, these fundamental frequencies ranged between 87 and 178 Hz for the male voice and between 158 and 338 Hz for the female voice. When converted to semitones (st), the difference between the maximum f0 value for the male voice in IDS and the maximum for the female voice in ADS was less than 3 st. The average values for the female and male IDS and ADS voices are reported in Table 3.1 in Hertz and Table 3.2 in semitones (relative to 100 Hz).

Table 3.1 Fundamental frequency characteristics in hertz (Hz). Average pitch level, pitch minimum, pitch maximum, and pitch span of the fundamental frequencies (Hz) of the ADS and IDS male and female voices.

Voice	Condition	Pitch level (Hz)	Pitch min (Hz)	Pitch max (Hz)	Pitch span (Hz)
Female	IDS	255.27	158.20	502.48	344.28
Male	IDS	166.43	102.86	289.98	187.12
Female	ADS	210.69	158.40	338.74	180.34
Male	ADS	116.47	87.71	178.19	90.47

Table 3.2. Fundamental frequency characteristics in semitones (st). Average pitch level, pitch minimum, pitch maximum, and pitch span of the semitones (st) of the ADS and IDS male and female voices.

Voice	Condition	Pitch level (st)	Pitch min (st)	Pitch max (st)	Pitch span (st)
Female	IDS	15.79	7.87	27.95	20.07
Male	IDS	8.23	0.48	18.41	17.93
Female	ADS	12.70	7.95	21.13	13.17
Male	ADS	2.41	-2.26	9.92	12.18

Apparatus and Procedure

Infants were seated on the lap of a parent in a dimly illuminated room, 60 cm away from a 22-inch computer screen. Parents were asked not to intervene, interact, or speak with their infant during the experiment. All parents signed a written consent form for their infant prior to the

experiment. The Ethical Committee of the Centres d'Investigation Clinique de l'inter-région Rhône-Alpes-Auvergne (authorization number IRB 2010-21) approved the study reported here.

The experiment consisted of four 12-s trials for each condition (i.e., IDS or ADS) with two different sets of faces and corresponding soundtracks. Infants participated in only one condition, either IDS or ADS. We divided the infants of each age category into two groups ($n = 15$ or 16 participants per group). On each trial, infants were presented with two side-by-side silent video-clips of faces (one male and one female) reciting a nursery rhyme. The faces were separated by a 15-cm gap. We used the same video-clips on Trials 1 and 3, and on Trials 2 and 4, but with switched left-right locations of the faces, so that each face was seen on the right and the left side of the screen (Figure 3.1). The side of gender presentation (male or female) was counterbalanced across infants on the first trial and reversed on the following trials (on Trials 3 and 4, see Figure 3.1). A soundtrack with a female or a male voice telling a story either in IDS or ADS was played at the same time as the video-clips. On each trial, infants listened to either a male or female voice, with the woman's voice presented twice over the four trials (i.e., for a given infant, across trials, the two female-male face pairings were presented once with a female voice and once with a male voice).

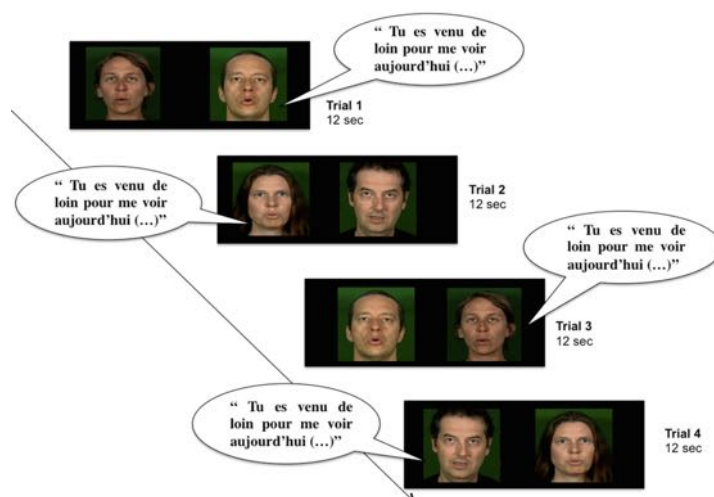


Figure 3.1. Schematic representation of the procedure. Infants saw two side-by-side silent video-clips presenting a male and a female face reciting a nursery rhyme. A soundtrack with a female or a male voice telling a story in IDS or ADS was played at the same time. The video-clips were repeated twice throughout the four trials, with left-right reversal for the positioning of the faces.

We used the E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA, USA) to conduct the experiment. Two loudspeakers (Dell A225) placed behind the screen, halfway between the two faces, transmitted the audio stimuli. A low-light video camera, located above the stimulus-presentation monitor, was used to record infant looking behavior. The video recordings were subsequently digitized and analyzed with a frame-by-frame coding procedure. We used a preferential looking technique to test multisensory gender perception of faces and voices and measured the total duration of looking time directed to each face.

3.4 Results

We performed analyses on the mean proportion of total looking time (PTLT) that each infant directed at the matching faces over the four test trials. For this calculation, we divided the amount of looking at the matching face by the total amount of looking at both faces during each trial and then averaged the two proportions for each gender over the four test trials. If infants perceived multisensory gender coherence, then they should direct more looking time to the face that matched the heard voice. The data were collapsed over story number (1 vs. 2) because it did not affect looking behavior, $F(1, 91) = .26$, $p = .60$, $\eta^2_p = .003$. Preliminary analysis of participant gender revealed no significant main effect or interaction of this factor, $F(1, 91) = 2.52$, $p = .13$, $\eta^2_p = .02$. This factor was therefore not included in the subsequent analyses.

A three-way mixed ANOVA was first conducted on the PTLTs directed at the matching face with condition (IDS, ADS) and age (6-, 9-, 12-months) as between-subjects factors and gender of the voice (male, female) as a within-subjects factor. Results first revealed that gender of the voice affected infant responsiveness, $F(1, 87) = 9.82$, $p < .005$, $\eta^2_p = .10$. As expected, infants looked longer at the matching face in the presence of a female ($M = 56.30\%$, $SD = 10.99\%$) than a male voice ($M = 51.26\%$, $SD = 10.93\%$). A one-sample t -test against the chance value of 50% revealed that the PTLTs directed at the female matching faces were significantly different from chance, $t(92) = 5.52$, $p < .001$ (two-tailed), Cohen's $d = .57$. This was not the case for the PTLTs directed at the male matching faces, $t(92) = 1.11$, $p = .27$, Cohen's $d = .11$. Results further showed that condition (IDS vs. ADS) affected responsiveness, $F(1, 87) = 6.33$, $p < .05$, $\eta^2_p = .06$, reflecting the fact that the PTLTs directed at the face that matched the heard voice were higher in the ADS ($M = 55.73\%$, $SD = 7.80\%$) than IDS condition ($M = 51.86\%$, $SD = 7.51\%$). There was also a statistically significant two-way interaction between condition and age, $F(2, 87) = 3.73$, $p < .05$, $\eta^2_p = .07$, indicating that the total amount of time spent on the face that matched the heard voice was not the same for the three age groups in each condition (Figure 3.2).

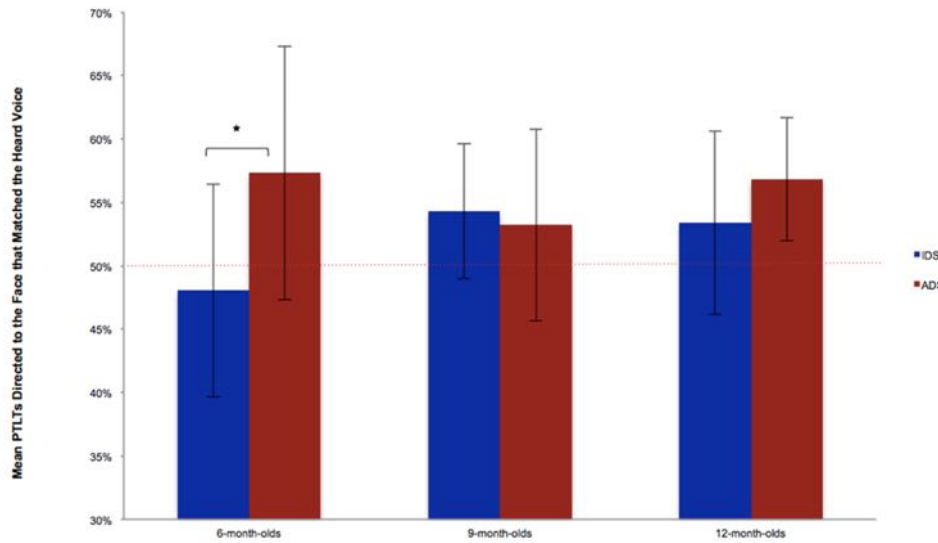


Figure 3.2. Mean proportion of total looking time to the face that matched the heard voice in each condition (IDS, ADS) and for each age group (6-, 9-, 12-month-olds). Note. * $p < .05$; Error bars represent SDs.

Post-hoc t-tests revealed that this effect was driven by the 6-months-olds who gazed significantly longer at the matching face in the ADS ($M = 57.32\%$, $SD = 9.98\%$) than in the IDS condition ($M = 48.06\%$, $SD = 8.40\%$), $t(29) = 2.80$, $p < .05$, Cohen's $d = .51$. For this age group, one-sample t-tests against the chance value of 50% revealed that the PTLTs directed at the face that matched the heard voice were significantly different from chance in the ADS condition, $t(14) = 2.84$, $p < .05$, Cohen's $d = .75$, but not in the IDS condition, $t(15) = -.92$, $p = .37$, Cohen's $d = -.23$. In 9-month-olds, no differences were found between the IDS and ADS condition, $t(29) = .45$, $p = .65$, Cohen's $d = .08$. In this age group, the PTLTs directed at the face that matched the heard voice were different from the chance value of 50%, $t(30) = 3.20$, $p < .005$, Cohen's $d = .58$. The same pattern of results was found for 12-month-olds, showing no differences between both conditions, $t(29) = -1.54$, $p = .13$, Cohen's $d = -.28$, but a statistically significant difference from the chance value of 50% for the PTLTs directed at the matching face, $t(30) = 4.42$, $p < .001$, Cohen's $d = .80$. The three-way interaction between condition, gender of the voice, and age was not statistically significant, $F(2, 87) = .19$, $p = .82$, $\eta^2_p = .004$.

3.5 Discussion

The aim of the present study was to investigate whether infant- (IDS) or adult-directed (ADS) speech influences the perception of multisensory gender coherence. There were two possibilities. Given the facilitative effects of IDS observed for language acquisition, it might be the case that

IDS could facilitate the extraction of gender from face and voice. However, as a language facilitator, it is also possible that IDS draws attentional resources to the prosodic features and the linguistic content of speech rather than to the gender of a talking face (Hillairiet de Boisferon et al., 2015). Based on this possibility and on the idea that the higher pitch in IDS is likely to reduce gender differences in the voices of speakers, one might observe a later emergence of multisensory perception of gender with IDS than with ADS. Our data are consistent with the latter possibility. Infants gazed longer at the face that matched the heard voice in the ADS than in the IDS condition. Furthermore, our results showed a significant interaction between age and condition (ADS, IDS) reflecting the fact that the time spent by each age group on the face that matched the heard voice was not the same in each condition. At 6 months, infants were more likely to match gender across face and voice with ADS than with IDS. At the two older age groups, there was no difference in matching performance between ADS and IDS. An additional expectation was that the perception of multisensory gender coherence might be specific to female faces. Our findings supported this prediction as they revealed that the gender of the voice affected infant responsiveness. Infants displayed advantages in matching female relative to male faces and voices.

On one hand, our findings that the perception of multisensory gender coherence emerges between 6 and 9 months of age when associated with IDS are in line with previous studies. For example, Patterson and Werker (2002) found that 8-month-old infants were able to match faces and voices when presented with infant-directed vowels. Similarly, Hillairiet de Boisferon et al. (2015) previously demonstrated that the perception of multisensory gender coherence emerges between 6 and 9 months of age when infants were exposed to faces reciting a nursery rhyme. On the other hand, our findings differ somewhat from the findings reported by Walker-Andrews et al. (1991) who reported successful matching of auditory and visible attributes of gender by 6 months of age when infants were presented with infant-directed nursery rhymes. The earlier emergence of matching observed by Walker-Andrews et al. (1991) may be due to the fact that the voices heard by the infants were the actual voices of the two speakers presented in the videos, raising the possibility that idiosyncratic relations between faces and voices might have facilitated matching. Prosodic bits of information, such as acoustic parameters, intensity, lip motion, and jaw movement are known to differ across individuals (e.g., Kelso et al., 1985; Loevenbruck, 1999; Loevenbruck, 2000) as is the pronunciation of segmental and suprasegmental information (e.g., Dohen, Loevenbruck, & Hill, 2009; Keating et al., 1994). Given that the actors in Walker-Andrews et al. (1991) were asked to dub their voice onto their own video recordings to ensure perfect synchrony, the possibility that they produced idiosyncratic articulatory or respiratory patterns which facilitated auditory and visual gender matching cannot be ruled out.

To control for possible idiosyncratic intersensory relations, speaker identity cues, or individually-specific respiratory patterns, we selected different voices that belonged to none of the actors presented in the video recordings. In this way, and as in previous studies (Hillairret de Boisferon et al., 2015; Patterson & Werker, 2002), we ensured that infants were genuinely representing gender across face and voice. Also, the auditory and visual information differed in that the auditory information was a different utterance than the one told by the actors in the videos. We chose these types of stimulus materials so to ensure that the common information across the auditory and visual modalities was gender. It is additionally worth noting that in the Walker-Andrews et al. (1991) procedure, infants saw only one pair of female and male faces. In contrast, in our study as well as in previous ones (Hillairret de Boisferon et al., 2015; Poulin-Dubois et al., 1998; Poulin-Dubois et al., 1994), infants were exposed to two pairs of female and male faces. As different identities were included rather than only one prototypical exemplar of a specific category (e.g., long-haired blond woman with blue eyes), which probably facilitated specific matching, our procedure more likely tested more general infant perception of multisensory gender coherence.

There are several differences between IDS and ADS that may contribute to the developmentally earlier matching of audible and visible cues in infants listening to ADS. Substantial evidence has shown that IDS influences the cognitive development of infants, promoting language acquisition, inclusive of lexical comprehension (Thiessen et al., 2005), word recognition (Fernald et al., 1998), and segmentation in a sentence (e.g., Aslin, 2000; Thiessen et al., 2005). As noted, these characteristics might suggest that IDS should facilitate infant perception of multisensory gender coherence. However, IDS may not facilitate all aspects of language acquisition. For example, Trainor and Desjardins (2002) reported that the high pitch typical of IDS impaired vowel discrimination in 6- to 7-month-olds, because of acoustic parameters and the space between harmonics in high-pitched sounds (Trainor & Desjardins, 2002). Moreover, when using exaggerated infant-directed nursery rhymes, Hillairret de Boisferon et al. (2015) recently demonstrated that infants did not match audible and visible attributes of gender until the age of 9 months, and only for female faces. As observed by Hillairret de Boisferon (2015), a combination of various factors, including attentional resources, and stimulus and task complexity, might account for the later age estimate of emergence for multisensory gender coherence. IDS is characterized by slower tempo, shorter utterances, longer pauses, better articulation, as well as higher pitch (e.g., Fernald, 1992; Lewkowicz, 1998; Trainor et al., 1997; Trehub & Nakata, 2001-2002). Critically, the typically higher pitch of IDS (Andruski & Kuhl, 1997; Fernald & Simon, 1984; Papousek, Papousek, & Haekel, 1987) may affect the perception of multisensory gender coherence in infants. The usual fundamental frequency of a male voice varies between 90 and 140 Hz and that of a female voice between 170 and 290 Hz for ADS (Fernald et al., 1989). For IDS, this fundamental frequency can range

from 120 to 190 Hz for male voices and from 250 to 450 Hz for female voices (Fernald et al., 1989). The acoustic analyses of the particular voices heard by the infants in our study revealed similar fundamental frequencies for ADS and IDS (see Tables 3.1 and 3.2). The higher pitch in IDS thus brings male voices closer to the usual daily life frequencies of female ADS, which might create a possible basis for why younger infants have more difficulty matching gender when presented with an IDS utterance corresponding to either the male or female face.

Compared to ADS, IDS has also been argued to be effective for attracting and holding infant attention (e.g., Fernald & Simon, 1984; Rose, Feldman, & Jankowski, 2003). It has been associated with higher social responsiveness as compared to ADS (Dunst & Hamby, 2012). Again, one might expect that increased attention would lead to stronger perception of multisensory gender coherence in infants. However, as IDS also promotes learning (e.g., Csibra & Gergely, 2006; Kaplan et al., 2002; Kaplan, Bachorowski, & Zarlengo-Strouse, 1999) and language acquisition (e.g., Liu, Kuhl, & Tsao, 2003; Thiessen et al., 2005; Werker et al., 2007), one can speculate that greater attentional resources are allocated to the meaning of the speech, to the extraction of segmental and suprasegmental information (Dohen et al., 2009), and to the acquisition of new words, rather than to the gender of the face. Our results might therefore suggest that with IDS, the attention of younger infants is driven away from the gender of the talking face and towards the linguistic content of the speech. Further studies using, for example, eye-tracking would be necessary to verify this suggestion. Whatever findings such future work might yield, a potential implication of the difference in findings with ADS versus IDS is that younger infants may not be able to extract multisensory gender information when adults are directly talking to them (as is typical with IDS), and are more likely to do so when looking at adults engaged in conversation, using ADS with one another. Another implication would be that when addressed with IDS, younger infants may be able to extract multisensory gender information, but only when provided with additional facilitative cues such as articulatory patterns or body-related information. This latter implication is in accord with the study of Walker-Andrews et al. (1991), showing that from 6 months of age, infants are able to match faces and voices when presented with synchronous speech. Different from the Walker-Andrews et al. study, idiosyncratic cues or physical information such as gender-related clothing or make-up, were not available in our study.

Finally, Schachner and Hannon (2011) have shown that IDS conveys cues for social selection, guiding the preferences of 5-month-olds towards potential partners. After hearing an individual speak in IDS, infants attend more to that individual than to a novel person. This social preference is not found in the situation where an individual speaks in ADS (Schachner & Hannon, 2011). An adult using IDS transmits affective components that may relate to safety (Fernald & Kuhl, 1987; Papousek, Papousek, & Symmes, 1991). Interestingly, Kim and Johnson

(2013) recently reported that infant preference for infant-directedness is not only restricted to speech but also extends to ID faces and that this preference might be mediated by the emotion conveyed by the face (Kim & Johnson, 2014). Based on these findings, our results suggest that the attention of younger infants is directed to IDS cues related to learning, emotion, and safety, rather than to the gender of the talking face.

It is additionally worth noting that we did not observe any differences between IDS and ADS conditions for 9- and 12-month-old infants. Correspondingly, Newman and Hussain (2006) found a decrease in infant preference for IDS over ADS during the second half of the first year of life, with no preference observed at the age of 9 months. This latter outcome is in accord with the finding that only younger infants in the present study were affected by speech manner. It may be that IDS engages the attention of younger infants to the extent that it interferes with processing of other attributes of the interaction including the gender of the speaker's face and voice.

Like in previous studies (Hillaiet de Boisferon et al., 2015; Poulin-Dubois et al., 1998; Poulin-Dubois et al., 1994), our results reveal an asymmetrical responsiveness to female versus male faces and voices. Infants looked longer to female faces in the presence of a female voice, but not longer to male faces in the presence of a male voice. The most likely explanation for this result is that infants typically have predominant experience with female faces (e.g., Ramsey-Rennels & Langlois, 2006; Rennels & Davis, 2008; Sugden et al., 2014). Even though infants are from birth able to match individual voice characteristics with particular faces (Coulon, Guellai, & Streri, 2011; Guellai, Coulon, & Streri, 2011), and perceive multisensory coherence of visible and audible speech syllables from 2 to 4 months of age (Kuhl & Meltzoff, 1988; Patterson & Werker, 1999, 2002), our results provide further evidence that it is not until the middle of the first year of life that they acquire the necessary skills to perceive more complex attributes such as gender. It may be that increased experience is needed before multisensory coherence emerges for such attributes (Hillaiet de Boisferon et al., 2015; Lewkowicz et al., 2015; Poulin-Dubois et al., 1998). More extensive experience with women may lead to an earlier emergence of the category knowledge of female faces because infants are exposed to multiple exemplars of the relevant attributes for categorization, e.g., long hair, thin and soft curved eyebrows, rounded cheeks, longer eye-to-eyebrow distance, and high-pitched voice (Homa, Sterling, & Trepel, 1981). Complex categorization skills seem to be therefore highly dependent on the degree of perceptual experience, inclusive of daily life experiences with relevant exemplars of a category. The association between male faces and voices would similarly require learning of particular correlations, e.g., deeper voice with prominent Adam's apple (Rosch et al., 1976). These associations may be more difficult to acquire because infants typically spend less time with male as compared to female faces during the first year of life (Rennels & Davis,

2008). Poulin-Dubois et al. (1998) found that it was not until the age of 18 months that infants were able to perceive multisensory gender coherence for male faces and voices. Given that Poulin-Dubois et al. (1998) used static images in their study, it would be informative to clarify whether the perception of multisensory gender coherence for male faces and voices also emerges around 18 months of age with dynamic faces. Future research could also investigate whether infants raised with fathers as primary caregivers would exhibit the opposite pattern of asymmetrical responsiveness to male versus female faces and voices. This possibility is raised by the Quinn et al. (2002) findings showing that infants raised primarily by male caregivers attend more to male faces than to female ones, suggesting that daily experience with a particular gender affects face processing.

Finally, the finding that infants only matched the gender of female faces and voices does not seem to reveal an overall preference for female faces, but rather a true matching of audible and visible gender attributes. If there were an overall preference for female faces, 6-month-old infants should also have looked longer to female faces in the IDS condition. Interestingly, Liu et al. (2015) recently demonstrated a preference for female own-race faces in 3- and 6-month-old infants, but no such gender preference in 9-month-olds, suggesting a decrease in the female face preference in older infants. The results of Liu et al. (2015) were taken as evidence that by 9 months of age, male face experience has accumulated to the point where there is no longer a preference for female faces. That reasoning would suggest that 12-month-olds would also show no baseline preference, an observation that would need to be confirmed with additional evidence.

3.6 Conclusions

The present study used dynamic faces coupled with infant- or adult-directed speech to test multisensory gender perception in 6-, 9- and 12-month-old infants. The results revealed that by the middle of the first year of life, infants are able to perceive multisensory gender coherence for female faces and voices in the case of ADS. However, this ability emerges later with IDS. Altogether these results indicate that the ability to efficiently perceive gender in audiovisual speech is influenced by speech manner and emerges earlier for female faces. Our data imply that younger infants may be less likely to extract multisensory gender information when adults are directly talking to them in IDS than when looking at adults engaged in ADS with another. Overall, our findings imply that the circumstances of social interaction as well as predominant experience with female faces and voices, shape early multisensory abilities to perceive gender.

3.7 Acknowledgements

A special thanks goes to all the infants and their parents for their participation in this study, as well as to all the people who helped for testing and coding the videos. This research was supported by grant R01 HD-46526 from the National Institute of Child Health and Human Development as well as by grant F14/06 from the Rectors' Conference of Swiss Universities (CRUS).

CHAPTER 4

RECONSTRUCTING DYNAMIC MENTAL MODELS OF FACIAL EXPRESSIONS IN PROSOPAGNOSIA REVEALS DISTINCT REPRESENTATIONS FOR IDENTITY AND EXPRESSION

Richoz, A. R., Jack, R. E., Garrod, O. G., Schyns, P. G., & Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. *Cortex*, 65, 50-64.

Research highlights

- ◊ Reconstruction of dynamic mental models of facial expressions in prosopagnosia
- ◊ Contrary to face identity, the patient's mental models were normal for expression
- ◊ Categorization deficit in the patient for static but not for dynamic expressions
- ◊ Evidence of distinct representations/cortical pathways for face identity and expression
- ◊ Evidence of a direct and sufficient cortical pathway to process dynamic facial expressions in the pSTS

Keywords: Prosopagnosia, Facial expressions of emotion, Reverse correlation

4.1 Abstract

The human face transmits a wealth of signals that readily provide crucial information for social interactions, such as facial identity and emotional expression. Yet, a fundamental question remains unresolved: does the face information for identity and emotional expression categorization tap into common or distinct representational systems? In order to address this question, we tested PS, a *pure* case of acquired prosopagnosia with bilateral occipitotemporal lesions anatomically sparing the regions that are assumed to contribute to facial expression (de)coding (i.e., the amygdala, the insula, and the posterior superior temporal sulcus – pSTS). We previously demonstrated that PS does not use information from the eye region to identify faces but relies on the suboptimal mouth region. PS's abnormal information use for identity, coupled with her neural dissociation, provides a unique opportunity to probe the existence of a dichotomy in the face representational system. In order to reconstruct the mental models of the six basic facial expressions of emotion in PS and the age-matched healthy observers, we used a novel reverse correlation technique of tracking information use on dynamic faces. PS was comparable to controls, using all facial features to (de)code facial expressions. PS's normal (de)coding of dynamic facial expressions suggests that the face system relies either on distinct representational systems for identity and expression or on dissociable cortical pathways to access them. Interestingly, PS showed a selective impairment for categorizing many static facial expressions, which could be attributed to her lesion in the right inferior occipital gyrus. PS's advantage for dynamic facial expressions might instead relate to a functional and sufficient cortical pathway that directly connects the early visual cortex to the spared pSTS. Altogether, our data provide critical insights into the healthy and impaired face systems, question evidence of deficits obtained from patients using static images, and offer new routes for patient rehabilitation.

4.2 Introduction

The human face transmits a wealth of visual signals that are relevant for the identification and the categorization of facial expressions of emotion. The brain, as a decoder, flexibly filters the incoming visual information transmitted by a face in order to rapidly achieve complex perceptual categorizations (Schyns, Petro, & Smith, 2009). For example, the uniqueness of the facial features characterizing a given individual and their overall organization in the face constitute the core information for identification and also for the differentiation of familiar faces from unfamiliar ones. Other signals can also be extracted from faces, which include the cues disclosing age (e.g., George & Hole, 1995), gender (e.g., Brown & Perrett, 1993; Ekman & Friesen, 1976; 1978; Schyns, Bonnar, & Gosselin, 2002; Tranel, Damasio, & Damasio, 1988), race (Caldara & Abdi, 2006; Caldara et al., 2004; Vizioli et al., 2010a; Vizioli, Rousselet, & Caldara, 2010b), and emotional state (e.g., Bruce & Young, 1986; Calder & Young, 2005; Ekman & Friesen, 1976; 1978; Smith, Cottrell, Gosselin, & Schyns, 2005). Overt emotional states can also be extracted from face signals; they are mostly conveyed through facial expressions of emotion. The basic signals (i.e., “happiness”, “surprise”, “fear”, “disgust”, “anger”, and “sadness”) are only weakly correlated with one another in order to minimize confusions for their decoding (Smith et al., 2005), and we recently reported cross-cultural tunings in the way the emotion signals are transmitted and decoded (Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Jack, Caldara, & Schyns, 2012; Jack, Garrod, Yu, Caldara, & Schyns, 2012). Yet, a primary question remains unanswered: does the face information used to recover identity and emotional expressions tap into common or distinct representational systems?

According to the influential cognitive (Bruce & Young, 1986) and the neuroanatomical (Haxby, Hoffman, & Gobbini, 2000) models of face processing, two distinct functional and neural systems carry out the recognition of facial identity and facial expression. The first system – performing facial identification (Haxby et al., 2000) – has been proposed to mainly involve the inferior occipital gyri and the lateral fusiform gyrus, whereas the second system – performing facial expression categorization – has been proposed to involve the inferior occipital gyri, the posterior superior temporal sulcus (pSTS), and the amygdala (for a review, see Calder & Young, 2005; Pessoa, 2008). However, some authors have questioned the idea of independence among these systems by mainly relying on the results from computational modelling and neuroimaging evidence (Calder, 2011; Calder & Young, 2005). A single model based on a Principal Component Analysis (PCA) can achieve independent coding of facial identity and facial expression, which suggests the possible existence of a multidimensional system wherein the independence is more partial than absolute (Calder, Burton, Miller, Young, & Akamatsu, 2001). These simulations have thus challenged the view that an independence exists between the coding for identity and expression, or at least suggests that these models are

less strongly supported than what is often assumed (Calder & Young, 2005). Complementary to this position, Palermo and colleagues (2013) have recently put forward the idea of a first common step in the processing of expression and identity and the occurrence of a splitting at a later stage – a view that is in agreement with the functional involvement of the inferior occipital gyrus as the entry level for both tasks (Calder & Young, 2005; Haxby, Hoffman, & Gobbini, 2000; Pitcher, 2014). However, although a neural dissociation for the processing of identity and emotional expression is supported by electrophysiological studies in primates (e.g., Hasselmo, Rolls, & Baylis, 1989), functional neuroimaging in humans (e.g., Winston, Henson, Fine-Goulden, & Dolan, 2004) and brain-damaged patients (Haxby et al., 2000), recent evidence suggests that the neural computations occurring in the inferior occipital gyrus and the right pSTS are functionally distinct and have a causal involvement in the processing of facial expressions (Pitcher, 2014). To summarize, more evidence is necessary to clarify this debate and, as acknowledged by Calder and Young (2005), further studies with brain-damaged patients are necessary in order to probe the hypothesis of distinct visuo-perceptual systems for facial identity and facial expression categorization.

As a consequence of brain lesions, some patients lose the ability to detect facial identity, despite having no other obvious impairments of the visual system and a preserved identification through other modalities (e.g., voice, gait, and so forth). The specificity of this face recognition deficit is spectacular and rare and has elicited considerable attention within the neuropsychological literature since the first clinical observations (Quaglino, 1867; Wigan, 1844) and the introduction of the term “prosopagnosia” by Bodamer (1947). Acquired prosopagnosia typically follows brain damage to bilateral occipitotemporal areas (e.g., Damasio, Damasio, & Van Hoesen, 1982; Farah, 1990; Landis, Regard, Bliestle, & Kleihues, 1988; Sergent & Signoret, 1992). Anatomical descriptions of prosopagnosia endorse the necessary and sufficient role of the right hemisphere (Landis et al., 1988; Sergent & Signoret, 1992) in the occipitotemporal pathway of face processing (for a review, see Bouvier & Engel, 2004). The clinical and anatomical conditions of prosopagnosia have always received a considerable amount of interest in the domain of cognitive neuroscience, as they clarify the neurofunctional mechanisms of normal face processing. The different sub-functions of the cognitive architecture of face processing have been isolated by the occurrence of distinct double dissociations in brain-damaged patients, for instance, a functional segregation between the ability to recognize unfamiliar and familiar faces (e.g., Malone, Morris, Kay, & Levin, 1982) and between lip reading and face identification (Campbell, Landis, & Regard, 1986). Yet, the neuropsychological literature remains controversial on the spared ability of prosopagnosic patients to identify facial expressions despite their inability to recognize facial identity, and on patients showing impaired facial expression recognition with preserved facial identity recognition (for a detailed review, see Calder, 2011). Some acquired prosopagnosic patients displayed a marked impairment in the

categorization of facial expressions (Bowers, Bauer, Coslett, & Heilman, 1985; De Gelder, Pourtois, Vroomen, & Bachoud-Levi, 2000; De Renzi & Di Pellegrino, 1998; Humphreys, Donnelly, & Riddoch, 1993). Other studies reported preserved recognition of emotion in acquired prosopagnosia (Bruyer et al., 1983; Cole & Perez-Cruet, 1964; Mattson, Levin, & Grafman, 2000; Sergent & Villemure, 1989; Shuttleworth, Syring, & Allen, 1982; Tranel et al., 1988; Young, Newcombe, de Haan, Small, & Hay, 1993). In addition, as pointed out by Calder and Young (2005) and Calder (2011), the decoding of face identity and facial expressions of emotion activates a similar network of regions in the occipito-temporal cortex. Facial expression impairments in patients are often correlated with an inability to decode emotions from other modalities, which suggests a general, multi-modal deficit in these patients, instead of a selective impairment of facial expression representations. In addition, a better understanding of the patients' information use (i.e., representations) for both tasks is necessary in order to clearly understand the very nature of the deficits in the face processing system (Calder and Young 2005; Calder et al., 2011). Consequently, the question of dissociation between identity and expression systems in acquired cases of prosopagnosia remains unclear.

In order to address this issue, we tested PS – a *pure* case of acquired prosopagnosia. PS is a 67-year-old woman (born in 1950), who sustained a closed-head injury in 1992. PS shows normal object recognition (e.g., Busigny, Graf, Mayer, & Rossion, 2010; Rossion et al., 2003) and relies on atypical cues to determine the identity of a person, which include voice, clothes, or other salient non-face features (e.g., glasses, haircut, beard, posture). She has major lesions on the left mid-ventral and the right inferior occipital cortex. Minor lesions on the left posterior cerebellum and the right middle temporal gyrus were also detected (for a complete anatomical description, see Rossion, 2008; Sorger, Goebel, Schiltz, & Rossion, 2007), whereas the regions that are assumed to be critical for the decoding of emotional expressions (i.e., the amygdala, the insula, and the pSTS) are anatomically spared. Note that even if the occipital temporal regions do not play a central role for facial expression decoding, the right inferior occipital gyrus is damaged in PS and represents the entry level for expression and identity in posited neuroanatomical models (Haxby et al., 2000; Pitcher, 2014). Thus, it remains to be clarified whether these lesions in the patient also have an impact on the processing of facial expressions. Of interest, we previously used a response classification technique – Bubbles – to reveal the diagnostic information used by PS for face identification (Caldara et al., 2005). Bubbles is a response classification technique, which samples the information in a 3D space (2D image x spatial frequencies) to present sparse versions of the faces as stimuli (Gosselin & Schyns, 2001, for more details, see Box 6). Observers categorize the sparse stimuli, and Bubbles keep track of the samples of information that lead to the correct and the incorrect identification responses. From this information, we can establish the way in which each region of the input space contributes to face identification performance and depict the diagnostic information that is used

to effectively decode the stimulus. In contrast to healthy observers, PS did not use information from the eye region to identify familiar faces, but instead used the lower part of the face, including the mouth and the external contours. In summary, PS's well-established bias to use information from the mouth to identify faces and her anatomical neural dissociation provide a unique opportunity to probe the existence of a dichotomy in the representations used for facial identity and expression categorization.

Here, we have first assessed her categorization performance of the six facial expressions of emotion, using the classical Ekman and Friesen (1976) FACS (Facial Action Coding System) static face database. The FACS provides an anatomical taxonomy of the human muscles that are activated during the transmission of facial expressions of emotion (Ekman & Friesen, 1978), by quantifying facial movements for every expression in terms of so-called action units (AUs), each of which relate to a particular muscle. We then modelled PS's 3D dynamic mental representations of the six classic facial expressions by using a dynamic FACS-based generative face grammar (GFG) (see Figure 4.1, the methods section, and Yu, Garrod, & Schyns, 2012) on the AUs, combined with a reverse correlation technique (see the methods and also Jack et al., 2012). The use of dynamic facial expressions provides a more ecologically valid approach for studying the perception and processing of facial expressions, as our natural environment is surrounded by dynamic, temporal, and multimodal information (Johnston, Mayes, Hughes, & Young, 2013). Pertinently, it has also been recently demonstrated that there is a causal involvement of the right pSTS in the processing of dynamic facial information (Pitcher et al., 2014), a region anatomically spared in the case of PS.

The main goal of our study was to test whether the representations for identity and facial expressions are distinct or common. Mapping out the facial features used by a prosopagnosic patient to perform facial expression categorization is necessary to achieve this goal (Calder & Young, 2005). As PS shows a sub-optimal use of facial information for identity (i.e., by using the mouth region and the external contours), we put forward the hypothesis that if she adequately uses all facial features for expression categorization, this observation would support the hypothesis of distinct sets of representations for identity and emotion recognition.

4.3 Participants

PS's Case Report

PS is a 67-year-old case of acquired prosopagnosia with normal object recognition. Despite the multiple and extensive brain lesions in the occipitotemporal cortex, PS recovered well in the months following her accident and with support from neuropsychological rehabilitation, she restarted working as a kindergarten teacher. PS's low-level vision is well preserved with a good

visual acuity in both eyes, except for a small right paracentral scotoma. She reads normally (although slowly) and has normal object perception and recognition, even for subordinate-level discriminations (Rossion et al., 2003). However, as a consequence of the lesions, she reports that her face recognition is severely impaired, even for close relatives (husband, children, mother, father) and herself. PS can categorize a face as a face and discriminate faces from objects and from a complex scene background, even at brief presentations (Schiltz et al., 2006). Her ability to categorize gender is impaired, in terms of both accuracy and sensitivity (Rossion et al., 2003). Her performance in categorizing facial expressions is not as good as those of controls. In a 3AFC expression categorization task where the stimuli were presented for a maximum of 10 seconds, PS's reaction times were significantly slower than those of the controls and her response accuracy was below the range for the first block of expressions comprising happiness, fear, and anger. However, her performance was collapsed across the expressions presented within a block and she was normal for the second block that tested disgust, sadness, and surprise (Rossion et al., 2003). Therefore, we had to assess PS's categorization accuracy of static facial expressions properly through the use of a 7AFC task (Experiment 1). PS is able to correctly draw a schematic face and perfectly point out all the single features and estimate age in the normal range. This latter result contrasts her inability to recognize previously seen or familiar faces and to match unfamiliar faces, including changes in viewpoints (Rossion et al., 2003). As reported by Rossion and colleagues (2003), she is ranked as highly impaired on the Benton Face Matching Test (Benton & Van Allen, 1972) scoring 27/54 (percentile 1). She is also impaired on the Short Recognition Memory Test for Faces, a set of the Camden Memory Tests (Warrington, 1984), scoring 18/25 (percentile 3). When confronted with the pictures of 60 famous people (all known by the patient), she was able to classify 14 of them as familiar and correctly classified all the unfamiliar ones (Rossion et al., 2003). Nevertheless, when asked to report the individual names of the faces classified as familiar, as well as their semantic information, she was correct regarding only four of them. Finally, PS has been tested extensively with simultaneous and delayed face and non-face (cars and novel objects) matching tasks in previous studies (Rossion et al., 2003). Although she is consistently impaired and slowed down for the face conditions, her performance with the non-face objects is in the normal range. Given the restriction of her deficit to the face category and the fact that PS is alert, cooperative, and without any learning difficulties (Caldara et al., 2005), she represents an ideal case to isolate the nature of the facial information extracted by an impaired face system for the processing of facial expressions.

Control participants

The age-matched healthy observers who voluntarily took part in our experiments had normal or corrected to normal vision with no neurological or psychiatric history. The number of healthy observers and their age is reported in the method section of each experiment. For all our

experiments, PS and the control participants signed a consent form describing the main goals of our experiments. The Ethical Committee of the Department of Psychology of the University of Fribourg approved all the studies reported here.

4.4 Experiments

4.4.1 Preliminary experiment: Categorization of the six classic facial expressions of emotion

We first assessed the categorization accuracy of static facial expressions using a standard set of posed facial expression stimuli – the Pictures of Facial Affect series (POFA) (Ekman & Friesen, 1976; 1978).

4.4.1.1 Material and methods

Control participants

Twelve age-matched healthy control subjects (eight females) participated in the experiment ($M_{\text{age}} = 59.41$; $SD = 3.98$).

Stimuli

We selected seven posed facial expression images from 22 identities (11 females) – one per emotion category (happiness, surprise, fear, disgust, anger, sadness, and neutral) – from the Pictures of Facial Affect (POFA) database (Ekman & Friesen, 1976; 1978). The pictures were in grayscale and were not cropped.

Procedure

Observers categorized each stimulus according to one of seven possibilities (happiness, surprise, fear, disgust, anger, sadness, and neutral) in a 7AFC task using a computer keyboard in which we labelled the keys accordingly. Each image was presented for 2000 ms in random order in the center of the observers' visual field with 30 repetitions of each expression, resulting in a total of 210 trials. Faces subtended a visual angle of 9.54° (vertical) and 8.11° (horizontal) on the screen.

4.4.1.2 Results and discussions

In order to determine whether PS's average accuracy is significantly different from that of the age-matched healthy participants, we used a modified independent samples t-test for single case studies (Crawford & Garthwaite, 2002). The statistical level of significance is $p < 0.05$. As shown in Table 4.1, PS's categorization of static, posed, FACS-coded facial expressions is significantly impaired for anger ($t(11) = -4.38$, $p < 0.05$), fear ($t(11) = -2.41$, $p < 0.05$), surprise ($t(11) = -2.69$, $p < 0.05$), and sadness ($t(11) = -3.27$, $p < 0.05$), compared to the age-matched healthy controls.

Table 4.1. Recognition accuracy of PS and age-matched controls on Ekman and Friesen’s facial expression recognition test

PS showed impaired recognition of anger, fear, surprise, and sadness, compared to healthy controls.

	PS		Age-matched controls	
	Score %	Modified <i>t</i> -test	Mean %	SD %
<i>Ekman and Friesen (1978)</i>				
Anger	36.33	-4.38*	71.95	8.46
Disgust	88.55	1.53	79.45	6.16
Happiness	97.78	0.55	96.11	3.12
Neutral	79.17	1.47	72.40	4.77
Fear	45.67	-2.41*	67.22	9.30
Surprise	48.89	-2.69*	70.00	8.16
Sadness	47.22	-3.27*	70.00	7.24

* $p < 0.05$

4.4.2 Experiment 1: Reconstructing PS’s dynamic mental models of facial expressions of emotion

Using a novel method, we modelled the 3D dynamic mental representations of the six facial expressions plus the neutral of PS and control participants, by using the FACS-based generative face grammar (GFG) computer graphics platform and a reverse correlation technique.

4.4.2.1 Material and methods

Control participants

Five healthy age-matched controls (three women) participated in the experiment ($M_{\text{age}} = 60.2$; $SD = 3.27$).

Stimuli

Using the GFG (Jack, Garrod, et al., 2012; Yu et al., 2012), we synthesized a series of photorealistic facial animations by randomly selecting a subset of groups of muscles that move in synergy on the face – action units (AUs) (Ekman & Friesen, 1976; 1978). Practically, they were measured and modelled in 4D (3D face plus time) from the mapping of actors trained by Ekman to produce them. Existing as mathematical models in the GFG, we can animate each AU over time independently, using six temporal parameters (onset, acceleration, peak amplitude, peak latency, deceleration, offset) (see Figure 4.1 and Movie 1 for an example stimulus).

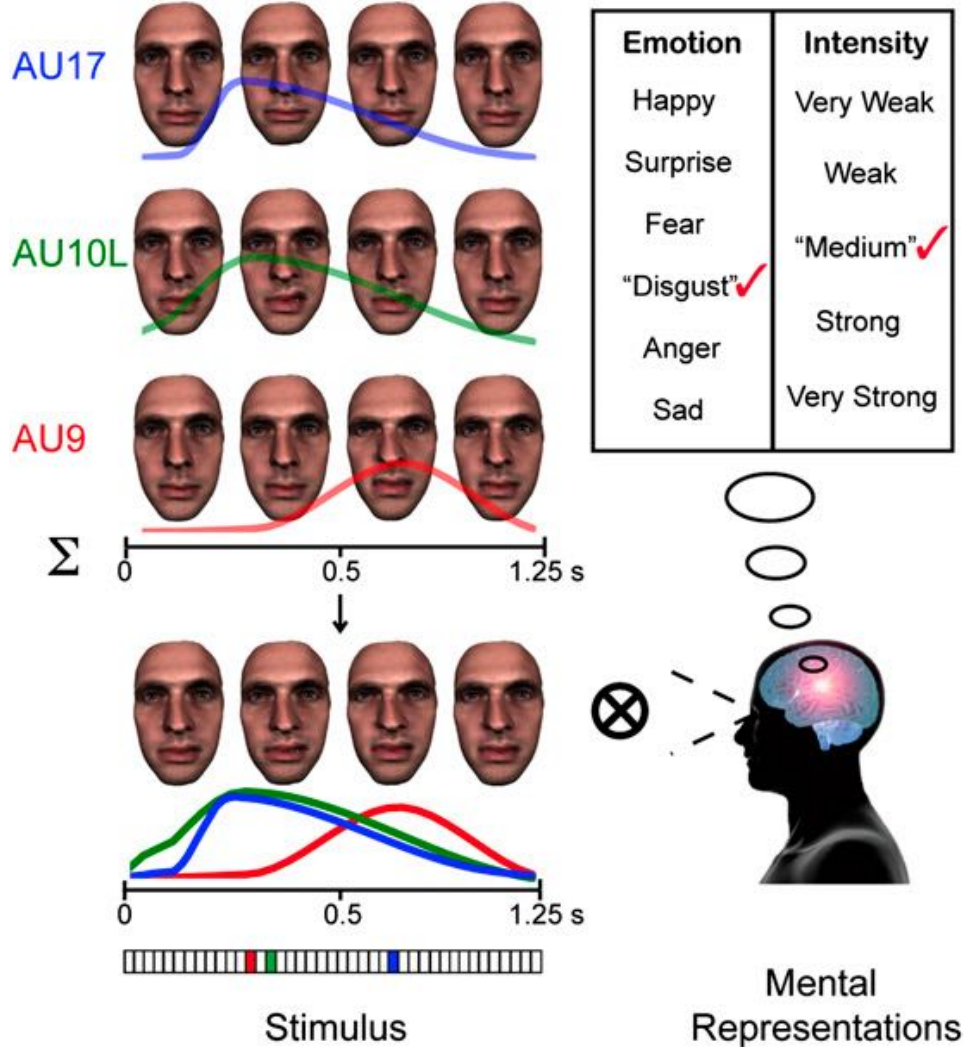


Figure 4.1. FACS-based generator of random facial muscle activations. *Left:* Stimulus generation – On each trial, the generative face grammar (GFG) randomly selected a subset of action units (AUs; AU17, AU10L, and AU9 are shown here with color-coded labels) and values for six temporal parameters (see the color-coded AU curves that illustrate the amplitude and acceleration or deceleration of movement over time). The color-coded vector at the bottom of the figure represents the three (of 41) randomly selected AUs that make up the stimulus on this illustrative experimental trial. We then applied the random facial animation to one of eight neutral-expression face identities using the procedure described in Yu, Garrod, and Schyns (2012). *Right:* Mental representations – Observers categorized each random facial animation according to the six basic emotion categories (plus “don’t know”) and rated the emotional intensity on a five-point scale. Observers will interpret the random facial animation as a meaningful facial expression (here, “disgust,” “medium intensity”) when the facial movements correspond to the observer’s mental representation of that facial expression.

Each animation is displayed on one of eight unfamiliar 3D photorealistic white Caucasian face identities (four females), which were acquired with a 3D photorealistic capture system (Dimensional Imaging). In any given trial, the GFG selects a subsample of AUs amongst the 41 core AUs using a binomial distribution ($n = 5$, $p = 0.6$, median = 3). The mental representation of each facial expression was then modelled for each observer by reverse correlating the random AUs and their temporal parameters with the observer's emotion responses. This technique has been validated in previous studies (Gill, Garrod, Jack, & Schyns, 2014; Caldara et al., 2012; Garrod et al., 2012).

Procedure

The experiment consisted of 12 sessions of four blocks that ran on a computer with the use of a code written with Adobe Flash. Each block included 50 trials, consisting of 4D facial animations (3D + time), displaying a random subset of AUs movements. All the observers categorized a total number of 2400 of such animations, comprising 30 frames (24 frames/second) of 1.25 seconds duration. The facial animations subtended approximately 9.54° (vertically) and 6.68° (horizontally) of visual angle. For each trial, observers viewed the facial animation and categorized it according to the six basic emotions: happiness, surprise, fear, disgust, anger, and sadness and a "don't know" response. Furthermore, observers rated the intensity of the perceived emotion on a five-point rating scale ("Very weak" to "Very strong"). We adapted the response to our senior population using the Geneva Emotion Wheel (GEW), which was introduced by Scherer (2005). We labelled the emotion categories in the outer boundaries of a circle, with the "don't know" response option in the center. Additionally, we used five circles gradually increasing from the center towards the respective emotional category placed on the border in order to allow participants to report the intensity of the perceived facial expression. Observers navigated with a mouse to select the basic emotion and its intensity. We did not provide feedback and did not place any time pressure on participants (including PS) to respond. All of the observers were familiar with using a computer and did not have difficulty with this interface. The whole experiment lasted for about six hours, over a period of six weeks.

4.4.2.2 Results

Representation of the six basic emotions. For each participant, we modelled the mental representations of the six basic emotions, by reverse correlating the Action Units (AUs) randomly selected for each trial with the responses of the observers. For each observer, this resulted in a 41-dimensional (one dimension per AU) ON/OFF vector, with ON AUs being significantly correlated with an emotion category. The p -value of 0.05 was used to determine whether the correlation coefficient between each AU's presence or absence and the participant's emotional response was significantly larger than zero. For the final models, only those AUs whose correlation coefficient was significantly larger than zero were displayed, but the regression

coefficients for each AU were not affected by whether or not other AUs were deemed significantly correlated or not. We then regressed for each ON AU, the six temporal parameters with the intensity responses of observers in order to model the activation dynamics (Jack et al., 2012; Garrod et al., 2012; Yu et al., 2012). Figures 4.2 and 4.3 report the results. Figure 4.2 illustrates the static version of the dynamic stimuli, computed by collapsing the highest amplitude of each AU involved in the models of the “Very Strong” intensity judgment. Figure 4.3 represents a static rendering of the facial parts that move (with the amplitude of movement represented in millimetres).

Figures 4.2 and 4.3 indicate that the models of PS are within the range of those of controls. That is, for each individual expression, at least one control is very similar to that of PS (e.g. as Figure 4.3 reveals, PS’s “happiness” is similar to all controls’ “surprise” to control 5, “fear” to 5 again, “disgust” to at least control 3, “anger” to control 4, and “sadness” to control number 3). Movie 2 presents the 4D dynamic version of the mental models of PS and Movies 3, 4, 5, 6, 7, those of the five control subjects. In order to test whether the patients’ AUs preferences for each emotion fell within the normal range of the population, we computed Pearson correlations between each of the patient’s AUs ON/OFF vectors and the corresponding vectors for the control group (Table 4.2, column 1).

Table 4.2. PS’s AUs preferences fell within the normal range of controls. For each expression, Pearson correlations fell between PS’s ON/OFF AU vectors and age-matched controls (column 1). Permutations of one control with other controls produced the expected correlations (column 2). Z-score difference between patient-controls correlations and the expected correlations (column 3). None of those differences was significant.

	PS vs. Age-matched controls, Pearson Correlations		
	PS vs. Controls	Expected Correlation	Z-scored difference
<i>Static reconstructed model</i>			
Anger	0.3028	0.2020	0.9096
Disgust	0.4131	0.3585	0.3630
Happiness	0.5027	0.5679	0.8744
Fear	0.0175	0.0745	0.4265
Surprise	0.4185	0.3638	0.5040
Sadness	0.1040	0.1053	-0.0096

A permutation test confirmed that the 95% confidence interval of the expected correlation (between any control and the remaining controls, Table 4.2, column 2) for the null hypothesis (Patient = Controls) contained the correlation between the patient and the controls for each of the six emotions. Table 4.2, column 3 contains the difference between the patient-

controls correlations and the mean of the permutation test correlations expressed as Z-scores. PS fell in the normal range for the AU preference.

4.4.2.3 Discussion

The reconstructed mental models of PS reveal that her representations of facial expressions are comparable to those of the age-matched controls. The fact that PS uses information from the eye region to represent basic emotions, as demonstrated by the analysis on the deviation maps, is of interest. These results conflict with her inability to accurately categorize most of the static, posed facial expressions of the Ekman and Friesen (1978) face database (i.e., anger, fear, surprise, and sad). How can we reconcile such a discrepancy?



Figure 4.2. PS's and controls' reconstructed mental models of happiness, surprise, fear, disgust, anger, and sadness. Positive correlation coefficients between active Action Units and emotional responses are represented as AU intensity. Active AUs for a given observer and emotion correspond to those whose correlation coefficient is significantly greater than zero ($p < 0.05$, one-tailed). *Rows:* Observer (top row PS, bottom rows controls). *Columns:* Emotion (happiness, surprise, fear, disgust, anger, sadness).

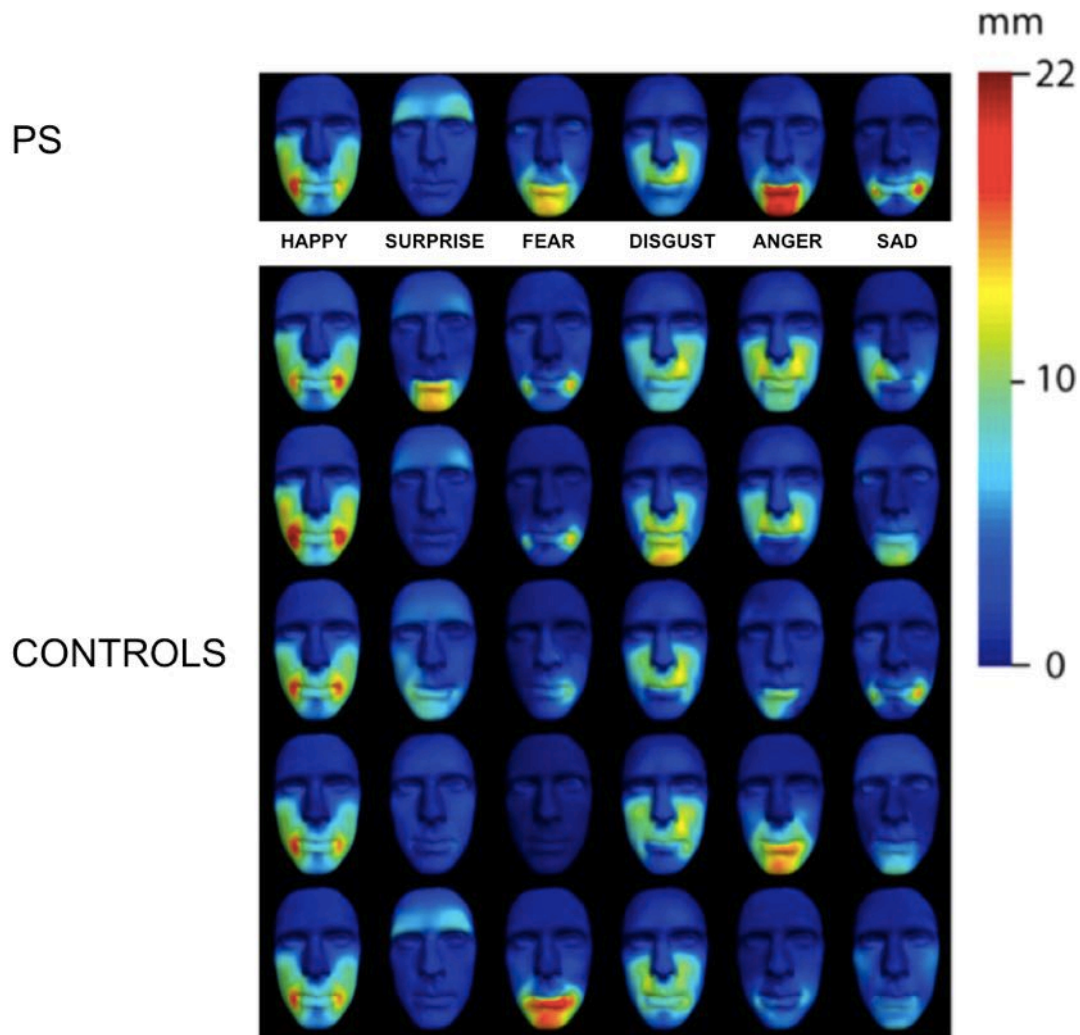


Figure 4.3. Deviation maps of PS's and controls' reconstructed mental models of happiness, surprise, fear, disgust, anger, and sadness. Colour scale represents peak magnitude of facial movement relative to neutral over the course of the expression at the corresponding position on the face. Facial movements are the combinations of active AUs weighted by positive Pearson correlation coefficients (negative correlations are not shown) with emotional responses. Active AUs for a given observer and emotion correspond to those whose correlation coefficient is significantly greater than zero ($p < 0.05$, one-tailed). Rows: Observer (top row PS, bottom rows controls). Columns: Emotion (happiness, surprise, fear, disgust, anger, sadness).

4.4.3 Experiment 2: Assessing the categorization of facial expressions of emotion with PS's static and dynamic reconstructed mental models

In order to clarify our previous observations, we conducted a verification task in which we presented PS and a new group of age-matched controls with static and dynamic mental models of the patient and asked all observers to categorize the facial expressions of these stimuli.

4.4.3.1 Material and methods

Control participants

A new group of 10 age-matched controls (six women) and PS participated in the experiment ($M_{\text{age}} = 58.4$; $SD = 4.19$).

Stimuli

Static reconstructed models. This experiment comprised static and dynamic versions of the same facial expressions. That is, for each dynamic 4D model (3D + time), we created a static version by collapsing time, keeping only the highest amplitude of each AU involved in the model (see Figure 4.2 and Movies 2 & 3 for example). We used PS's mental models of each of the basic emotions (happiness, surprise, fear, disgust, anger, and sadness) at three levels of intensity (low, medium, and high). The judgment of intensity was only kept to maintain attention and provide observers with a fine-grained scale, but its analysis is beyond the scope of the present work. This resulted in 18 facial expressions, presented six times, for a given total of 108 images. The faces were subtending the same visual angles as in the previous experiment and were presented for 1250 ms.

Dynamic reconstructed models. Each dynamic facial expression consisted of 30 frames (24 frames/second) and lasted for 1250 ms. The facial animations and their static versions were covering the same visual angle of the previous experiment.

Procedure

Static reconstructed models. The stimuli were presented in a random order, one at a time, on the computer screen. PS and the controls were instructed to categorize each stimulus as accurately as possible by pressing one of the six labelled keys on the computer keyboard (one per expression). No feedback was provided.

Dynamic reconstructed models. We used 12 animations for each emotion (four per intensity) for a total of 72 trials. These animations were taken from the same identities used for the static presentation. After each presentation, observers categorized the dynamic facial expression according to the six classic emotions (i.e., happiness, surprise, fear, disgust, anger, and sadness),

as accurately as possible by pressing one of the six labelled keyboard keys (one per expression). No feedback was provided.

4.4.3.2 Results and discussion

Static and dynamic reconstructed models. PS could not categorize the static version of many of her own mental models of facial expressions (Figure 4.4).

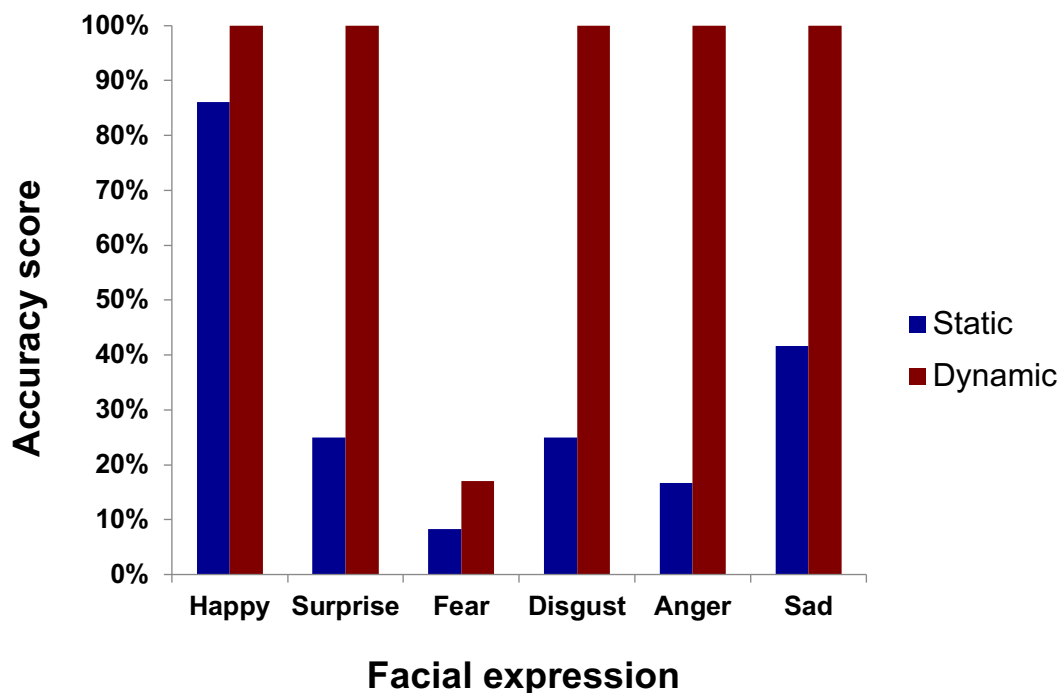


Figure 4.4. Assessing categorization of facial expressions of emotion with PS's reconstructed mental models. The figure shows PS's percentage accuracy score for the recognition of happiness, surprise, fear, disgust, anger, and sadness, from her static and dynamic reconstructed mental models. With the exception of fear, PS showed the maximum accuracy score of 100% for the recognition of all the dynamic emotions.

As shown in Table 4.3, PS was significantly impaired in the recognition of the static images of her own mental models of anger ($t(9) = -2.04$, $p < 0.05$), disgust ($t(9) = -3.008$, $p < 0.05$) and surprise ($t(9) = -2.57$, $p < 0.05$). In contrast, the controls categorized the static models of PS successfully, with the exception of "fear". In addition, PS categorized her own dynamic mental models (with the exception of fear) without difficulty. She was perfectly accurate for "anger, disgust, happiness, surprise, and sadness". A one-way repeated measures analysis of variance (ANOVA) was performed in order to examine whether dynamic presentations facilitated

performance for controls in the same way as PS. There was no significant effect of display presentation (static vs. dynamic) ($F_{(1,18)} = 0.74$, $p = 0.4$) and no significant interaction of emotion x display presentation ($F_{(1,18)} = 0.14$, $p = 0.98$).

Table 4.3. Percentage accuracy score of PS and age-matched controls for PS's static and dynamic reconstructed mental models. The scores are given for the recognition of anger, disgust, happiness, fear, surprise, and sadness (Experiment 2).

	PS		Age-matched controls	
	Score %	Modified t-test	Mean %	SD %
<i>Static reconstructed model</i>				
Anger	16.67	-2.04*	80.03	32.19
Disgust	25.00	-3.008*	90.00	22.49
Happiness	86.11	-1.04	96.67	10.54
Fear	8.33	-0.05	10.00	31.62
Surprise	25.00	-2.57*	83.33	23.57
Sadness	41.67	-0.80	73.33	40.98
<i>Dynamic reconstructed model</i>				
Anger	100	0.56	81.67	33.75
Disgust	100	0.49	93.33	14.05
Happiness	100	0.33	96.67	10.54
Fear	17	-	0	0
Surprise	100	0.76	74.97	34.52
Sadness	100	0.85	68.27	38.95

* $p < 0.05$

4.4.4 Experiment 3: Assessing categorization of facial expressions of average static and dynamic reconstructed mental models

In order to verify whether PS's ability to categorize dynamic emotional expressions is restricted to the recognition of the expressions reconstructed from her mental model, we assessed her ability to categorize average static and dynamic reconstructed mental models. Given the limited number of mental models from the healthy controls in the present experiment, we reconstructed an average mental model by using the data of 30 Westerner observers of our previous study (Jack et al., 2012). In line with Experiment 2, this comprised the categorization of static and dynamic versions of the same facial expressions.

4.4.4.1 Material and methods

Control participants

PS and 12 age-matched control subjects (seven women) took part in this experiment ($M_{\text{age}} = 59$; $SD = 3.71$).

Stimuli

Average static and dynamic 4D mental models. We used an average of the mental models of 30 Westerners for each of the basic emotions (Jack et al., 2012), all presenting the highest level of intensity. We then rendered these models on 10 Caucasian actors (five females). The faces randomly expressed the six basic emotional expressions of fear, happiness, anger, disgust, sadness, and surprise. As done in the previous experiment, we created a static version of the average models by collapsing the time factor and retaining only the highest amplitude of each AU involved in the model (Figure 4.5). All observers were required to categorize both the static and dynamic facial expressions of the stimuli. We then randomly sampled all the mental models of each of the basic emotions 10 times on each face, which resulted in a total of 60 trials in both static and dynamic conditions. The faces subtended the same visual angle as seen in the previous experiment, and each stimulus was presented for 1250 ms in random order at the centre of the screen.

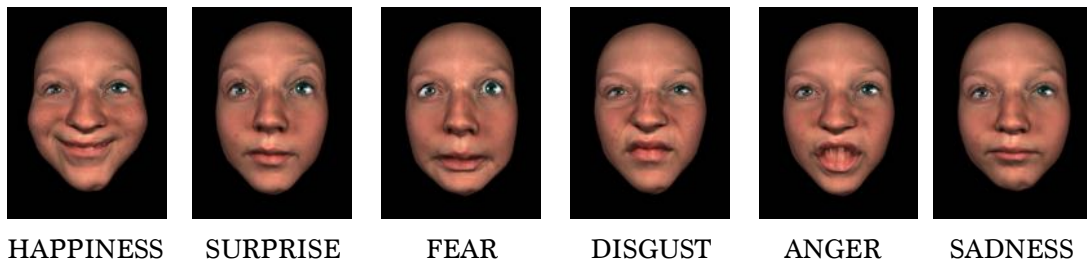


Figure 4.5. Average *static* reconstructed mental model for the expressions of happiness, surprise, fear, disgust, anger, and sadness. The *static* version of the average models was created by collapsing the time factor and keeping only the highest amplitude of each AU involved in the model.

Procedure

We followed the same procedure as the one in Experiment 2.

4.4.4.2 Results and discussion

PS was impaired in categorizing the static emotions of fear ($t(11) = -2.15, p < 0.05$), as well as sadness ($t(11) = -2.70, p < 0.05$) from the average mental models, compared to the age-matched healthy control observers (see Table 4.4). However, PS was accurate in categorizing all the dynamic emotions. Her performance for decoding fear from the dynamic average of healthy mental models was preserved.

For the controls, a repeated-measures ANOVA yielded a statistically significant main effect of the presentation (static vs. dynamic) ($F_{(1, 22)} = 9.48, p < 0.05$), reflecting that the average

performance with dynamic presentations was higher ($M = 83.61\%$, $SD = 9.55\%$) than with static ones ($M = 71.25\%$, $SD = 19.94\%$). A highly significant emotion \times presentation interaction ($F_{(1, 22)} = 7.45$, $p < 0.001$), indicated that the accuracy scores for dynamic presentations were higher for some emotions but not for others. Post-hoc t-tests revealed that this effect was driven only by “surprise”, which was significantly higher with the dynamic presentation ($M = 89.17\%$, $SD = 13.79\%$) than with its static version ($M = 39.17\%$, $SD = 17.30\%$) ($t(11) = 7.54$, $p < 0.001$).

Table 4.4. Percentage accuracy score of PS and age-matched controls for average static and dynamic reconstructed mental models

The scores are given for the recognition of anger, disgust, happiness, fear, surprise, and sadness.

	PS		Age-matched controls	
	Score %	Modified t-test	Mean %	SD %
<i>Static reconstructed models</i>				
Anger	53.33	-0.20	60.00	-0.20
Disgust	76.67	-0.23	80.83	-0.23
Happiness	90.00	-1.72	97.50	-1.72
Fear	33.33	-2.15*	71.67	-2.15
Surprise	40.00	-0.05	39.17	0.05
Sadness	56.57	-2.70*	78.33	-2.70
<i>Dynamic reconstructed models</i>				
Anger	90	0.40	79.17	25.39
Disgust	100	1.21	73.33	21.03
Happiness	100	0.41	98.33	3.89
Fear	90	0.60	75.00	23.93
Surprise	100	0.75	89.17	13.79
Sadness	100	0.66	86.66	19.23

* $p < 0.05$

4.5 General discussion

Influential theoretical (Bruce & Young, 1986) and neuroanatomical (Haxby et al., 2000) models of face processing have suggested the existence of distinct cortical pathways for face identification and expression categorization, a view that has also been challenged (Calder & Young, 2005). The main goal of the present study was to clarify whether the face information of identity and emotional expression categorization tap into common or distinct representational systems. To tackle this issue, brain-damaged patients can be very informative, as from their lesions, their specific behavioural impairments, and information use it is possible to infer the critical role played by the damaged regions in the healthy operating system. We first assessed the ability of PS – a *pure* case of acquired prosopagnosia – to categorize static facial expressions.

PS was impaired for many facial expressions of the Ekman and Friesen's faces. Second, we modelled her mental representations of happiness, surprise, fear, disgust, anger, and sadness. Our overarching goal was thus to map out the facial features used by the patient in order to achieve this biologically relevant task, considering the fact that we previously reported a suboptimal use of the eye region for face identification in the same patient (Caldara et al., 2005). Surprisingly, her dynamic mental models of facial expressions revealed an appropriate use of all facial features and Action Units (AUs – facial muscles), for all facial expressions. Since PS does not merely use the mouth during this task, these results suggest that the mental representations of facial expressions are anatomically separate, or that they are common but can be accessed from a distinct (cortical) route from face identification. Regardless of either potential explanation, our data support the view of a discrete (neural) coding for expression and identity.

The normal representations of dynamic facial expressions in PS, coupled with the impairment in the categorization of many of the static Ekman faces, came as a surprise to us and raised unexpected novel questions. How could PS have appropriate dynamic models of facial expressions and not be able to recognize them when presented in the form of static images? We thus designed subsequent verification tasks in which the patient was asked to categorize her and average static and dynamic reconstructed mental models of facial expressions of emotion. These experiments revealed a clear-cut dissociation. PS was selectively impaired in the categorization of many facial expressions of her and the average static reconstructed models. In stark contrast, she excelled with the very same dynamic models, a pattern of results that can be explained by her particular set of lesions. Moreover, in general, this result points to the use of dynamic stimuli as being critical in the assessment of facial expression recognition with brain-damaged patients and questions the sole use of static face images towards this aim. We now discuss, in turn, each of these findings and their implications.

4.5.1 Does the face system rely on common or distinct representations for identity and expression?

In order to investigate whether facial identity and expression decoding is tapping into common or distinct representational systems, we mapped out the different facial features/AUs used by a *pure* case of acquired prosopagnosia for categorizing facial expressions. PS displayed abnormal performance in categorizing most of the classic facial expressions (i.e., anger, fear, surprise, and sadness) of the well-known Ekman and Friesen database (Ekman & Friesen, 1976; 1978), whereas the age-matched control group displayed a normal range performance (Calder et al., 2003). The analysis of the reconstructed models clearly revealed that PS is comparable with the age-matched control observers and uses similar facial muscles (i.e., AUs) and similar temporal

dynamics in order to represent the six classic emotions. Thus, contrary to our previous observations for face recognition (Caldara et al., 2005), PS used all facial features/AUs to reconstruct facial expressions, favouring the view of distinct representations for expression and identity¹. In line with these positions, neuroimaging studies have revealed that the fusiform gyrus is sensitive both to facial identity and facial expression information (e.g., Cohen Kadosh, Henson, Cohen Kadosh, Johnson, & Dick, 2010; Fox, Moon, Iaria, & Barton, 2009; Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001). Calder (2011) suggested that the fusiform gyrus contributes primarily to the analysis of the visual form of faces for both the tasks, whereas the pSTS is involved in the coding of the changeable aspects of faces (Haxby, Hoffman, & Gobbini, 2000). PS thus relies on the fusiform gyrus during face identification and the pSTS during facial expression recognition. This explanation may be partially accounted by a previous neuroimaging study that used a neural adaptation paradigm with the patient. Schiltz and colleagues (2006) observed that contrary to normal observers, the neural activations of PS's right fusiform gyrus could not discriminate between the repetition of identical and different faces. This result suggests that the fine-grained visual form analysis necessary to effectively perform face identification requires the integrity of the face network and/or an optimal use of all the facial features. In contrast, the pSTS is anatomically and functionally spared in PS (Sorger et al., 2007) and could account for her normal facial information use for facial expression categorization. Despite favouring this view, we cannot completely and firmly rule out an alternative scenario in which the face system has a common set of representations, and the access may occur from distinct cortical pathways – one for face identification (through the inferior occipital gyrus) and one for expressions (through the pSTS). Importantly, both explanations suggest the existence of a discrete coding for expression and identity, and a future neuroimaging study is necessary to provide a clear-cut picture of those scenarios and precisely isolate the brain regions dedicated to the decoding of static and dynamic faces in PS, for both face identity and expression categorization.

On a general note, the use of dynamic reverse correlation techniques represents a unique tool for the understanding of patient impairments and their rehabilitation. For instance, a growing number of studies have indicated that schizophrenic patients are impaired in the recognition of the six basic facial emotional expressions (e.g., Kohler, Bilker, Hagoort, Gur, & Gur, 2000; Kohler et al., 2003; Kohler, Turner, Gur, & Gur, 2004; Martin, Baudouin, Tiberghien, & Franck, 2005; Sachs, Steger-Wuchse, Kryspin-Exner, Gur, & Katschnig, 2004). A selective impairment in recognizing fear and disgust has also been shown for unaffected relatives (Mendoza et al., 2011). Revealing the precise (defective) facial information use in these

¹According to this scenario, it could be possible that there are shared mechanisms represented in a multiple node network, with some nodes being responsible for the processing of identity and others for the recognition of facial expressions. Some of the nodes related to face identification would be severely damaged in PS.

populations might help in tailoring rehabilitation-training programs. Gaze-contingent eye tracking paradigms (Caldara, Zhou, & Mielliet, 2010; Mielliet, Caldara, & Schyns, 2011; Mielliet et al., 2012) will be well suited for this purpose, as they can force the viewing of a particular diagnostic features during a categorization task (see also Adolphs et al., 2005), which hopefully improves performance.

4.5.2 Advantage in decoding dynamic facial expression in prosopagnosia

Our study clearly points to an advantage in decoding dynamic versus static images for the categorization of facial expressions in PS only. Dynamic facial expressions provide observers with richer, more unique, and ecologically-valid representations, which should facilitate their processing (e.g., Giard & Peronnet, 1999; Johnston, Mayes, Hughes, & Young, 2013; LaBar, Crupain, Voyvodic, & McCarthy, 2003; Paulmann, Jessen, & Kotz, 2009; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Trautmann, Fehr, & Herrmann, 2009). However, the advantage in the processing of the dynamic over the static emotional stimuli in healthy observers is controversial (for a review, see Fiorentini & Viviani, 2011). Some studies have showed an advantage (e.g., Ambadar, Schooler, & Cohn, 2005; Cunningham & Wallraven, 2009; Giard & Peronnet, 1999; Knappmeyer, Thornton, & Bulthoff, 2003; Paulmann et al., 2009; Wehrle, Kaiser, Schmidt, & Scherer, 2000), whereas others have revealed that the benefits of dynamic properties in processing facial information may be minimal (e.g., Gold et al., 2013) or inexistent (e.g., Fiorentini & Viviani, 2011).

Importantly, the dynamic advantage that we observed in PS could be attributed to her specific set of lesions. It has been recently noted that the right inferior occipital gyrus – anatomically damaged in PS – is causally engaged in the processing of static face images (Pitcher et al., 2014) and expressions (Pitcher, 2014). On the contrary, the pSTS – anatomically spared in PS – has been related to the processing of dynamic faces (Pitcher et al., 2014) and expressions (Pitcher, 2014). Consequently, our data suggest that the inferior occipital gyrus plays a critical role in the decoding of static facial expressions (as the patient shows an impairment), but the pSTS is sufficient for effectively decoding facial expressions from dynamic visual inputs (for which the patient shows a normal performance). Greater responses for dynamic compared to static facial expressions have been found in the fusiform gyrus and the pSTS in the right hemisphere (e.g., Johnston et al., 2013; Kessler et al., 2011; LaBar et al., 2003; Sato et al., 2004; Schultz & Pilz, 2009), as well as the visual motion area in the occipito-temporal regions (more commonly referred to as V5/MT) (Johnston et al., 2013). Our observations also reinforce the existence of a direct and functionally distinct cortical pathway connecting the early visual cortex to the pSTS, which would not require structural information from the right inferior occipital gyrus (since this region is damaged in PS) in order to decode expressions effectively. This

advantage for directly processing dynamic visual inputs seems to be specific to facial expressions, as the patient cannot recover identity through dynamic visual information in everyday life. However, this hypothesis remains to be clarified with the help of a future functional neuroimaging study that measures the extent to which the activations observed in the pSTS in PS (Sorger et al., 2007) would be significantly modulated by the presentation of dynamic as well as static facial expressions.

Our results also feed the debate of the dynamic vs. static advantage with new data, by suggesting that dynamic information may give an advantage to patients only. In fact, our results are in line with those found by Humphreys, Donnelly, and Riddoch (1993) with an agnosic patient. This patient was impaired at discriminating facial identity and facial expressions when exposed to static images, whereas he performed normally when asked to judge a subset of facial expressions (i.e., smiling, frowning, or surprise) and gender from dynamic faces animated by light dots. However, whether this agnosic patient could correctly categorize all the basic facial expressions and his fine-grained information use was not assessed. Despite these methodological and theoretical differences, the observation that emerges from both the studies is that the recognition of emotions seems to imply a complex mechanism, which is facilitated by dynamic information. In fact, patients suffering from depression or schizophrenia also benefit from dynamic presentation of facial emotional expressions (e.g., Atkinson, Dittrich, Gemmell, & Young, 2004; Kan, Mimura, Kamijima, & Kawamura, 2004; Schaefer, Baumann, Rich, Luckenbaugh, & Zarate, 2010). Dynamic stimuli enclose information that cannot be completely rendered by static representations and force the observers to shift their attention to different facial features. This may enhance attention and motor simulations particularly in fragile or neurologically impaired face systems, which explains the increased performance for dynamic faces in those populations.

We should acknowledge that we did not objectively assess whether the amount of physical information conveyed by our static and dynamic stimuli is identical or different for the categorization of facial expressions. We therefore cannot completely rule out the hypothesis that temporal properties provide a considerable amount of information that cannot be completely rendered by static facial cues. Interestingly, Gold and colleagues (2013) have observed that dynamic stimuli do not seem to provide additional information for the recognition of facial expressions than what is already offered by static facial cues in normal observers, by measuring the amount of information carried by static and dynamic facial expressions. In our study, we created a static version of the stimuli by retaining only the highest amplitudes of each AU involved in the model, a procedure that should have led to the representation of “optimal” static signals. In fact, by using a comparable approach, Fiorentini and Viviani (2011) have revealed the absence of an advantage between the categorization of static and dynamic facial expressions

in normal observers. In line with those results, our findings show equal performance for our healthy control observers between static and dynamic stimuli, with the exception of surprise. Thus, PS's ability to correctly categorize dynamic facial expressions does not seem to stem from the physical information available but rather from an adequate psychological ability to make use of this information – a process that most probably occurs in the pSTS.

4.5.3 The special case of “fear”

Our results support the hypothesis of a defective processing of fear in brain-damaged patients. PS was strongly impaired in recognizing the dynamic facial expression of fear from her mental model, the static version from Ekman and Friesen's faces, and the static version of the average mental model. It is worth noting that our age-matched control group displayed an effective categorization of emotions from PS's static and dynamic reconstructed models with respect to all facial expressions except for fear. This observation suggests defective internal representations for this expression in the patient. However, when PS was stimulated by optimal fear dynamic inputs (i.e., the average dynamic mental models), she performed correct categorization.

The impairment for the categorization of fear for static images resonates with the findings obtained in the case of another brain-damaged patient. Adolphs and colleagues (2005) discovered selective impairment in fear recognition with a patient (SM) presenting bilateral damage of the amygdala. SM did not spontaneously use the eyes during fear decoding in static images, but rather the mouth, which resulted in impairment in fear recognition. SM was able to categorize fear only when forced to look at the diagnostic region for this expression, i.e., the eye (Smith et al., 2005). Similarly, the analysis comparing PS's model for fear with the Ekman model showed a clear tendency to emphasize information from the mouth region in PS, as compared to the eyes. However, PS's lesions are anatomically sparing of the amygdala (Sorger et al., 2007) and cannot account for this deficit. In fact, Adolphs (2013) suggests that the processing of fear involves a complex distributed network featuring interactions among diverse cortical regions, rather than a single, localized "fear center". However, little is known regarding which structures play a key role in this brain circuit and the way these different regions interact together. Thus, PS's selective impairment in fear recognition could also stem from her brain lesions, micro-lesions such as axonal shearing (Sorger et al., 2007), or the missing interactions between her lesions and other cortical structures.

PS's impairment for fear is inconsistent with results obtained in a previous fMRI study investigating her emotional attention (Peelen, Lucas, Mayer, & Vuilleumier, 2009). In this study, a target face was presented among an array of distractor faces in a visual search task. The target differed from the distractor either in terms of identity, identity and emotional expression

(fear or happiness), or identity and colour. PS and the controls showed a faster detection for fear and happiness in comparison with the neutral face conditions. PS also displayed a similar advantage for fear in a second change detection task, detecting significantly more changes when the changed face was fearful, as compared to when it was neutral. This behavioural pattern of results was paralleled with according neural response biases that indicate emotional attention in the brain areas coding for faces, which suggests a normal processing for fear. However, paying attention to a particular facial expression of emotion does not necessarily involve an explicit correct categorization for this expression. PS could have reached this pattern of results by using low-level cues that distinguish fear from neutral or happy signals, as she is relatively preserved for those expressions. PS's deficit for fear might also be clarified by recording her brain activation during the decoding of her and the average models in a future neuroimaging study.

4.6 Conclusions

The adequate categorization of facial expressions is a critical feature for adaptive social interactions. Our general goal was to understand whether face information that is used for identity and emotional expression categorization tap into common or distinct representational systems. We isolated information use for facial expressions in a *pure* case of acquired prosopagnosia with a lesion encompassing the right inferior occipital gyrus. PS's reconstructed mental models revealed a normal use of all facial features and muscles (i.e., AUs of the FACS) for the representation of facial expressions. This is in stark contrast with the suboptimal information that she uses for retrieving face identity (i.e., the mouth and the external contours). These data suggest that the face system does not rely on a unique representational system to code facial features for identity and expression, or at least, it relies on distinct cortical pathways to access them, flexibly adapting to visual and task constraints. In addition, our observations indicate that these cortical routes are modulated by the use of dynamic information, which facilitates the correct categorization of facial expressions in the patient. The inferior occipital gyrus plays a critical role for the decoding of static images, and the patient presents a selective impairment in the decoding of static expressions. Conversely, the patient displays a normal performance in effectively decoding facial expressions from dynamic faces. The posterior superior temporal sulcus, which is spared in the patient, would be sufficient to effectively achieve this task. This result reinforces the view of the existence of a cortical pathway carrying out directly face signals from the early visual cortex to the pSTS, thus providing new insights into the normal face operating system. Altogether, our data also question the conclusions obtained from patients by using unnatural static images and emphasize the need for a future neuroimaging study with the same patient in order to consolidate and provide a fine-grained picture of the present findings.

4.7 Acknowledgements

We are extremely grateful to PS for her devoted participation in this research study as well as to all our healthy participants. This study was supported by the National Center of Competence in Research (NCCR) Affective Sciences, financed by the Swiss National Science Foundation (n° 51NF40-104897).

4.8 Supplementary movies

Movie 1. Randomly generated facial animation. An example of 4D stimulus (3D + time) generated by the computer graphics platform on one experimental trial is presented here. In this example, the platform randomly selected AUs 2L, 7R, 39, and 14L and random values for each of the six temporal parameters, and then combined these parameters with the shape and texture information of a Western Caucasian male identity.



http://perso.unifr.ch/roberto.caldara/cortex/Movie_1.mov

Movie 2. PS's dynamic reconstructed mental models of emotions. Rows: happiness, surprise, fear, disgust, anger, and sadness. Columns: emotion intensity “very low”, “medium”, and “very high”.



http://perso.unifr.ch/roberto.caldara/cortex/Movie_2.mov

Movies 3, 4, 5, 6 and 7. Dynamic reconstructed mental models of emotions of each control observer. *Rows:* happiness, surprise, fear, disgust, anger, and sadness. *Columns:* emotion intensity “very low”, “medium”, and “very high”.



PS:

http://perso.unifr.ch/roberto.caldara/cortex/control_1.mov

and the other controls:

http://perso.unifr.ch/roberto.caldara/cortex/control_2.mov

http://perso.unifr.ch/roberto.caldara/cortex/control_3.mov

http://perso.unifr.ch/roberto.caldara/cortex/control_4.mov

http://perso.unifr.ch/roberto.caldara/cortex/control_5.mov

CHAPTER 5

TRACKING THE RECOGNITION OF STATIC AND DYNAMIC FACIAL EXPRESSIONS OF EMOTION ACROSS LIFE SPAN

Richoz, A. R., Lao, J., Pascalis, O., & Caldara, R. (submitted). Tracking the recognition of static and dynamic facial expressions of emotion across life span.

Research highlights

- ◇ Assessing static and dynamic facial expression recognition throughout life span
- ◇ Data-driven estimation of the peak efficiency for every expression
- ◇ Quantification of the steepness of improvement and decline
- ◇ New evidence for a dynamic advantage for facial expression recognition (FER)
- ◇ The use of dynamic faces is critical while assessing FER in young children and elderly people

Keywords: Facial expressions of emotion, Life span, Static, Dynamic, Aging

5.1 Abstract

In everyday life, the effective transmission and decoding of dynamic facial expressions of emotion is omnipresent and critical for adapted social interactions. Thus, common intuition would suggest an advantage for dynamic facial expression recognition (FER) over the static snapshots routinely used in most experiments. However, while many studies reported an advantage in the recognition of dynamic over static expressions in clinical populations, results obtained from healthy participants are contrasted. To clarify this issue, we conducted a large cross-sectional study to investigate FER across the lifespan in order to determine if age is a critical factor to account for such discrepancies. Over 400 observers (age range 5–96) performed recognition tasks of the six basic expressions in static, dynamic, and shuffled (temporally randomized frames) conditions, normalized for the amount of energy sampled over time. We applied a Bayesian hierarchical step linear model to capture the nonlinear relationship between age and FER for the different viewing conditions. While replicating the typical accuracy profiles of FER, we determined the age at which peak efficiency was reached for each expression and found greater accuracy for most dynamic expressions across lifespan. This advantage in the elderly population was driven by a significant decrease in performance for static images, which was twice as large as for the young adults. Our data posit the use of dynamic stimuli as being critical in the assessment of FER in the elderly population, inviting caution when drawing conclusions from the sole use of static face images to this aim.

5.2 Introduction

Human faces convey a wealth of dynamic signals that are critical for an adequate and rapid categorization of the emotional states of others. Yet, the vast majority of studies investigating expression recognition have relied on static images that commonly display the apex or the highest state of a given expression. In everyday life, however, facial expressions are rarely transmitted and decoded through static snapshots of internal states. Natural human interactions are a highly dynamic (and multimodal) phenomenon, with faces evolving over time while transmitting distinct signals to convey diverse emotional states. Dynamic expressions provide observers with additional cues related to their inherent temporal properties, such as their unfolding speed (slow *vs.* fast) (Bould & Morris, 2008; Bould et al., 2008; Kamachi et al., 2001), rise time (from the neutral to the highest state) (Recio, Schacht, & Sommer, 2013; Jack, Garrod, & Schyns, 2014) or intensity (Bould et al., 2008), critical for an adequate categorization. Therefore, dynamic faces are richer and ecologically more valid depictions of the way expressions are encountered in everyday life compared to static images (e.g., Johnston, Mayes, Hughes, & Young, 2013; Paulmann, Jessen, & Kotz, 2009; Trautmann, Fehr, & Herrmann, 2009). Interestingly, from an evolutionary perspective, humans have had more experience with dynamic faces, as static pictures only appeared during the last century with the advent of photography and the rapid expansion of digital tools and social networks. The decoding of static faces is also a learnt behavior that develops throughout life. As the human visual system is from birth on steadily stimulated by dynamic signals from faces with minimal exposure to static faces, common intuition would suggest the existence of a particular expertise to decode such events, with the presence of an advantage for the recognition of dynamic over static expressions.

Previous studies that have attempted to investigate this question have yielded inconsistent findings (for a review, see Alves, 2013; Fiorentini & Viviani, 2011; Kätsyri, 2006; Krumhuber, Kappas, & Manstead, 2013). Some behavioral studies have revealed an advantage (e.g., Ambadar, Schooler, & Cohn, 2005; Cunningham & Wallraven, 2009; Giard & Peronnet, 1999; Knappmeyer, Thornton, & Bulthoff, 2003; Paulmann et al., 2009; Wehrle, Kaiser, Schmidt, & Scherer, 2000) whereas others have revealed that the benefits of dynamic cues in facial expression recognition may be minimal (e.g., Gold et al., 2013) or inexistent (e.g., Fiorentini & Viviani, 2011). These contrasting findings suggest that the dynamic advantage for facial expression recognition is not as straightforward as it may appear. Rather, it seems that the physical properties of the stimuli presented as well as clinical or neuropsychological conditions influence the extent to which dynamic displays lead to processing benefits (Ambadar et al., 2005; Bould, Morris, & Wink, 2008; Wallraven, Breidt, Cunningham, & Bülthoff, 2008).

Several studies have shown that the beneficial effects of dynamic events are particularly relevant in suboptimal situations when the physical information available is limited (Ambadar et al., 2005; Bould et al., 2008), deteriorated, or blurred (Ehrlich, Schiano, & Sheridan, 2000; Kätsyri & Sams, 2008; Wallraven et al., 2008). For example, Wallraven et al. (2008) found that dynamic events increased recognition accuracy of computer animated facial expressions whose texture or shape were systematically degraded. Similarly, by comparing the ability of observers to recognize expressions from schematic and natural faces, Kätsyri and Sams (2008) and Ehrlich, Schiano, and Sheridan (2000) discovered a recognition advantage for dynamic expressions with schematic but not natural faces. Along the same lines, other studies have revealed that dynamic events provide compensatory cues when subtle facial expressions are presented (Ambadar et al., 2005; Bould et al., 2008). With subtle expressions, additional temporal information may be essential to disambiguate the uncertainty introduced by the lack of intensity.

Similarly, an advantage is noticeable when it comes to clinical conditions, as dynamic information provides compensatory cues in suboptimal situations. Dynamic presentations facilitate the recognition of facial expressions in adults and children with intellectual disability (Harwood, Hall, & Shinkfield, 1999), pervasive developmental disorder (Uono, Sato, & Toichi, 2010), and autism (Back, Ropar, & Mitchell, 2007; Gepner, Deruelle, & Grynfeldt, 2001; Tardif, Lainé, Rodriguez, & Gepner, 2007; but see Kätsyri, Saalasti, Tiippana, von Wendt, & Sams, 2008, for Asperger Syndrome). In neuropsychology, several brain injury studies have shown increased recognition performance when dynamic expressions were used (Adolphs, Tranel, & Damasio, 2003; Humphreys, Donnelly, & Riddoch, 1993; Richoz, Jack, Garrod, Schyns, & Caldara, 2015). For example, Humphreys, Donnelly, and Riddoch (1993) reported the case of an agnosic patient who was significantly impaired at identifying facial identity and facial expressions when exposed to static images. In contrast, his performance was proficient when asked to judge a subset of facial expressions (i.e., smiling, frowning, or surprise) from dynamic faces animated by light dots. On the same line, we recently investigated the ability of a prosopagnosic patient – the well-studied case of PS – with multiple and extensive brain lesions in the occipitotemporal cortex – to recognize facial expressions from static and dynamic faces. Our findings revealed that the patient PS was selectively impaired in decoding static expressions, while showing normal performance for the decoding of dynamic emotional expressions. This observation favours the existence of distinct representational systems for static and dynamic expressions, or dissociable cortical pathways to access them (Richoz et al., 2015). Noteworthy, the advantage for processing dynamic faces in PS is related to a suboptimal information use for static (i.e., bias towards the mouth) compared to dynamic faces (i.e., all face features) (Fiset et al., 2017).

While several neuropsychological studies have shown that the dynamic properties of human facial expressions provide significant processing advantages, other behavioral studies involving healthy observers suggest that this might not be the case (Bould & Morris, 2008, for expressions of high intensity; Christie & Bruce, 1998; Fiorentini & Viviani, 2011; Gold et al., 2013; Jiang et al., 2014; Kamachi et al., 2001, Experiment 2). By using a threshold model, Fiorentini and Viviani (2011), for example, reported that neither reaction times nor identification accuracy were more accurate for the dynamic as compared to the static expressions. Similar findings were reported in a later study by Gold et al. (2013). Their results revealed that recognition rates were nearly identical when participants were exposed to static, dynamic, shuffled (temporally randomized expressions), or reversed expressions. This suggests that the temporal properties provided by moving faces are not necessary for observers to reliably categorize emotional expressions. Altogether, these studies suggest that a healthy visual system seems to be powerful enough to efficiently recognize intense expressions from static faces, leaving only a non-significant benefit to the processing of dynamic facial expressions. By contrast, in clinical conditions, the muscular movements associated with the temporal unfolding of an expression may force the observers to shift their attention to different facial features. This may enhance attention and motor simulations in fragile or neurologically impaired face systems, which may explain the increased performance with dynamic signals in these populations.

Interestingly, there are stages in healthy observers during which the perceptual system is also particularly fragile or immature. For example, from early infancy to late adolescence, the brain undergoes a wide array of anatomical and functional changes as it develops (e.g., Blakemore, 2012; Blakemore & Choudhury, 2006; Casey, Tottenham, Liston, & Durston, 2005; Durston et al., 2001). Similarly, during normal aging, the cognitive functions decline, which is induced by age-related loss of synaptic contacts, neural apoptosis (e.g., Raz, 2000; Rossini, Rossi, Babiloni, & Polich, 2007), reduction in cerebral blood flow (e.g., Chen, Rosas, & Salat, 2011), or volume reduction in different brain regions (e.g., amygdala, hippocampus, frontal cortex, Calder et al., 2003; Jack et al., 1997; Navarro & Gonzalo, 2017; Ruffman, Henry, Livingstone, & Phillips, 2008). Considering the increased vulnerability of the brain under neural architectural changes (Andersen, 2003; Hof & Morrison, 2004), it is possible that healthy young children and normal aging adults also benefit from the presentation of dynamic faces. However, only a few developmental studies have compared facial expression recognition in children using both static and dynamic stimuli (Nelson, Hudspeth, & Russell, 2013; Nelson & Russell, 2011). These studies yielded equivocal results, none of them revealing a significant advantage for dynamic over static stimuli; two studies even pointed to differences favoring static stimuli (Nelson & Russell, 2011; Widen & Russell, 2015). Nevertheless, most of these studies tested facial expression recognition with the use of a single actor and provided additional information about face, body movements, and vocal intonations, which may have facilitated expression

recognition. In the aging literature, a small number of studies examined facial expression recognition with static and dynamic faces (Grainger, Henry, Phillips, Vanman, & Allen, 2015; Krendl & Ambady, 2010; Sze, Goodkind, Gyurak, & Levenson, 2012). Although most of these studies pointed to a dynamic advantage for the recognition of facial expressions, they did not use a database of static and dynamic stimuli controlled for the amount of *low-level* visual information carried over time (Grainger et al., 2015; Sze et al., 2012), were limited to a subset of emotional expressions (Krendl & Ambady, 2010), included participants in only one condition (Krendl & Ambady, 2010) or relied on dynamic movies that were not displaying natural expressions (Grainger et al., 2015). These methodological issues considerably limit firm conclusions on the potential benefits of dynamic cues for the recognition of facial expressions in elderly people.

Developmental studies have reported an early tuning to cultural specific expressions (Geangu, et al., 2016 – for a review, see Caldara, 2017) and emotion-dependent differences in the development of facial expression recognition abilities, with some expressions being recognized earlier (e.g., happiness) than others (e.g., fear) (Durand, Galloway, Seigneuric, Robichon, & Baudouin, 2007; Gao & Maurer, 2010; Gross & Ballif, 1991; Herba & Phillips, 2004; Rodger, Vizioli, Ouyang, & Caldara, 2015). Similarly, studies with elderly people have shown that the recognition of some expressions decreases with increasing age, while the recognition of others remains stable or even improves (Calder et al., 2003; MacPherson, Phillips, & Della Sala, 2002; Sullivan & Ruffman, 2004b; Zhao, Zimmer, Shen, Chen, & Fu, 2016). Most of these studies were however conducted with static posed images and only little is known about the effects of aging on the recognition of genuine dynamic emotional expressions.

To fill this gap in the developmental literature, we investigated whether the advantage for dynamic stimuli extends to other populations with immature (i.e., young children) or fragile (i.e., elderly adults) face processing systems. We conducted a large cross-sectional study involving over 400 observers (age range 5–96) in order to investigate facial expression recognition from early to elderly age. Observers performed categorization tasks of the six basic expressions (anger, disgust, fear, happiness, sadness, and surprise) in three conditions: static, dynamic, and shuffled (temporally randomized frames, Gold et al., 2013). Importantly, we relied on a specific database of static, dynamic, and shuffled stimuli created by Gold et al. (2013). Our experimental choice was driven by the fact that these authors also used an ideal observer model to objectively measure the amount of *low-level* physical information carried by the stimuli. It is worth noting that most studies investigating the presence of a dynamic advantage (e.g., Ambadar et al., 2005; Bould & Morris, 2008; Bould et al., 2008; Cunningham & Wallraven, 2009; Fiorentini & Viviani, 2011; Kätsyri & Sams, 2008) directly compared participants' recognition rates in the static and dynamic conditions without controlling the amount of *low-level* information physically

available to the observers. As mentioned by Gold et al. (2013), the absence of an objective measure of stimulus information makes it difficult, in most cases, to determine whether increased recognition rates are due to adequate categorization skills, to the amount of physical information available, or a combination of both the factors. By comparing human expression recognition scores with the performance of a statistically ideal observer, Gold et al. (2013) reported that their dynamic stimuli did not provide additional *low-level* information than what was already offered by their static snapshots (for additional information, see Gold et al., 2013). In addition to this approach, we modelled the relationship between age and facial expression recognition by using a hierarchical Bayesian approach, with a step linear model. Our results revealed emotion-specific advantages for dynamic stimuli. More specifically, while participants displayed nearly identical categorization performance for the static and dynamic expressions of fear and sadness, *all* the other expressions were more readily labelled as correct when featuring dynamic displays. Overall, the results of this study provide a comprehensive and detailed view of the way in which static and dynamic expressions are recognized across the human life span.

5.3 Material and methods

The experiment script, raw data, and analysis codes are open to access on Github (https://github.com/iBMLab/Static_dynamic).

5.3.1 Participants

A total of 444 healthy observers participated in the current study. Subjects who did not respond at least once to all expressions on the first condition/block were excluded from the analyses ($N = 32$), leaving a total number of 412 participants. Their exclusion is based on the difficulty to determine whether they actually did not recognize the expression presented or did not correctly understand the task. A future research paper will investigate the systematic errors of the participants that were excluded.

We intended to collect data from 20 participants in each age group ranging from 5 to 96 years of age. The groups were comprised as follows: 5–6 year old ($N = 27$, 17 females), 7–8 year old ($N = 24$, 17 females), 9–10 year old ($N = 22$, 11 females), 11–12 year old ($N = 22$, 14 females), 13–14 year old ($N = 24$, 10 females), 15–16 year old ($N = 21$, 8 females), 17–18 year old ($N = 21$, 16 females), 19–20 year old ($N = 31$, 27 females). From the age of 21 to the age of 96, six different groups were created: 21–30 year old ($N = 31$, 23 females), 31–40 year old ($N = 23$, 13 females), 41–50 year old ($N = 33$, 22 females), 51–60 year old ($N = 30$, 18 females), 61–80 year old ($N = 31$, 25 females), and 81–96 year old ($N = 40$, 30 females).

All participants had normal or corrected to normal vision, with no neurological or psychiatric history. Children were recruited from primary and high schools in the area of Fribourg, Switzerland. Parental consent was required for all children under the age of 16. Participants above 16 were recruited at the University of Fribourg, through social networks or advertisements. Observers from the University obtained course credits for their participation. All participants signed a consent form that described the main goals of our experiment.

Elderly people were recruited and tested in senior housings in the Fribourg region. We used the Mini-Mental State Examination (MMSE, Folstein, 1975) in order to determine the eligibility of the elderly people aged 60 and over. This brief cognitive screening test, which has been extensively used and validated since its creation in 1975, allows the assessment of different cognitive functions, such as memory, orientation, attention, language, and recall, through 11 questions, with a maximum score of 30. Elderly people with a score below 24 were excluded from our study ($N = 3$), as this score has been set as the most commonly used cutoff score for cognitive impairment (Mitchell, 2009). The Ethical Committee of the Department of Psychology of the University of Fribourg approved the study reported here.

5.3.2 Stimuli

We used the same stimuli as those used by Gold et al. (2013). In order to create their database, Gold et al. (2013) asked eight individuals (four females) to record the six basic facial expressions of emotion (i.e., anger, disgust, fear, happiness, sadness and surprise) (Ekman & Friesen, 1976). The dynamic faces evolved from a neutral state to a full-blown expression at a frame rate of 30 frames/s. All expressions reached their apex within 30 frames. If the fully articulated expression was reached before 30 frames, one to four supplementary apex frames were appended, but as the actors were asked to maintain the apex for several seconds, this happened for only seven out of 48 movies (for more details, see Gold et al., 2013). Faces were presented in black and white and cropped at the hairline to present only the internal facial features. Previous experiments have shown that external features attract children's attention (Leitzke & Pollak, 2016). Moreover, the faces were centred and seen through an oval aperture, which was placed in the middle of a grey-coloured background. The borders of the oval aperture were slightly blurred in order to produce a progressive transition between the background and the faces (Gold et al., 2013). The faces were resized from the original experiment and measured each 768 pixels in height and 768 pixels in width. They subtended a visual angle of 12° on the screen, at a viewing distance of 65 cm. All faces were equated for luminance and contrast.

Based on these dynamic sequences, Gold et al. (2013) generated two other sets of stimuli: a set of frozen images (static condition) and a set of temporally randomized dynamic frames

(shuffled condition) (see Figure 5.1; supplementary videos related to this article can be found under the specific links). In the static condition, movies were created by taking the apex frame of each dynamic sequence and replicating it 30 times in a row. In the shuffled condition, movies were generated by randomly selecting the individual frames of the dynamic sequences. This condition was originally designed to assess whether human observers were sensitive to the temporal development of an expression over time (i.e., order of frames). The results reported by Gold et al. (2013) revealed that recognition efficiency did not significantly differ between the dynamic and shuffled expressions in young adults, suggesting that young adults are insensitive to the temporal properties associated with the unfolding of an expression.

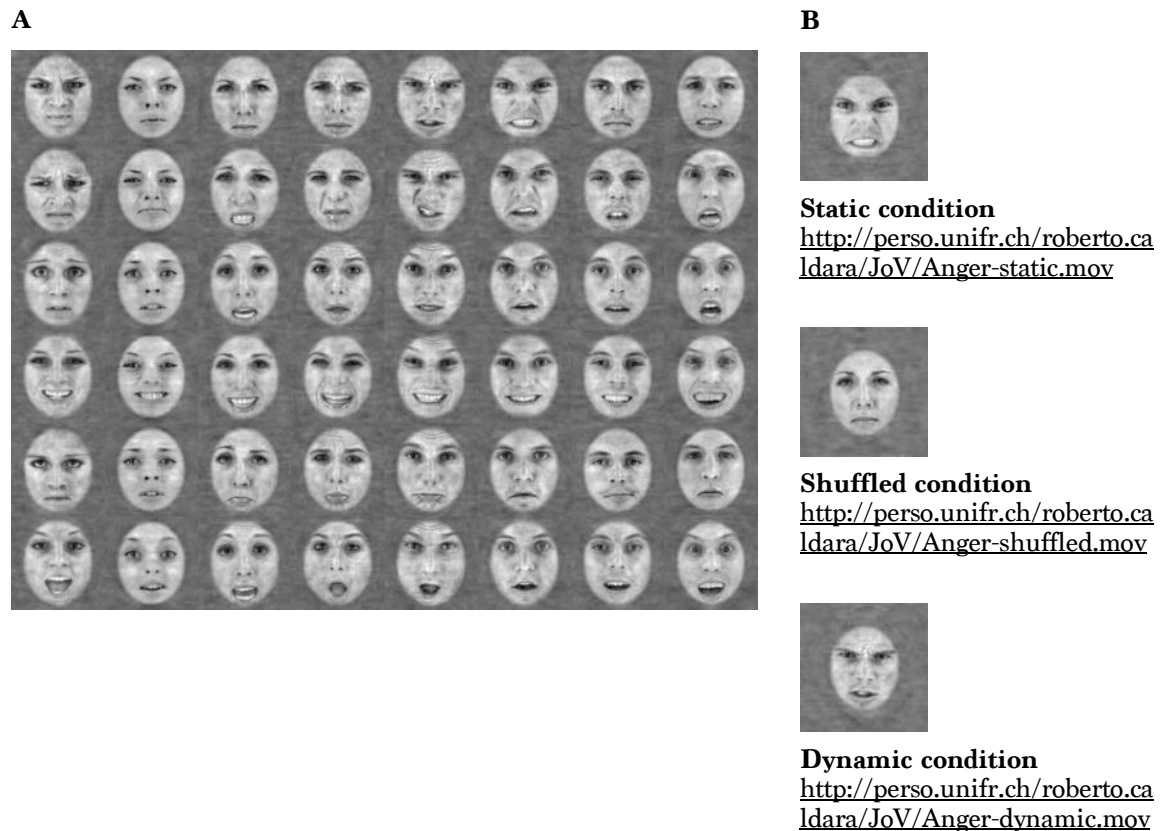


Figure 5.1. Examples of the three different kinds of stimuli used in our study. Panel A represents the static images of each actor (column) for the six expressions (row). Panel B represents stimuli examples for the static, dynamic and shuffled conditions. Please note the insertion of noise in the static condition in order to normalize the amount of energy sampled over time across conditions. Click on the specific links to visualize the videos. Adapted with permission from Gold et al. (2013).

We normalized all stimuli for their low-level properties and the amount of energy sampled over time, even for the static condition. The video stimuli were normalized across all frames and all expressions using the SHINE toolbox with the default option (Willenbockel et al., 2010). In order to partly account for the differences in visual input between static and dynamic stimuli, we computed the raw pixel intensity differences between each frame of the dynamic movies. We then added these intensity differences to each frame at random permuted locations in the static images. The stimuli were displayed on a color liquid-crystal display (LCD) with a resolution of 1440 x 900 pixels and a refresh rate of 60 Hz. The whole experiment was programmed in Matlab (Matlab 2014B) using the Psychophysics Toolbox (PTB-3, Brainard & Vision, 1997; Kleiner et al., 2007).

5.3.3 Procedure

Participants were told that they would see faces expressing different kinds of emotions on a computer screen, and their task would be to categorize them as accurately as possible, according to the six following possibilities: anger, disgust, fear, happiness, sadness, and surprise.

In order to familiarize children with the faces and ensure that they understood the conceptual meaning of all expressions, we presented them with printed sheets of the different expressions and asked them to tell us how the person presented on the image was feeling.

All participants sat 65 cm away from a computer screen in a quiet room. Each trial started with a white fixation cross presented at the centre of the screen for 500 ms. The stimuli were then presented in a random order, one at a time, on the centre of the computer screen for a duration of 1 s each (for a schematic representation of the procedure, see Figure 5.2). We used the same stimuli presentation time in all three conditions in order to fully replicate the study by Gold et al. (2013). Note that a presentation time of 1 second was also previously used in other studies with dynamic faces (Adolphs, Tranel, & Damasio, 2003; Recio et al., 2013; Richoz et al., 2015). After each presentation, a response window was displayed on the screen and remained there until the participant answered. Observers categorized each stimulus by using a computer keyboard in which we labelled the keys accordingly. They could press a “I don’t know” labelled key if they were unsure, had not had enough time to see the expression, or did not know the answer. We decided to introduce a “I don’t know” option in order to reduce the noise and response bias produced by the lack of such a key. We gave our participants as much time as required to categorize the expressions and told them that judgement accuracy was important, not the response time. Children under the age of 10, participants who were not familiar with computers, and elderly people above 65 gave their answers verbally to the experimenter who keyed them in. No feedback was provided. The stimuli were blocked by condition. Each

condition consisted of two blocks of 48 trials (eight actors, six expressions) presented twice (96 expressions for each condition), for a total of 288 trials. Participants took part in all three conditions in a counterbalanced random order. The testing was done in one session for adolescents and adults, two or three sessions for participants under 10 or above 65. Before starting the testing phase, participants completed 12 practice trials for each condition.

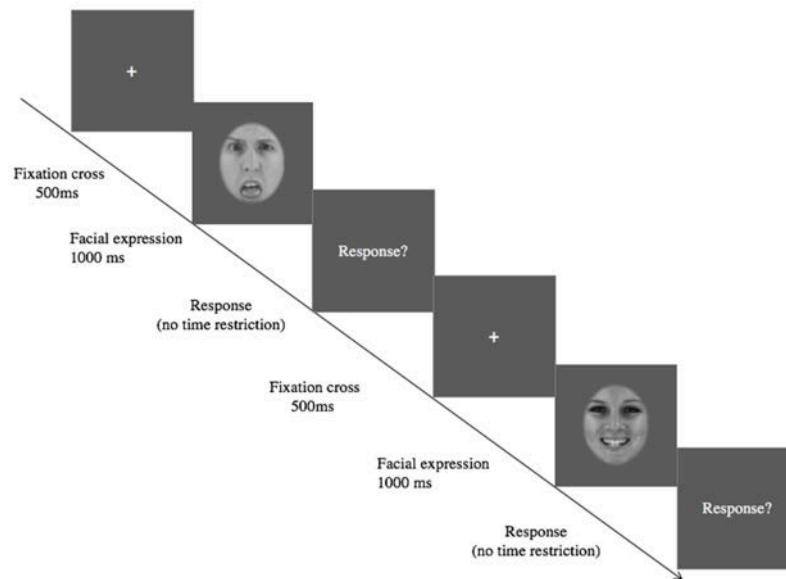


Figure 5.2. Schematic representation of the procedure. Each trial began with a white fixation cross that was presented for 500 ms, followed by a face presented for 1 second, which expressed one of the six basic facial expressions of emotion: anger, disgust, fear, happiness, sadness, and surprise. After each trial, participants were asked to categorize the previously seen expression.

5.3.4 Data Analysis

Data analysis was performed in Python using Jupyter Notebook. Summary statistics by groups are displayed as confusion matrices (Supplementary Figure 5.11A, B, C, D, E) and line plots (Figure 5.4) for each condition.

Bayesian modelling was performed using PyMC3 version 3.2, and the results were displayed using Seaborn and Matplotlib. The main aim of the current study was to determine the underlying function between expression recognition ability and age, conditioned on the basis of different types of visual stimuli. More specifically, we were interested in modelling expression recognition ability as a function of age, expression, and stimuli type (static, dynamic, or shuffled):

$$\text{Recognition Ability} = f(\text{age}, \text{expression}, \text{stimuli type}).$$

Here, recognition abilities were measured using the correct identification (i.e., hit). Importantly, as the target function f is a nonlinear function, in order to capture the increase and then the decrease in the recognition abilities displayed in the data, we constructed a simple step linear function with two linear equations. The first equation captures the increase in recognition abilities before a breakpoint, defined as the momentum in age capturing peak efficiency, whereas the second equation captures the decrease in recognition abilities.

$$\begin{aligned} f_1(\text{age}, \text{expression}, \text{stimuli type}), \text{age} < \tau \\ f_2(\text{age}, \text{expression}, \text{stimuli type}), \text{age} \geq \tau \end{aligned}$$

Here, the breakpoint τ is expressed as a latent variable that is estimated from the model. Both f_1 and f_2 are linear functions of age during the recognition of specific expressions and stimuli type. Thus, the slope of the function f_1 and f_2 (coefficient for age) captures the change in recognition abilities, whereas the intercept of the function captures the general recognition abilities before and after the age τ . We estimated the general dynamic advantage by computing contrasts of the intercepts between the different stimuli type (i.e., static, dynamic, shuffled) and quantified the interaction between stimuli type and age (i.e., whether there is a stronger dynamic advantage in young/old age) by computing contrasts of the slopes. Importantly, considering that the breakpoint τ could occur at different age stages among the expressions, we modelled each expression independently, thus turning the target function into:

$$\begin{aligned} f_1(\text{age}, \text{stimuli type} \mid \text{expression}), \text{age} < \tau \\ f_2(\text{age}, \text{stimuli type} \mid \text{expression}), \text{age} \geq \tau \end{aligned}$$

where recognition ability of different types of stimuli is a step linear function of age conditioned on a specific expression.

In practice, we formulated functions f_1 and f_2 as logistic regressions, with the function output being the success probability p in each trial in the Binomial distribution. The total number of correct responses for one participant during the presentation of one expression and one stimuli type follows a Binomial distribution:

$$k \sim \text{Binomial}(p, n)$$

Thus, this is an extended beta-binomial model with latent variables. The full model is formulated as below:

$$\begin{aligned}
 & i \text{ for each task (dynamic, static, shuffle), } j \text{ for each participant.} \\
 & \text{Hyperpriors of the slope} \\
 & \quad \mu_\beta \sim \text{StudentT}(3, 0, 10) \\
 & \quad \sigma_\beta \sim \text{HalfNormal}(10) \\
 & \text{Hyper-prior of the breakpoint} \\
 & \quad \tau \sim \text{Uniform}(0, 100) \\
 & \text{Hyperprior of the recognition ability at } \tau \\
 & \quad \theta \sim \text{Uniform}(0, 1) \\
 & \quad \kappa \sim \text{Uniform}(0, Nt) \\
 & \text{(reparameterized as the mode of Beta Distribution)} \\
 & \quad a = \theta * (\kappa - 2) + 1 \\
 & \quad b = (1 - \theta) * (\kappa - 2) + 1 \\
 & \text{For each stimuli type } i \in \text{static, dynamic, shuffled:} \\
 & \quad \text{Prior of the recognition ability at } \tau \\
 & \quad \quad \theta_i \sim \text{Beta}(a, b) \\
 & \quad \text{Priors of the breakpoint} \\
 & \quad \quad \tau_i \sim \text{Normal}(\tau, 10) \\
 & \quad \text{Priors of the slopes (} a \text{ indicates before age } \tau \text{)} \\
 & \quad \quad \beta_i^{a,b} \sim \text{Normal}(\mu_\beta, \sigma_\beta) \\
 & \quad \text{The intercepts before and after age } \tau \\
 & \quad \quad b_i^{a,b} = \text{logit}(\theta_i) - \tau_i * \beta_i^{a,b} \\
 & \quad \text{Linear function and invlogit transform} \\
 & \quad \quad \beta_i, \text{Intercept}_i = \begin{cases} \beta_i^a, b_i^a & \text{if } \text{age} < \tau_i \\ \beta_i^b, b_i^b & \text{if } \text{age} \geq \tau_i \end{cases} \\
 & \quad \quad \theta_{i,j} = \beta_i \otimes \text{age}_j + \text{Intercept}_i \\
 & \quad \quad \hat{y}_{i,j} = \text{invlogit}(\theta_{i,j}) \\
 & \quad \text{Observed accurate categorizations} \\
 & \quad \quad k_{i,j} \sim \text{Binomial}(\hat{y}_{i,j}, n_{i,j})
 \end{aligned}$$

As shown above, the slope of each condition is regularized using a weakly informative hyper-prior. The prior of each slope is a normal distribution, with the mean distributed as a zero mean Student- t distribution, with three degrees of freedom and 10 standard deviations, and the standard deviations distributed as a half-normal distribution. The hyper-prior of the breakpoint τ is a Uniform distribution from 0 to 100, which is the overall mean of the condition-specific breakpoint that follows a normal distribution with 10 standard deviations as prior. Importantly, the intercept of the two linear functions f_1 and f_2 is determined by the recognition ability θ at the breakpoint τ . The condition-specific recognition ability θ_i follows a

Beta distribution as prior. Moreover, we re-parameterized the Beta distribution by the mode θ and the concentration κ (Kruschke, 2014, cf. Eq. 9.4, pp. 223). Here, the mode θ follows a Uniform prior between 0 and 1, and κ follows a Uniform prior with 2 as minimum and the number of trials as maximum.

The probabilistic model was built using PyMC3 and we sampled from the posterior distribution using NUTS with automatic differentiation variational inference (ADVI) initialization. We ran four MCMC chains with 3000 samples each; the first 1000 samples were used for tuning the mass matrix and step size for NUTS, and were discarded following this. Model convergence was diagnosed by computing Gelman and Rubin's convergence diagnostic (R-hat, 1992), examining the effective sample size, inspecting the mixing of the traces, and checking whether there is any divergent sample that has been returned from the sampler.

From the posterior distribution, we estimated a) the peak efficiency, namely the point at which observers' recognition performance reaches its maximum before declining; b) the steepness of increase and decrease in recognition abilities; c) differences in the steepness of increase and decrease between different conditions (e.g., dynamic vs. static); and d) the overall processing advantage of the dynamic over the static and the shuffled stimuli. By performing statistical inference directly on the full posterior distribution, we were able to properly quantify the dynamic stimuli effects and their associated uncertainty. A conceptual representation of the model is provided in Figure 5.3.

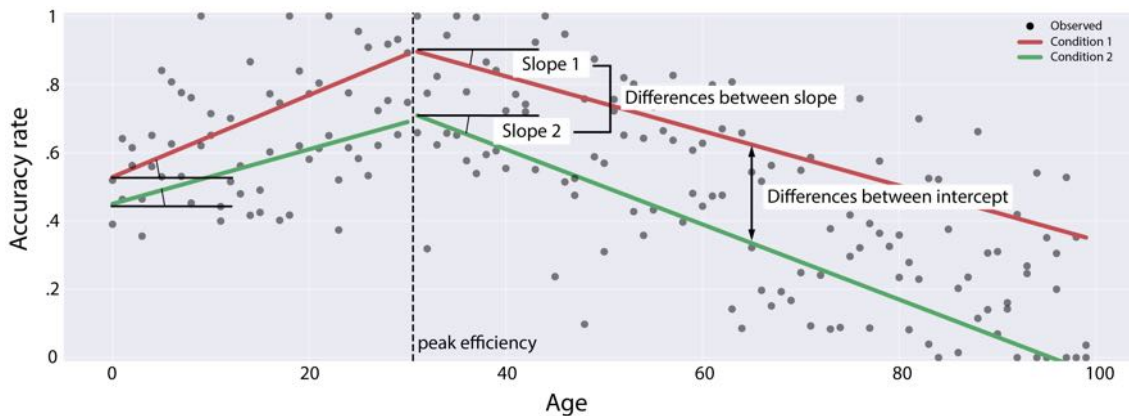


Figure 5.3. A conceptual representation of the step linear model. We are interested in the posterior distribution of the peak efficiency and the contrasts between the posterior distribution of different slopes and the different intercepts.

5.4 Results

The group average categorization performance for each condition is presented in Figure 5.4. The non-linear relationship between age and recognition ability is clearly demonstrated, with differences among conditions clearly visible for some expressions. When the model returns a concave pattern, we refer to the breakpoint as a peak efficiency. This value relates to the point at which recognition performance reaches its apex, also relating to the age at which observers are the most efficient.

For the Bayesian modeling, trace plots, posterior distributions for the key parameters in the model, contrasts of interest, and full numerical reports of the parameter estimations are available in the supplementary results. Below, we report the key findings of the step linear model.

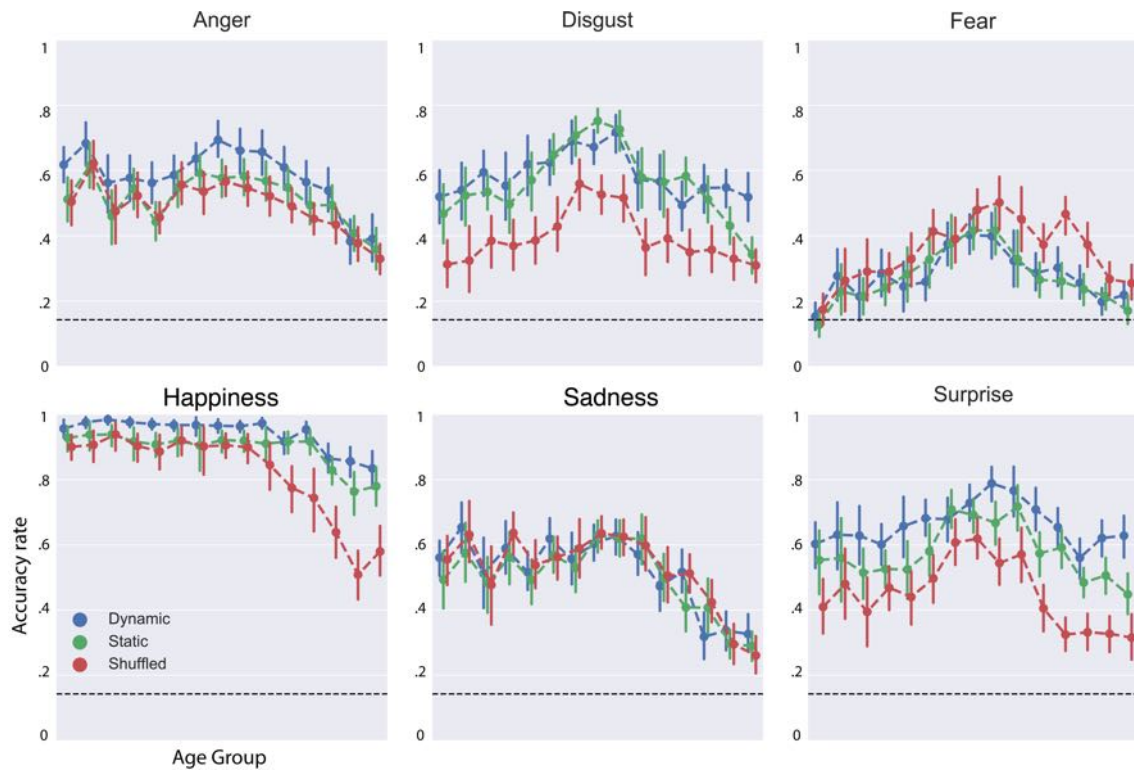


Figure 5.4. Accuracy across age groups for each expression in the three different conditions. Error bars show 95% bootstrap confidence interval for the mean. Age groups were created as follows: 5–6 / 7–8 / 9–10 / 11–12 / 13–14 / 15–16 / 17–18 / 19–20 / 21–30 / 31–40 / 41–50 / 51–60 / 61–70 / 71–80 / 81–96.

Anger

The posterior model fit for the raw data is shown in Figure 5.5. By sampling the full posterior distribution, we estimated that the overall recognition ability for the expression of anger peaks at age 36.13 [22.23, 51.21], 95% highest posterior density interval (hpd). The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 39.17 [31.03, 46.84], static 35.70 [23.59, 46.41], and shuffled 33.22 [18.13, 49.62]. The overall recognition ability of anger at peak efficiency is 0.605 [0.350, 0.872], and the average peak accuracy for each condition is given as follows: dynamic 0.660 [0.627, 0.696], static 0.592 [0.554, 0.628], and shuffled 0.539 [0.489, 0.587]. On an average, participants showed better performance in the dynamic condition as compared to the static and the shuffled conditions, both before (dynamic – static: 0.075 [0.042, 0.106]; dynamic – shuffled: 0.096 [0.060, 0.132]) and after (dynamic – static: 0.043 [0.012, 0.072]; dynamic – shuffled: 0.085 [0.051, 0.118]) peak efficiency. In contrast, the difference between the static and shuffled conditions is quite small (shuffled – static before peak efficiency: -0.021 [-0.061, 0.019]; after peak efficiency: -0.042 [-0.070, -0.013]). The slopes of the step linear functions are the following: dynamic 0.0069 [-0.0011, 0.0156], static 0.0107 [0.0003, 0.0239], shuffled 0.0040 [-0.0072, 0.0180] before peak efficiency; and dynamic -0.0251 [-0.0300, -0.0201], static -0.0183 [-0.0236, -0.0123], shuffled -0.0159 [-0.0211, -0.0108] after peak efficiency. Moreover, the differences of the slope across different conditions are mostly negligible; most of the posterior contrasts are distributed around zero, with the exception of the contrast: dynamic – shuffled: -0.0092 [-0.0162, -0.0020] after peak efficiency (Figure 5.5).

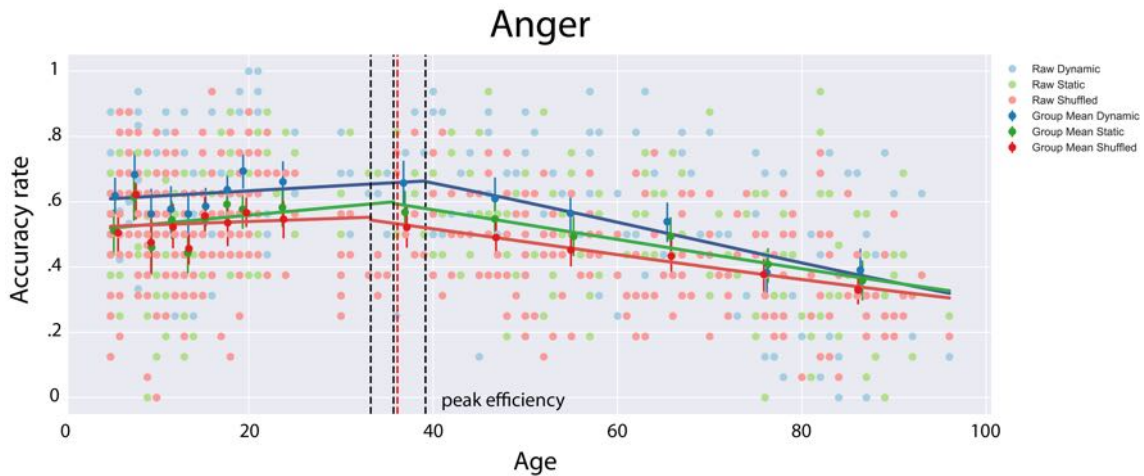


Figure 5.5. Anger. The posterior model fit (solid line) for the expression of anger with the individual performance (scatter plot) and the group average performance (dots with error bars) is given here. The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

Disgust

The overall recognition ability for the expression of disgust peaks at age 18.87. The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 18.02 [15.38, 20.75], static 19.71 [18.19, 21.38], and shuffled 18.14 [16.55, 19.67]. The overall recognition ability of the expression of disgust at peak efficiency is 0.644 [0.380, 0.920], and the average peak accuracy for each condition is the following: dynamic 0.665 [0.644, 0.687], static 0.724 [0.702, 0.744], and shuffled 0.500 [0.475, 0.522]. On an average, participants showed better performance in the dynamic and static conditions as compared to the shuffled conditions, both before (dynamic – shuffled: 0.199 [0.160, 0.240]; static – shuffled: 0.183 [0.146, 0.217]) and after (dynamic – shuffled: 0.191 [0.170, 0.214]; static – shuffled: 0.152 [0.125, 0.177]) peak efficiency. The difference between the dynamic and the static conditions is quite small before peak efficiency (dynamic – static: -0.016 [-0.037 , 0.072]); it is, however, substantial after peak efficiency (0.040 [0.013 , 0.069]). The slopes of the step linear functions are the following: dynamic 0.0494 [0.0318 , 0.0680], static 0.0824 [0.0671 , 0.0993], shuffled 0.0686 [0.0511 , 0.0860] before peak efficiency; and dynamic -0.0104 [-0.0130 , -0.0078], static -0.0225 [-0.0254 , -0.0196], shuffled -0.0128 [-0.0152 , -0.0102] after peak efficiency. Moreover, the slopes of the static condition are steeper than the ones in the dynamic and shuffled conditions. The contrasts of the slopes before peak efficiency are given as follows: dynamic – static: -0.0330 [-0.0577 , -0.0096], shuffled – static: -0.0138 [-0.0360 , 0.0123]; and the contrasts of the slopes after peak efficiency are the following: dynamic – static: 0.0121 [0.0084 , 0.0161], shuffled – static: 0.0097 [0.0059 , 0.0135] (Figure 5.6).



Figure 5.6. Disgust. The posterior model fit (solid line) of the expression of disgust with the individual performance (scatter plot) and the group average performance (dots with error bars) is given here. The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

Fear

The overall recognition ability of the expression of fear peaks at around age 20.83. The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 20.87 [18.71, 23.18], static 19.72 [17.63, 21.43], and shuffled 21.79 [20.17, 23.51]. The overall recognition ability of fear at peak efficiency is 0.446 [0.168, 0.697]; the average peak accuracy for each condition is the following: dynamic 0.399 [0.372, 0.430], static 0.416 [0.390, 0.443], and shuffled 0.526 [0.494, 0.557]. On an average, participants showed better performance in the shuffled condition compared to the other two conditions, both before (shuffled – dynamic: 0.056 [0.013, 0.099]; shuffled – static: 0.055 [0.024, 0.083]) and after (shuffled – dynamic: 0.094 [0.071, 0.118]; shuffled – static: 0.113 [0.089, 0.135]) peak efficiency; however, the difference between the dynamic and static conditions is quite small (dynamic – static before peak efficiency: -0.002 [$-0.043, 0.033$]; after peak efficiency: 0.018 [$-0.003, 0.040$]). The slopes of all conditions are comparable: dynamic 0.0706 [0.0546, 0.0867], static 0.0927 [0.0747, 0.1139], shuffled 0.0913 [0.0764, 0.1056] before peak efficiency; and dynamic -0.0151 [$-0.0186, -0.0115$], static -0.0189 [$-0.0224, -0.0157$], shuffled -0.0176 [$-0.0213, -0.0141$] after peak efficiency. The maximum contrasts of the slopes before peak efficiency is given as follows: dynamic – static: -0.022 [$-0.0473, 0.0022$]; and the maximum contrasts of the slopes after peak efficiency is the following: dynamic – static: 0.0038 [$-0.0010, 0.0086$] (Figure 5.7).

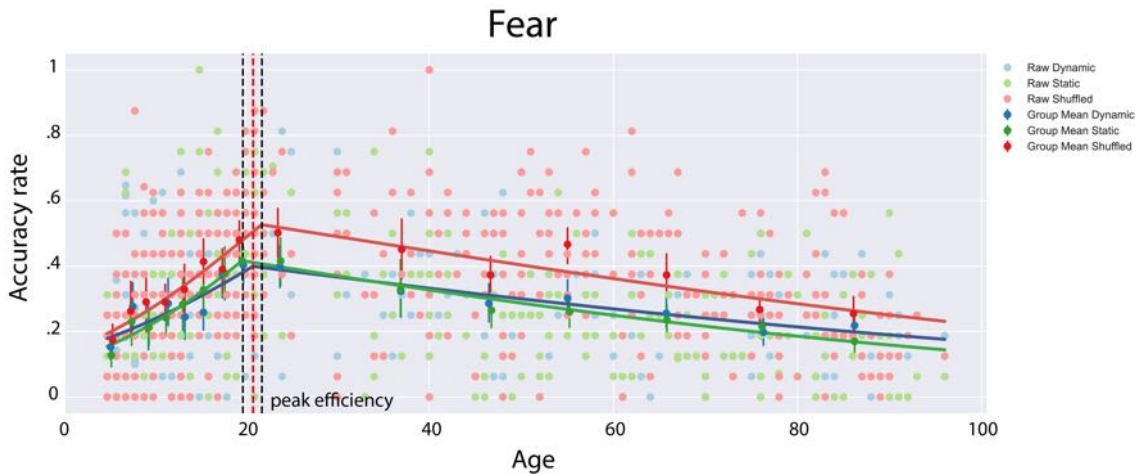


Figure 5.7. Fear. The posterior model fit (solid line) of the expression of fear with the individual performance (scatter plot) and the group average performance (dots with error bars) is presented here. The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

Happiness

Unlike for the other facial expressions, the overall recognition ability for the expression of happiness reaches optimal levels of task performance at a very young age, declining slowly throughout the life span. Nonetheless, our model identifies a breakpoint at around age 57.98, with a large uncertainty. Importantly, the accuracy rate estimated at this breakpoint is not the apex in recognition performance, but rather the peak of the decline (i.e., the model did not return a concave pattern). The posterior expectation of the age at this breakpoint is given as follows: dynamic 61.27 [24.18, 93.42], static 50.25 [35.90, 62.01], and shuffled 62.22 [21.95, 81.00]. The overall recognition ability of this expression at the breakpoint is 0.855 [0.664, 0.999], with dynamic 0.895 [0.819, 0.972], static 0.898 [0.868, 0.931], and shuffled 0.660 [0.541, 0.896]. Overall, participants performed better in the dynamic condition as compared to the static and shuffled conditions, both before (dynamic – static: 0.040 [0.027, 0.054]; dynamic – shuffled: 0.106 [0.063, 0.143]) and after (dynamic – static: 0.060 [0.029, 0.095]; dynamic – shuffled: 0.269 [0.203, 0.319]) the breakpoint. Participants also performed better in the static than in the shuffled condition (static – shuffled before the breakpoint: 0.066 [0.018, 0.105] and after the breakpoint: 0.209 [0.156, 0.255]). The slopes of the step linear functions are given as follows: dynamic -0.0248 [-0.0335 , -0.0121], static -0.0078 [-0.0172 , 0.0007], shuffled -0.0306 [-0.0375 , -0.0163] before the breakpoint; and dynamic -0.0239 [-0.0387 , -0.0023], static -0.0298 [-0.0381 , -0.0217], shuffled -0.0175 [-0.0355 , 0.0026] after the breakpoint. The differences of the slopes across the different conditions are mostly negligible. Most of the posterior contrasts are distributed around zero, with the largest contrasts being the following before the breakpoint: dynamic – static: -0.0170 [-0.0297 , -0.0036], and shuffled – static: -0.0228 [-0.0353 , 0.0087] (Figure 5.8).

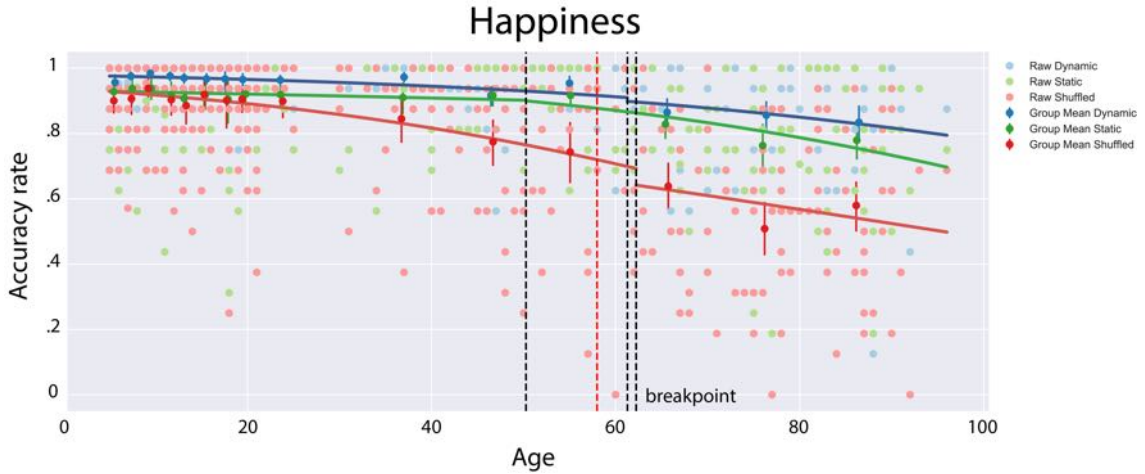


Figure 5.8. Happiness. The posterior model fit (solid line) of the expression of happiness with the individual performance (scatter plot) and the group average performance (dots with error bars) is presented here. The overall breakpoint is shown as the red vertical dashed line, and the condition-specific breakpoints are represented by the black dashed lines.

Sadness

The overall recognition ability for the expression of sadness peaks at the age of 28.96. The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 27.87 [21.44, 33.90], static 28.12 [23.50, 32.62], and shuffled 30.52 [26.05, 34.74]. The overall recognition ability of sadness at peak efficiency is 0.638 [0.408, 0.888]; the average peak accuracy for each condition is the following: dynamic 0.605 [0.572, 0.636], static 0.631 [0.602, 0.660], and shuffled 0.653 [0.622, 0.681]. The categorization accuracy rates of all conditions are comparable both before and after peak efficiency. The maximum contrasts of the average performance before peak efficiency is given as follows: shuffled – static 0.0277 [–0.0011, 0.0561]; the maximum contrasts of the average performance after peak efficiency is the following: shuffled – static 0.0097 [–0.0156, 0.0362]. Similarly, all conditions show comparable slopes: dynamic 0.0075 [–0.0031, 0.0181], static 0.0235 [0.0117, 0.0347], shuffled 0.0179 [0.0079, 0.0285] before peak efficiency; and dynamic –0.0223 [–0.0262, –0.0186], static –0.0265 [–0.0306, –0.0223], shuffled –0.0302 [–0.0346, –0.0261] after peak efficiency. The maximum contrast between slopes before peak efficiency is the following: dynamic – static: –0.0160 [–0.0320, –0.0010]; the maximum contrast after peak efficiency is given as follows: dynamic – shuffled: 0.0079 [0.0022, 0.0137] (Figure 5.9).



Figure 5.9. Sadness. The posterior model fit (solid line) of the expression of sadness with the individual performance (scatter plot) and the group average performance (dots with error bars). The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

Surprise

The overall recognition ability of surprise peaks at age 22.47. The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 23.55 [20.52, 26.91], static 24.30 [20.36, 28.26], and shuffled 19.34 [17.34, 21.76]. The overall recognition ability of surprise at peak efficiency is 0.692 [0.466, 0.953]; the average peak accuracy for each condition is: dynamic 0.758 [0.735, 0.783], static 0.700 [0.673, 0.725], and shuffled 0.575 [0.552, 0.599]. On an average, participants showed the best performance in the dynamic condition, and the worst in the shuffled condition. The results were the following: dynamic – static: 0.075 [0.048, 0.101], static – shuffled: 0.093 [0.058, 0.123] before peak efficiency; and dynamic – static: 0.107 [0.082, 0.133]; static – shuffled: 0.172 [0.146, 0.195] after peak efficiency. The slopes of the step linear functions are: dynamic 0.0442 [0.0315, 0.0565], static 0.0442 [0.0311, 0.0577], shuffled 0.0530 [0.0381, 0.0692] before peak efficiency; and dynamic -0.0126 [-0.0164 , -0.0092], static -0.0175 [-0.0213 , -0.0133], shuffled -0.0190 [-0.0220 , 0.0163] after peak efficiency. The slope between age and accuracy is similar across all conditions before peak efficiency, whereas after peak efficiency, the dynamic condition shows the most gradual slope: static – dynamic: -0.0048 [-0.0104 , 0.0004]; shuffled – dynamic: -0.0064 [-0.0109 , -0.0018] (Figure 5.10).

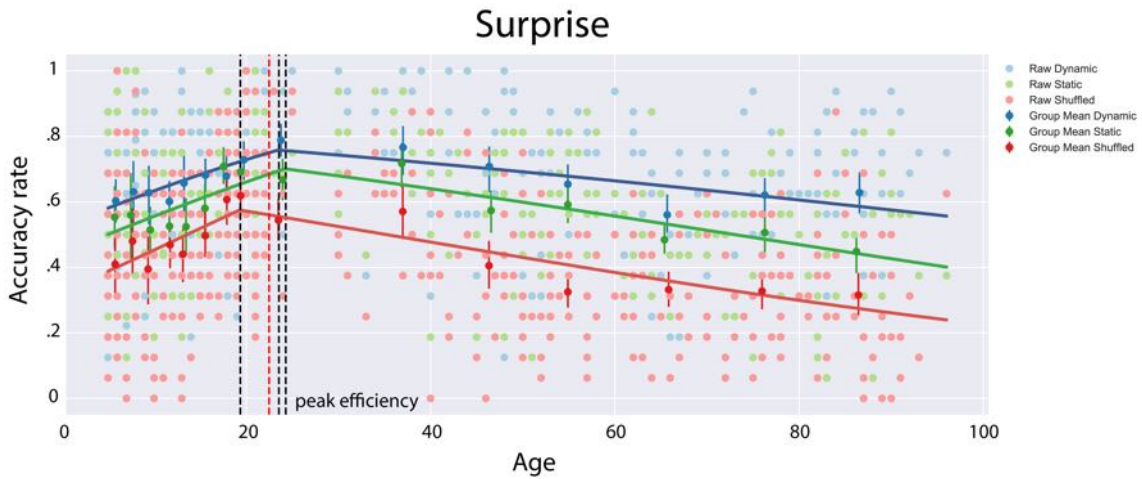


Figure 5.10. Surprise. The posterior model fit (solid line) for the expression of surprise with the individual performance (scatter plot) and the group average performance (dots with error bars) are given here. The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

5.5 Discussion

Our results present a fine-grained developmental tracking of human observers' ability to recognize the six basic emotions when presented with varying temporal properties: dynamic, static, and shuffled. Previous studies in the literature, examined expression recognition by using *arbitrary* age groups: ten years bins (Williams et al., 2009), stages of life (Horning, Cornwell, & Davis, 2012) or largely different age groups (e.g., 18-30, 58-70, Calder et al., 2003), while revealing either expression recognition improvement (Rodger et al., 2015) or decline (Calder et al., 2003; MacPherson et al., 2002; Malatesta, Izard, Culver, & Nicolich, 1987; Moreno, Borod, Welkowitz, & Alpert, 1993; Ruffman et al., 2008; Sullivan & Ruffman, 2004a). In contrast, our approach innovates by estimating the continuous developmental trajectory of facial expression recognition (from increase to decline), by considering age as a *continuum*, ranging from 5 to 96 years.

Using a Bayesian approach, we estimated for each condition and each expression individually the associated uncertainty and a) the peak efficiency, namely the point at which observers' recognition performance reaches its maximum before declining; b) the steepness of increase and decrease in recognition abilities; c) differences in the steepness of increase and decrease between different conditions (e.g., dynamic *vs.* static); and d) the overall processing advantage of the dynamic over the static and the shuffled stimuli. We will now discuss, in turn, each of these findings and their implications.

5.5.1 Recognition trajectory across development: increase, peak efficiency, and decrease

Our findings revealed unique developmental profiles and peak efficiency for the static, dynamic, and shuffled versions of each individual expression. Herein, we will focus on the dynamic and static trajectories and the differences between both conditions in a more detailed manner (i.e., static and dynamic). The results of the shuffled condition will be briefly considered at the end of the discussion.

Efficiency – increase

In both static and dynamic conditions, the sharpest rises in accuracy were observed for fear, followed by disgust and, to a lesser extent, surprise. These findings mirror the results of a previous developmental study that investigated the effects of age on the development of emotion processing in children, revealing that increasing age produced significant improvements in the recognition of fear and disgust (Herba & Phillips, 2004). We observed a more gradual increase for sadness and anger, but only in the static condition. Finally, we did not observe any increase for the expression of happiness, regardless of the experimental condition.

The steepest increase evidenced for fear might be accounted for by the very low recognition rates observed for this expression in young children, reaching only 13% in the 5–6 age group in the static condition (15% in the dynamic condition, Figure 5.11A). The expression of fear has been regularly reported in developmental (Herba & Phillips, 2004; Rodger et al., 2015; Widen, 2013), neuropsychological (Adolphs et al., 2003; Richoz et al., 2015), and behavioral studies (Calder et al., 2003) as being the most difficult expression to effectively recognize among all the expressions – a difficulty that is puzzling considering the evolutionary importance of an adequate and rapid categorization of this expression for survival. However, among the basic expressions, fear is probably the one that transmits the strongest multisensory perceptual cues. Multisensory and contextual information, such as environmental threats, may therefore play a crucial role in the decoding of this expression and be essential for an adequate categorization. Consistent with our findings, fear has been observed to display a sharp increase in some prior developmental studies (Herba, Landau, Russell, Ecker, & Phillips, 2006; Vicari, Reilly, Pasqualetti, Vizzotto, & Caltagirone, 2000), while other studies have revealed more gradual improvements (Gao & Maurer, 2009; Thomas, De Bellis, Graham, & LaBar, 2007) or stable, albeit low-task performance from early childhood to adulthood (Rodger et al., 2015). Differences across studies may be attributed to methodological considerations and task differences, as recognition rates have been proven to be task dependent (e.g., Montiroso, Peverelli, Frigerio, Crespi, & Borgatti, 2010; Vicari et al., 2000), with performance variations occurring even within the same study when the task is changed (Vicari et al., 2000). Importantly, the findings reported here provide further evidence that the recognition of fear has a special status within the framework of facial expression recognition (Richoz et al., 2015; Rodger et al., 2015).

Disgust also showed a steep increase in recognition accuracy, following a similar trajectory as fear. In line with our findings, steep improvements from childhood to adulthood were previously observed for disgust in a study by Rodger et al. (2015), which measured the quantity of information necessary for an observer to accurately recognize facial expressions, as well as in earlier studies that investigated expression recognition with matching (Herba et al., 2006) or labeling tasks (Vicari et al., 2000). As mentioned by Vicari et al. (2000), the steep improvement observed for disgust in children aged 5 to 10 may occur owing to the greater lexico-semantic abilities in older children. It might be also plausible that the very distinctive facial configurations of disgust convey signals about potentially contaminated food. These signals are crucial from an evolutionary perspective and hence the need to rapidly improve in the detection of this expression in order to stay away from harmful substances.

Finally, our findings also revealed a sharp increase for surprise, in both the static and dynamic conditions. Interestingly, the expression of surprise was already well recognized in

very young children aged 5 to 6, with recognition rates of 60% for the dynamic stimuli (55% for the static images, Figure 5.11A). High recognition rates in young children would rather accord with a more gradual developmental trajectory, as suggested by prior research that investigated the recognition of surprise from 5 up to 18 years of age (Rodger et al., 2015). Interestingly, however, the sharp increase observed for surprise in the current study may be accounted for by the very high recognition rates observed for this expression in participants above the age of 18, reaching up to 79% in the 21–30 age group for the dynamic stimuli (67% for the static images, Figure 5.11C).

We also observed a gradual increase for anger and sadness, although only in the static condition. These findings are generally consistent with previous reports (Herba et al., 2006; Rodger et al., 2015, for sad; Vicari et al., 2000). In the dynamic condition, children aged 5 to 6 were nearly as effective in recognizing anger (62%, Figure 5.11A) as young adults aged 17–18 (64%, Figure 5.11B), 21–30 (66%, Figure 5.11C), or 31–40 (66%, Figure 5.11D). The same pattern was observed for sadness, with identical recognition rates for young children aged 5–6 (56%) and young adults in the 17–18 age group.

Finally, our results did not reveal an increase for happiness in either condition, task performance remaining stable over ages, with our peak efficiency revealing the peak of the decline. The absence of improvement observed for happiness may be explained by the very high recognition accuracy already found in young children for this expression, which leaves little scope for improvement. Our findings for happiness are consistent with previous studies that revealed that children as young as five years of age recognize the expression of happiness just as effectively as adults (Gao & Maurer, 2009; Gross & Ballif, 1991; Herba & Phillips, 2004), even when the presentation time is as fast as 500 ms (Rodger et al., 2015). In order to capture the increase in recognition performance for happiness, we should have started earlier, below 5 years of age.

Additionally, it is worth noting that our findings revealed differences in the steepness of increase between the static and dynamic conditions for the expressions of disgust and sadness, the increase being steeper with the static stimuli. These findings might be accounted for by the low recognition rates found for the static images of disgust and sadness in very young children. An exposure to static images of disgust and sadness is rather uncommon in everyday life, particularly for young children, whereas an exposure to the dynamic versions of these expressions might be more frequent for children when their schoolfellows or siblings dislike some particular food they have to eat (disgust) or when they cry or express their sorrow (sadness).

Peak efficiency

The data-driven identification of the peak efficiency, the point at which observers' recognition performance reaches its optimum before declining, revealed a series of novel interesting findings. To the best of our knowledge, this is the first study that has effectively isolated the age at which observers are the most efficient for the recognition of the basic facial expressions of emotion across life span. We observed the earliest peak efficiencies for both the static and dynamic expressions of disgust (18–20 years) and fear (19–21 years) in young adults, followed by surprise (23–25 years) and sadness (27–28 years). Peak efficiency for the static expression of anger was found at 35 years of age, whereas the recognition of its dynamic version was reached at 39 years. The latest breakpoint that emerged from our data was observed for the dynamic expression of happiness at around 61 years of age (50 years for the static version). It is worth noting that the breakpoints found for each expression were nearly the same in both conditions, with the exception of anger and happiness, which reached their breakpoints later with dynamic expressions.

There are two explanations for the very early peak efficiencies found for fear and disgust. Firstly, as mentioned above, from an evolutionary perspective, these two expressions convey important signals about potential dangers or harmful substances, both important for survival. Disgust and fear can therefore be expected to reach their peak rapidly in order to ensure survival. Secondly, for fear and disgust, the point in time at which the peak efficiency emerges may be driven by the inherent properties of those expressions. Stimuli that are difficult to recognize for young people might be even more difficult for elderly people, as difficult tasks are likely to be more sensitive to cognitive decline (Calder et al., 2003; Ruffman et al., 2008). Changes in the slope of the lines may therefore be expected to occur earlier with difficult tasks. We examined response biases for each expression, computing confusion matrices across different age groups (see supplementary Figure 5.11A, B, C, D, E). The confusion matrices found for fear and disgust indeed revealed that these two expressions were particularly difficult for our observers to identify. Disgust was commonly confused with anger, with confusion rates ranging up to 28% in the 5–6 age group for the dynamic stimuli (Figure 5.11A), 20% for the 71–80 age group (Figure 5.11E). Previous studies also reported marked confusions between disgusted and angry faces, which were interpreted as a general bias towards angry faces (Recio, Schacht, & Sommer, 2013). Such a bias could explain the stable and high recognition rates found in the current research for angry faces from childhood onwards. Other confusions were observed between fear and surprise. In line with previous studies (Rodger et al., 2015), fear was found to be the most frequently confounded expression among all age groups, with confusion rates reaching up to 53% for the dynamic expression of surprise in the 5–6 age group (Figure 5.11A) or even 63% in the 71–80 age group (Figure 5.11E). As mentioned by Calder et al. (2003), age-related cognitive decline may reinforce these confusions due to perceptual or conceptual

difficulties (i.e., fear and surprise are conceptually very close and share facial signals that are morphologically similar, Delis et al., 2016). Note also that the reverse confusion was much less common. When presented with surprise, the confusion rates observed for fear reached only 3% for the dynamic stimuli (4.6% for the static expressions) in the 5–6 age group and 5% (3.3% for the static expressions) in elderly people aged 71–80.

Interestingly, our findings also revealed a later emergence of the peak efficiency for anger compared to the other expressions. As mentioned before, recognition abilities for anger showed no increase in the dynamic condition, task performance being already high in young children, and displayed only a slight increase for the static condition, recognition rates being also high in young children. A potential while speculative explanation for this observation may lie in the fact that we are daily exposed to the expression of anger, arguing with our partners, children, colleagues – an exposure that might postpone the recognition decrease of this expression and, therefore, the changes in the slope of the line.

Finally, the latest breakpoint found for happiness may be accounted for by the ceiling effect found for this expression from childhood onwards.

Altogether, this second set of findings offers novel insights into the development of human facial expression recognition. As observed, facial expression recognition develops following *emotion-dependent* trajectories that do not necessarily all reach their peak efficiency in early adulthood as predicted by previous studies (Calder et al., 2003; De Sonneville et al., 2002; Horning et al., 2012; Williams et al., 2009). The optimal level of task performance can indeed be reached at a very late point in development, depending also on the very nature of the diagnostic information of the facial expression, its temporal properties and evolutionary value.

Efficiency – decrease

Finally, we observed differences in the steepness of decrease in recognition performance across emotions and conditions. In the *dynamic* condition, the steepest decreases were observed for anger, happiness, and sadness, and less severe decreases for fear and surprise. Disgust showed the least severe decrease in this condition. Different patterns were observed in the *static* condition, the steepest decline being for happiness, followed by sadness and disgust. Less severe decreases were found for fear and anger, whereas the least severe decrease was observed for the expression of surprise. Similarly to the differences in the steepness of increase observed between static and dynamic conditions, differences in the steepness of decrease were observed between both conditions for the expression of disgust. The recognition of the static expression of disgust decreased from 51% to 34% between the ages of 61-70 and 81-90, whereas recognition accuracy of its dynamic version remained relatively stable (decrease from 55% to 52%, Figure 11E).

This pattern of results posits that the recognition of facial expression declines over time, which is consistent with previous models of aging. These models suggest that age-related structural changes in different brain regions, particularly in frontal and temporal volumes, as well as changes in neurotransmitters (Calder et al., 2003; Ruffman et al., 2008) might be responsible for older adults' impairment in the recognition of facial expression. For example, the amygdala, which plays a crucial role in the processing of fear and sadness (e.g., Adolphs et al., 2005; Yang et al., 2002), undergoes severe atrophy with age and becomes progressively less responsive to negative stimuli (De Winter et al., 2016; Mather et al., 2004; Ruffman et al., 2008). In contrast, the insula and basal ganglia, which underlie the processing of disgust, seem to be less vulnerable to aging, as evidenced by the preserved ability to recognize this expression in older adults (Calder et al., 2003; Horning et al., 2012; Ruffman et al., 2008). Interestingly, our findings revealed the least severe decrease in recognition performance for the dynamic expression of disgust. However, in contrast to previous studies that showed no reduction in the recognition of some expressions or even some improvements with increasing age (e.g., Calder et al., 2003), our findings revealed steep to moderate decreases for all the expressions, even for disgust being usually preserved in elderly people (Calder et al., 2003; Horning et al., 2012; Ruffman et al., 2008).

Methodological considerations may be responsible for the differences observed between the current study and previous ones. Indeed, previous studies investigated facial expression recognition across groups of ages (Calder et al., 2003), stages of life (Horning et al., 2012), or decades (Williams et al., 2009), whereas our study investigated elderly people's ability to categorize emotions by considering age as a continuum. This methodological approach overcomes the problem of defining arbitrary age boundaries, which are routinely used in the literature to relate to critical developmental ages.

Furthermore, the variability in findings between the current research and previous neuropsychological and behavioral studies can be accounted for by the age-ranges tested across the studies. For example, Calder et al. (2003) observed improved recognition abilities for disgust in their older adult age group, spanning from age 58 to 70 (mean age 65). In contrast, in our study, we tested participants up to the age of 96, giving rise to the possibility that the decline for disgust appears at a later point in development. This assumption is in line with a previous study that showed a decrease in the recognition of disgust in elderly people, aged 80 to 91 (Williams et al., 2009).

Additionally, the stimuli used across the different studies might also have impacted expression recognition performance. In the current study, we used a specific database of emotional expressions that are less prototypical than the Ekman and Friesen (1976) standard

set of facial photographs used in previous research (Calder et al., 2003; McDowell, Harrison, & Demaree, 1994; Sullivan & Ruffman, 2004b). Moreover, in contrast to previous reports that used only static images displaying the apex or the highest state of an emotional expression (Calder et al., 2003; Ruffman et al., 2008), we tested facial expression recognition with static, dynamic, and shuffled stimuli. Importantly, our stimuli were controlled for the amount of *low-level* discriminative information carried over time. In other words, the quantity of *low-level* information carried by our static, dynamic, and shuffled stimuli was identical across conditions and tasks (Gold et al., 2013). In line with previous studies (Krendl & Ambady, 2010; Sze et al., 2012), we found that elderly people were impaired in recognizing static but not dynamic expressions. However, in contrast to the findings reported by Krendl and Ambady (2010), we also observed steep to moderate declines for all the expressions, even in the dynamic condition. However, in their study participants were provided with additional aiding cues, such as body-related information or contextual cues, which might have facilitated expression recognition, given that the perception of a particular expression is strongly influenced by the context in which it occurs (Barrett & Kensinger, 2010; Horning et al., 2012). For example, Aviezer et al. (2008) found more consistent recognition performance for fear when person-related or contextual information were provided to the participants.

Finally, the divergence between our findings and those of previous research may also be due to the small number of trials presented (Horning et al., 2012; Moreno et al., 1993) as well as the differences in the settings used, with some studies relying on laboratory settings (Calder et al., 2003; Horning et al., 2012) and others on online tasks (Williams et al., 2009).

5.5.2 Static versus dynamic expressions

A dynamic advantage before peak efficiency

Our findings revealed a dynamic face advantage for the recognition of anger, surprise, and happiness before peak efficiency. These results are inconsistent with previous developmental studies, which revealed that dynamic presentations did not increase children's recognition performance (Nelson et al., 2013; Nelson & Russell, 2011; Widen & Russell, 2015) and with some experiments showing even an overall advantage for static expressions (Nelson & Russell, 2011, Study 1; Widen & Russell, 2015). Such advantage also differs from the results reported by previous studies in young and healthy adults (e.g., Christie & Bruce, 1998; Jiang et al., 2014; Kättsyri & Sams, 2008), showing that the recognition of facial expressions is not facilitated by the dynamic information provided by moving faces.

The lack of consistency between these studies and the present work may be accounted for by methodological factors. For instance, in some of the aforementioned developmental

studies, only a single actor was selected to record the facial expressions (Nelson et al., 2013; Nelson & Russell, 2011), raising the possibility that the results found could be biased by the acting performance. Compared to the Ekman and Friesen (1976) standard set of facial expressions, the expressions of the single actor used in the study by Nelson and Russell (2011) were indeed more readily labeled as correct by adults, as they were perceived as clearer and more intense. Asking children to categorize facial expressions of a single actor in their dynamic and static forms might also have impacted their recognition performance since they might have been more likely to choose the same label in both the conditions, by using a picture matching strategy. In addition, compared to the current research in which children were asked to choose the correct answer among six possibilities, previous developmental studies used free labeling as a measure of recognition (Nelson & Russell, 2011; Widen & Russell, 2015), which raises the possibility that vocabulary performance rather than children's true ability to understand the emotions of others were tested.

Importantly, most developmental studies that revealed an overall static advantage for facial expression recognition (Nelson & Russell, 2011; Widen & Russell, 2015), directly compared children's performance for static expressions to their scores with dynamic expressions. In most cases, these direct comparisons can be problematic because they make it difficult to determine whether increased recognition rates are caused by psychological or physical factors (Gold et al., 2013). For instance, Nelson and Russell (2011) and Widen and Russell (2015) created their static images by presenting a single frame of the highest amplitude of the dynamic sequences, a procedure that might have created "optimal" static images. The overall static advantage found in their research may be due to an increased quantity of discriminative information provided by the stimuli rather than an enhanced psychological ability to perceive the static expressions. In order to control for this general confounding of physical and psychological factors, we decided to rely on a database of stimuli created by Gold et al. (2013), who controlled for the amount of *low-level* information carried by their stimuli over time by carefully dissociating these two factors with the use of a psychophysical approach. Compared to previous studies (Nelson et al., 2013; Nelson & Russell, 2011; Widen & Russell, 2015), our results, therefore, offer a more reliable view and a better understanding of the way in which temporal properties influence facial expression recognition from childhood onwards.

A dynamic advantage after peak efficiency

Our results revealed processing benefits of dynamic stimuli after peak efficiency for all the expressions, with the exception of sadness and fear. Interestingly, our data evidenced that these results were driven by a suboptimal performance for the recognition of static expressions in elderly people rather than increased abilities to recognize dynamic expressions (see Supplementary Table 5.1 for the example of surprise).

In everyday life, facial expressions are dynamic events that unfold over time in some particular ways, representing a richer and more valid approach to study facial expression recognition. Previous fMRI studies have also suggested that different neural substrates underlie the processing of dynamic and static expressions (e.g., Johnston et al., 2013; Kessler et al., 2011; LaBar, Crupain, Voyvodic, & McCarthy, 2003; Paulmann et al., 2009; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Schultz & Pilz, 2009; Trautmann et al., 2009). Dynamic faces have been found to selectively elicit higher neural responses in the pSTS, in the anterior STS, and in the inferior frontal gyrus (Bernstein, Erez, Blank, & Yovel, 2017; Fox, Iaria, & Barton, 2009; Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011). A very recent fMRI study that used multi-voxel pattern analysis revealed that dynamic expressions were associated with increased activation in both face-selective and motion-selective areas as well as higher categorization accuracies, compared to static expressions (Liang et al., 2017). Given that dynamic faces elicit higher neural responses (Bernstein et al., 2017; Fox et al., 2009; Pitcher et al., 2011) and cause the activation of a wider network of regions in the brain (Arsalidou, Morris, & Taylor, 2011; Liang et al., 2017), their decoding may be less vulnerable to age-related degeneration compared to the decoding of static images.

In contrast, suboptimal performance for static stimuli could be explained by age-related structural changes in brain regions responsible for the processing of static emotional expressions. For example, De Winter et al. (2016) recently evidenced that age-induced atrophy to the amygdala of patients with frontotemporal dementia affected emotion processing in distant face-selective areas. More specifically, their findings evidenced a positive correlation between grey matter volume in the left amygdala and emotion-related brain activity in the fusiform face area, a core region in the face processing network involved in the decoding of static stimuli (Pitcher et al., 2011) and emotional expressions (Ganel, Valyear, Goshen-Gottstein, & Goodale, 2005; Xu & Biederman, 2010).

Dynamic faces also include information that cannot be completely rendered by static images, forcing the observers to shift their attention to different facial features. Multiple shifts on different facial areas are likely to benefit expression recognition, given that the facial signals critical for the recognition of emotional expressions can be found throughout the face. In clinical conditions or normal aging populations, when slower or suboptimal processing takes place, dynamic stimuli may provide additional cues, attracting and holding attention as well as enhancing motor simulations. The increased attention inherently elicited by moving faces may compensate for the apparent age-related deficits found in elderly populations on expression recognition tasks with static images. Dynamic face stimuli may naturally drive the focus of attention towards the diagnostic information in a bottom-up fashion (i.e., the mouth for

surprise), whereas static face images require the observers to move towards those features based on top-down internal representations.

It is also important to note that the advantage we observed for dynamic expressions cannot simply be attributed to an overall larger amount of discriminative information carried by the dynamic stimuli. As reported above, the stimuli used in the current experiment were created by Gold et al. (2013), who used an ideal observer approach to effectively measure the amount of information provided by the stimuli. Gold et al.'s (2013) findings revealed that their dynamic stimuli did not offer more discriminative information to the observers as compared to their static images. Thus, the dynamic advantage for anger, disgust, happiness, and surprise observed in our participants is unlikely to be the result of physical factors. Rather, this dynamic advantage most probably comes from an adequate ability to use the available perceptual and diagnostic information.

We did not find a dynamic advantage for fear and sadness, neither before nor after peak efficiency. From a sociobiological perspective, the expression of fear is critical for human survival (LoBue, 2010) and it has been shown to potentiate early visual processing of perceptual events (e.g., Phelps, Ling, & Carrasco, 2006) and enhance attention (e.g., Carlson & Mujica-Parodi, 2015; Pourtois, de Gelder, Bol, & Crommelinck, 2005). Interestingly, using a single-trial Repetition Suppression (stRS) approach, we very recently revealed that this expression boosts the early coding of individual faces, regardless of attentional constraints (Turano et al., 2017). Similar findings were reported in a very recent developmental study that examined detection thresholds for happy and fearful faces presented with noise. The superior ability for detecting fearful faces was observed already in infants aged 3.5 months (Bayet et al., 2017). Our brain may be particularly tuned to detect this expression, regardless of its temporal properties. This assumption may explain the absence of a dynamic advantage for the decoding of this expression. However, enhanced processing of fear would also predict increased categorization performance, a prediction that is inconsistent with our findings. Our results indeed revealed very low recognition rates throughout the life span. As mentioned above, among all expressions, fear is probably the most powerful for transmitting multisensory information. Broader contextual information may therefore be necessary to reliably categorize it. This assumption is in line with an emerging literature that suggests that isolated facial signals may not be sufficient for observers to adequately perceive the emotions of fear and disgust, and that additional information regarding the context in which the expression occurs is critical (e.g., Barrett & Kensinger, 2010).

The absence of a dynamic advantage for the processing of sadness is consistent with previous findings, which revealed that the expression of sadness is better recognized through

static pictures (Bould et al., 2008; Recio et al., 2013; Widen & Russell, 2015) or when evolving slowly (Kamachi et al., 2001; Recio et al., 2013). Ekman (2003) suggested that among all the expressions, sadness is the one lasting the longest over time, a property that may explain why slowness or stillness may increase recognition performance. Our results further confirm that the idiosyncratic properties of this expression are inherently slow.

We should acknowledge that we did not assess whether elderly people's cognitive abilities influenced recognition performance. Fluid intelligence (e.g., Horning et al., 2012; Sullivan & Ruffman, 2004a), processing speed (e.g., Orgeta & Phillips, 2007), verbal memory (e.g., MacPherson et al., 2002), or discrimination of visual information (Mill et al., 2009) are all cognitive faculties that are critical for the recognition of human facial expressions and have been found to decrease with increasing age (e.g., Mill et al., 2009; Salthouse, 2004). Because our stimuli were only presented for 1 s, reduced processing speed in elderly people may have influenced their recognition performance for static face images, as the number of facial features extracted from the faces in this limited presentation time might have been lower than that in younger adults. Interestingly, a recent cross-sectional study that examined the influence of different cognitive abilities on facial expression recognition observed that these faculties contributed to the performance but did not fully account for the impairments observed in older adults (Horning et al., 2012). Additionally, in a later study, Zhao et al. (2016) observed that the slower processing speed in elderly people was not responsible for facial expression recognition deficits (Zhao et al., 2016; but see Suzuki & Akiyama, 2013; West et al., 2012).

Note also that differences in recognition abilities could stem from differences in the cohorts that were tested, such as educational shifts, cultural norms, or social differences. To the best of our knowledge, no prior research has ever examined the extent to which the recognition of facial expression is influenced by these cohort effects (Ruffman et al., 2008).

The shuffled condition

We introduced this condition to fully replicate the study conducted by Gold et al. (2013), without having clear predictions for this experimental condition. Our results revealed similar developmental trajectories for the six basic expressions in the shuffled condition (i.e., increase, peak efficiency and decrease) although recognition rates were generally lower than those observed in the other conditions. More specifically, our findings revealed a recognition advantage for the dynamic expressions of anger, disgust, happiness and surprise over the shuffled ones and a recognition advantage for the static expressions of happiness, disgust and surprise over the shuffled ones. In contrast, we observed better recognition performance for the expression of fear in the shuffled compared to the static or dynamic conditions. As reported by our participants, this advantage for fear could be accounted for by the properties of the stimuli

themselves. The shuffled expressions were generated by temporally randomizing the frames of the dynamic movies. This procedure leads to the impression that the actors performing the emotional expressions are shaking giving the feeling that they are afraid.

Interestingly, differences in recognition performance across conditions are inconsistent with the results reported by Gold et al. (2013) who observed similar performance in all three conditions. In that prior study however, the authors did not consider the recognition rates of the individual expressions effectively, collapsing them across the six expressions in each condition. Our findings offer therefore new evidence that the temporal progression of information (i.e., the order of the frames) provided by genuine natural expressions is more important for the recognition of some expressions (e.g., anger, disgust happiness, surprise) than others (e.g., fear, sad). Given the very non-ecological nature of the stimuli, we will not further discuss these results as their contribution is limited in a theoretical point of view.

Methodological considerations

In the current study, we used a hierarchical Bayesian model with weakly informative priors. The flexibility and power of the Bayesian approach in dealing with time series data and building nonlinear models was also recently demonstrated in emotion research (e.g., see Krone, Albers, Kuppens, & Timmerman, 2017, for an application in personal emotion dynamics) thanks to the rapid development in probabilistic language programming. It is worth noting, that there are alternative candidates models that could capture the nonlinear relationship between age and some psychological or behavioral measurements, including Latent Growth Curve Models (for an introductory text, see Duncan, Duncan, & Strycker, 2013), Generalized Additive Mixed Models including spline regression, or Quadratic Linear Mixed-Effect Models (Wood, 2006). Some of these models have been previously applied to investigate similar questions, such as the estimation of the peak efficiency of diverse recognition abilities (i.e., change point estimation, e.g., Cohen, 2012; Cudeck & Klebe, 2002). The Bayesian modeling framework we used here provides a coherent mathematical language to describe our model and assumption, while giving the flexibility to potentially extend part of the component to build more complex models. Moreover, it allowed us to properly quantify the uncertainty and regulate the estimation across different conditions using hyper-priors.

To identify inverse “U-shape” patterns, as those observed in our study, previous modeling methods occasionally involved the testing of a quadratic relationship (e.g., a significant regression coefficient of age^2), even if such practice is not always valid (Simonsohn, 2017). Instead, Simonsohn (2017) suggested to fit two separate linear models and compare the coefficients of the two slopes as a more valid alternative. While our model is conceptually similar to Simonsohn (2017)’s model, there are two major differences. First, the inference proposed by

Simonsohn (2017) involved multiple model fitting steps, by initially identifying the breakpoint (i.e., peak efficiency in our case), and then estimating the coefficients of the two linear functions. In contrast, with a full model that jointly estimates the breakpoint and the linear functions, we could better estimate the parameters and quantify the associated uncertainty. Second, the intercepts of our step linear function are linked and represented as one value (i.e., the recognition ability at peak efficiency), whereas in Simonsohn (2017)'s model the two linear functions are not linked. The linked linear function is more appropriate in our case, as it is unlikely to have a sudden increase or decrease in recognition ability in a short span during natural development. Nonetheless, an implicit yet important assumption present in both models is that the peak efficiency is found somewhere in the middle of the life span (or more precisely, not at either of the two extrema). Indeed, if the peak efficiency is at the lower or upper limit (e.g., too young or too old), the parameter estimation may not be accurate.

Our model estimation performed well, except for the expression of happiness, because of the ceiling effect we observed for this expression. The divergence in the trace and the multimodal in the posterior distribution of the peak efficiency both indicate that the current step linear model is not the best suited to represent changes in recognition abilities across the life span for this expression. Currently, all the expressions are estimated independently. While modeling this way is easier to interpret, we ignored the random effect in the subjects across the expressions. Future studies are necessary to take into account the random effect from each subject (intercept and slope). This could be done by directly modeling the full confusion matrix from each subject (instead of only looking at the diagonal in the current study), presumably with some matrix decomposition trick or a Dirichlet-Categorical model.

Finally, our model allowed us to estimate the overall advantage of one condition over another before and after the peak efficiency. However, since we decided to consider the age as a continuum and not rely on specific age groups on the basis of arbitrary boundaries, our model did not allow us to finely estimate at which precise age the dynamic advantage emerges or disappears.

5.6 Conclusions

Current knowledge about facial expression recognition primarily arises from studies that use static images. In our daily life, however, natural faces are dynamic; they evolve over time in some particular ways to convey crucial information for adapted social behaviors. Prior studies investigating the importance of dynamic cues for the processing of facial expressions have yielded equivocal results, with some studies suggesting that dynamic expressions are more readily recognizable than static images and others suggesting that they are not. In order to

clarify these results and to determine if age is a critical factor to account for such discrepancies, we conducted a large-cross-sectional study to investigate the recognition of facial expressions by participants aged 5 to 96. Over 400 observers were instructed to categorize static, dynamic, and shuffled expressions according to the six basic expressions. Our findings revealed that regardless of the age of the observer or temporal condition, happiness was the best recognized facial expression, whereas fear was the most difficult to effectively categorize, as this expression was commonly confused with surprise. Bayesian modelling allowed us to quantify the steepness of increase and decrease in performance for each individual expression in each condition. Our results also revealed a data-driven estimation of the peak efficiency for every expression and finally provided new evidence for a dynamic advantage for facial expression recognition, stronger for some expressions than others and more important around specific points in the life course. Notably, performance for static images was less effective in the elderly population. Altogether, our findings highlight the importance of using ecologically valid faces in exploring the recognition of facial expressions and invite caution while drawing conclusions from studies that use only static images to this aim.

5.7 Acknowledgements

We would like to thank Prof. Jason Gold and his colleagues from the Indiana University, Bloomington, for providing us with the stimuli used in this study. Very special thanks go to all our participants, especially the children and adolescents who participated in our study from the following schools: Ecole Enfantine et Primaire de Marly Grand Pré and Cycle d'Orientation de Domdidier in the area of Fribourg, Switzerland. We would also like to thank all the teachers for their patience and help, especially the head teachers, Claude Meuwly from Marly and Chantal Vienny Guerry from Domdidier as well as technician Zvonko Traykoski for meticulously organizing our visits and ensuring that everything worked well.

We would also like to express our gratitude to the retirement homes Foyer de Bouleyres and Maison Bourgeoisiale in Bulle and Foyer St-Martin in Cottens, especially to Christian Rime, Véronique Castella, Isabelle Montagnon, and Philippe Bourquin for letting us test their residents. We would also like to thank Maria Teresa Turano and Prof. Maria Pia Viggiano from Florence, Léa Poitrine, Claudia Wyler, Qendresa Shkodra, Vanessa Ferrari, Pauline Rotztetter, Pauline Schaller, Linda Pigozzi, Lauriane Beffa, Hugo Najberg, Christel Aichele, and Martina Studer for their precious help with testing. This study was supported by grant F14/06 from the Rectors' Conference of Swiss Universities (CRUS).

5.8 Supplementary Figures and Table

5-6 age group



7-8 age group



9-10 age group



Figure 5.11A. Confusion matrices – Response classification errors. Each row displays one of the six presented facial expressions, while each column shows the average frequency of the response given by the observers (Null indicates a "I don't know" response). For example, in the 5-6 age group, when presented with fear, the confusion rates observed for surprise reached up to 53% in the dynamic condition, 44% in the static, and 37% in the shuffled condition. There is a correspondence between the colourmap and the numbers displayed in the matrix, with dark blue tones indicating low frequency while blue-to-green shades indicate high frequency. The values in the main diagonal indicate the recognition performance for each expression.

11-12 age group



13-14 age group



15-16 age group



Figure 5.11B. Confusion matrices – Response classification errors. Each row displays one of the six presented facial expressions, while each column shows the average frequency of the response given by the observers (Null indicates a “I don’t know” response). There is a correspondence between the colourmap and the numbers displayed in the matrix, with dark blue tones indicating low frequency while blue-to-green shades indicate high frequency. The values in the main diagonal indicate the recognition performance for each expression.

17-18 age group**19-20 age group****21-30 age group**

Figure 5.11C. Confusion matrices – Response classification errors. Each row displays one of the six presented facial expressions, while each column shows the average frequency of the response given by the observers (Null indicates a "I don't know" response). There is a correspondence between the colourmap and the numbers displayed in the matrix, with dark blue tones indicating low frequency while blue-to-green shades indicate high frequency. The values in the main diagonal indicate the recognition performance for each expression.

31-40 age group



41-50 age group



51-60 age group



Figure 5.11D. Confusion matrices - Response classification errors. Each row displays one of the six presented facial expressions, while each column shows the average frequency of the response given by the observers (Null indicates a "I don't know" response). There is a correspondence between the colourmap and the numbers displayed in the matrix, with dark blue tones indicating low frequency while blue-to-green shades indicate high frequency. The values in the main diagonal indicate the recognition performance for each expression.

61–70 age group**71–80 age group****81–90 age group**

Figure 5.11E. Confusion matrices - Response classification errors. Each row displays one of the six presented facial expressions, while each column shows the average frequency of the response given by the observers (Null indicates a "I don't know" response).

There is a correspondence between the colourmap and the numbers displayed in the matrix, with dark blue tones indicating low frequency while blue-to-green shades indicate high frequency. The values in the main diagonal indicate the recognition performance for each expression

Table 5.1. Recognition performance for surprise between the dynamic and static conditions at peak efficiency and above the age of 80.

	Condition effect		
	Dynamic	Static	(Δ = Dynamic - static)
PE	75.8%	70%	5.8%
> 80	62.8%	45%	17.8%
Age effect on the conditions (Δ=PE - >80)			
	13%	25%	

Note. PE = recognition performance at peak efficiency; > 80 = recognition performance for all the observers above the age of 80 (N = 41).

As shown in Table 5.1, we observed an overall advantage for the processing of the dynamic facial expression of surprise over the static one. This advantage was even more marked for the above 80-year-old observers (Δ Dynamic - static). Importantly, the recognition performance of the dynamic expression of surprise decreased from 75.8% at peak efficiency to 62.8% after the age of 80 (Δ = 13%). For the static expression of surprise, the recognition performance dropped from 70% at peak efficiency to 45% after the age of 80 (Δ = 25%). When the performance of the observers above the age of 80 is compared to the performance at the peak efficiency (Δ PE - >80), then the difference in the static condition is nearly twice as large that in the dynamic condition. This pattern of results favors the view that the dynamic advantage for the recognition of facial expressions of emotion is driven by a suboptimal performance for static stimuli.

5.9 Supplementary Results

The experiment script, raw data, and analysis codes are open to access on Github (https://github.com/iBMLab/Static_dynamic).

CHAPTER 6

GENERAL DISCUSSION

6.1 Main findings

The aim of this thesis was to use dynamic faces to investigate different aspects of face processing in different populations and age groups. More precisely, our work examined three questions: (1) Is the extraction of gender from *dynamic* faces and voices influenced by the nature of speech signal in 6-, 9-, and 12-month-old infants? (2) Are the *dynamic* internal representations of facial expressions in a prosopagnosic patient similar to those in healthy controls? Further, are these *dynamic* expressions better recognized by patient PS than *static* images? (3) Besides prosopagnosic individuals, do other populations with immature or fragile face processing systems also benefit from the presentation of *dynamic* emotional signals? The following sections will summarize our main findings in order to answer each of these questions.

Is the extraction of gender from *dynamic* faces and voices influenced by the nature of speech signal in 6-, 9-, and 12-month-old infants?

In the first experimental study, we used movies of dynamic faces in combination with short stories told in infant- or adult-directed speech to investigate infants' ability to perceive multisensory gender coherence. Our findings revealed that by the middle of the first year of life, infants displayed the ability to match female faces and voices when the stories were told in adult-directed speech. This ability emerged later, between 6 and 9 months of age, when the stories were told in infant-directed speech and restricted to female faces. Altogether, our results suggest that the speech manner strongly influences infants' ability to reliably integrate gender information in audio-visual communication. They indicate that 6-month-old infants may be more likely to match audible and visible gender attributes when looking at adults communicating with each other, than when adults are directly speaking to them. Overall, these findings suggest that the early multisensory ability to perceive gender is shaped by the very nature of social interactions (Richoz et al., 2017).

Are the *dynamic* internal representations of facial expressions in a prosopagnosic patient similar to those in healthy controls? Further, are these *dynamic* expressions better recognized by patient PS than *static* images?

In the second experimental contribution, we reported a single-case study of the patient PS suffering from later-life acquired prosopagnosia. By using a novel 4D technique, we examined whether the dynamic internal representations of the six basic expressions in patient PS were similar to those in healthy controls. We did so to shed further light on the widely debated question as to whether emotional and identity judgments rely on common or distinct representational systems. In a previous study, Caldara et al. (2005) demonstrated that patient

PS used facial information in a suboptimal manner when asked to recognize familiar faces. Here, in contrast, we found that she was within the typical range of controls, using *all* facial features when instructed to categorize *dynamic* emotional expressions. Moreover, our findings revealed that patient PS was selectively impaired in categorizing static expressions, while her ability to categorize dynamic expressions was preserved (with the exception of fear). Our findings cannot assert an independence between the extraction of face information related to identity and expression, as patient PS was selectively impaired in categorizing static emotional expressions. However, our findings support the idea that the face system relies on distinct representational systems to extract information from static (identity and static expressions) and dynamic faces or at least dissociable cortical pathways to access this information. Our findings also emphasize the importance of research featuring dynamic displays in the field of emotion recognition, particularly in the realm of brain-damaged or clinical conditions (Richoz et al., 2015).

Besides prosopagnosic individuals, do other populations with immature (i.e., young children) or fragile (i.e., elderly adults) face processing systems also benefit from the presentation of *dynamic* emotional signals?

Following on from our previous neuropsychological findings on a single-case of acquired prosopagnosia (**CHAPTER 4**), we decided to conduct another study to investigate whether dynamic signals offer processing benefits in other populations with fragile face processing systems, namely elderly people or young children. Our findings provide new evidence for a dynamic advantage for facial expression recognition, stronger for some expressions than others, and more important in young and elderly participants. In elderly adults, the dynamic advantage was driven by suboptimal performance for the recognition of static emotional expressions. Critically, this suboptimal performance could be used as a marker for impaired face processing associated with other aspects of general cognitive decline. Our results highlight again the importance of using dynamic stimuli when investigating the recognition of facial expressions in specific populations, inviting to caution when interpreting the results obtained from studies that only use static images.

6.2 Discussion

6.2.1 A key involvement of dorsal face-selective brain regions

The three studies presented in this thesis involved the processing of dynamic and multimodal face information. At the neuro-functional level, the ability we observed in processing such information from faces in our *different* experimental populations, might critically relate to a *common* involvement of the dorsal face-selective brain regions.

Recent fMRI studies investigating the role played by dynamic properties in the processing of faces have revealed that the dorsal pathway including the posterior Superior Temporal Sulcus (pSTS), the anterior part of the Superior Temporal Sulcus (aSTS) as well as the Inferior Frontal Gyrus (IFG), elicits stronger responses to dynamic as compared to static faces. In marked contrast, face-selective areas located in the ventral stream do not exhibit a preference for moving stimuli (Bernstein et al., 2017; Fox et al., 2009a; Pitcher et al., 2011). In our second study, we evidenced that PS, a single-case of acquired prosopagnosia, showed a selective impairment in the categorization of static expressions, whereas her performance with dynamic stimuli was within normal range. Based on the fMRI evidence mentioned above, PS's advantage for the categorization of dynamic facial expressions might relate to her functionally intact dorsal pathway, while her selective impairment with static images could be accounted for by her lesions in the ventral stream. Although our data provide further evidence for a direct cortical pathway connecting early visual areas to the pSTS, a future neuroimaging study with PS is necessary to support this assumption and isolate the brain regions underlying her performance.

In a similar way, in our third study, dorsal face-selective brain regions probably played a key role in elderly people's perseveration in the recognition of dynamic expressions. Our findings indeed revealed that the dynamic advantage for the recognition of anger, disgust, happiness, and surprise was driven by suboptimal performance with static images. As dynamic stimuli elicit elevated neural responses in the dorsal stream (Bernstein et al., 2017; Fox et al., 2009a; Pitcher et al., 2011) and cause the activation of a wider network of regions in the brain (e.g., Arsalidou et al., 2011; Liang et al., 2017), their decoding might be less vulnerable to age-related degeneration as compared to the decoding of static images. However, this assumption again needs to be further investigated through neuroimaging studies that would directly tackle this issue.

Dorsal face-selective brain regions do not only play a key role in the processing of dynamic faces but also in the integration of multimodal person-related information. By using

fMRI, Watson and colleagues (2014) recently reported that an important region spreading down the trunk of the right STS is selectively activated by both faces and voices. Similar findings were reported in a later study by Hasan and colleagues (2016), who investigated the integration of identity information from auditory and visual cues. These results suggest that dorsal face-selective brain regions play a crucial role in observers' ability to integrate information from multiple sensory sources. In light of these findings, a question related to our first study spontaneously arises: does the STS of young infants already display sensitivity to multimodal information?

Previous studies, which investigated the maturation of the cortical network for face perception, have frequently reported the STS as being the latest region among face-selective areas to reach levels of maturity comparable to those seen in adults (Cohen-Kadosh, 2011; Leroy et al., 2011). To the best of our knowledge, however, prior developmental fMRI research did not address the question as to whether the STS is involved in the integration of multimodal information in infants. As fMRI studies in the pediatric field are particularly challenging due to head movements, dropouts, poor performance, or increased noise (Cohen-Kadosh, 2011), such a study could be conducted using fNIRS. fNIRS is a growing non-invasive neuroimaging technique that allows the localization of brain activity in specific cortical regions by monitoring relative changes in oxy-, deoxy-, and total-hemoglobin concentrations. The human brain undergoes a number of physiological changes as it responds to environmental stimuli. These changes in blood levels and electrochemical activity also affect its optical properties. fNIRS captures cortical brain responses by measuring the transmission and absorption of near-infrared light through cerebral tissues. Although the spatial resolution of fNIRS is lower than that of fMRI, this technique has numerous unparalleled advantages. As demonstrated in previous developmental research, fNIRS is a flexible technique that offers new and interesting insights into the development of the human brain. Because of its flexibility to head movements, blinking, or other artefacts, fNIRS is particularly well suited to investigate the functional emergence of the face-selective cortical network in infants or young children (e.g., Nakato et al., 2011).

6.2.2 Our findings in the light of revised frameworks of human face processing

Based on TMS research (Pitcher, 2014; Pitcher et al., 2014), fMRI studies (Fox et al., 2009a; Pitcher et al., 2011), and recent functional connectivity evidence (Avidan et al., 2014; Gschwind et al., 2012; Pyles et al., 2013), Duchaine and Yovel (2015) have proposed a revised framework of the face processing network, which suggests that dorsal face areas are specifically tuned to the processing of facial motion, while the ventral regions are preferentially engaged in the processing of form information (see section 1.3.3, Figure 1.48). Inconsistent with the influential

neural model of face processing proposed by Haxby et al. (2000), Bernstein and colleagues (2017) have further evidenced that the OFA and FFA, although insensitive to motion information, display similar sensitivity to both changeable and invariant aspects of faces. In contrast, the pSTS elicits a stronger sensitivity to motion information and changeable aspects of faces. The revised neural framework proposed by Duchaine and Yovel (2015) provides some interesting explanations for the results evidenced in the second experimental contribution of this thesis. PS's preserved ability to recognize dynamic expressions may be related to the functionally intact dorsal face-selective brain regions that connect early visual areas to the pSTS and aSTS.

Duchaine and Yovel (2015) further suggested that due to its broad sensitivity to form information, the FFA may also contribute to the processing of emotional expressions, at least in the case of static faces (see also Bernstein et al., 2017; Cohen-Kadosh et al., 2010; Dalrymple et al., 2011; Fox et al., 2009b; Ganel et al., 2005; Xu & Biederman, 2010). Considering the fact that the FFA is also anatomically intact in PS (Rossion, 2008; Sorger et al., 2007), one might expect equally preserved recognition abilities for the processing of static expressions in patient PS. Our findings, however, revealed that patient PS was strongly impaired in that task. These findings point to the idea that the FFA alone is not sufficient for the processing of static emotional expressions and that the connections between the OFA and FFA may be necessary to reliably achieve this task.

Such an assumption is supported by another model of human face processing, the reverse hierarchical model proposed by Rossion (2008, 2015, see section 1.3.3, Figure 1.43). As stated in the literature review in **CHAPTER 1**, this model suggests that visual inputs are directly sent from early visual areas to the FFA through a route bypassing the OFA. After an initial broad representation of the face in the FFA where face-selectivity emerges, the representation is sharpened through re-entrant connections with the OFA (Rossion, 2008, 2015). In support of this model, Schiltz and colleagues (2006) have demonstrated that despite showing face-selective neural activation, the rFFA of patient PS is insufficient to adequately process faces, showing no adaptation to repeated presentations of the same face (Schiltz et al., 2006). Similar findings were observed by Steeves and colleagues (2006) in patient DF. These results suggest that an integrity of the entire face network is necessary to perform the fine-grained analysis required to discriminate individual faces and most probably static emotional expressions as well.

To sum up, my view is that neither of these two models can fully explain our observations. At the functional level, I defend an intermediate position, which suggests that both models, in combination, are necessary to provide an adequate and satisfactory interpretation to our findings. The model elaborated by Duchaine and Yovel (2015) offers an explanation for the preserved ability of patient PS to recognize dynamic facial expressions and

the processing advantage for dynamic stimuli occurring before and/or after peak efficiency in our healthy observers. In contrast, Rossion's reverse hierarchical view (2008, 2015), which focuses more on the ventral stream and the functional roles of the OFA and FFA, provides an explanation for the selective impairment found in patient PS with respect to static faces.

6.3 Limitations

6.3.1 Limitations to our first experimental contribution

One limitation of our first study ([CHAPTER 3](#)) lies in the fact that infants participated only in one condition, hearing either infant- or adult-directed speech stories. We decided to use this procedure rather than repeated measures due to the attentional difficulties and general fatigability usually observed with infants. Additionally, in our study, we did not assess infants' ability to match audible and visible attributes of gender with static faces. In the context of this thesis, it would have been of particular interest to assess whether infants exhibit a later sensitivity to audio-visual correspondences with static images as compared to dynamic movies.

6.3.2 Limitations to our second experimental contribution

In the second study, we demonstrated that PS used all facial features to represent dynamic expressions. In a previous research, in contrast, Caldara and colleagues (2005) revealed that she used only the mouth region when instructed to extract static information about identity. We interpreted these findings as evidence that distinct representational systems support the extraction of face information for identity and expression recognition ([CHAPTER 4](#)). However, since identity and expression recognition tasks are very different, it is possible that the mouth bias observed for identity is restricted to identity and does not extend to the recognition of static expressions, an assumption that we did not verify in our original experiment, as we reconstructed the mental models of the six basic expressions through the use of dynamic stimuli. We recently addressed this issue in another study in which we used the Bubble technique with eye-tracking to map out PS's facial information use for static expressions (Fiset et al., 2017, for more details on the technique, see [Box 6](#)). As shown in Figure 6.1A, compared to healthy controls, PS mainly relied on the lower part of the face, namely the mouth and the external contours, in order to recognize the expressions of fear, neutral, and happiness. These findings suggest that, in contrast to the view that we defended in the second experimental contribution of this thesis ([CHAPTER 4](#)), some common perceptual mechanisms are dedicated to the processing of identity and static expressions, at least at the level of the extraction of information from the eye region. These results also mirror previous findings reported by Adolphs and colleagues in 2005 with patient SM. This patient, suffering from bilateral amygdala damage,

was severely impaired in the recognition of fear and, like PS, used the lower part of the face to recognize facial expressions (Adolphs et al., 2005) (Figure 6.1B and C).

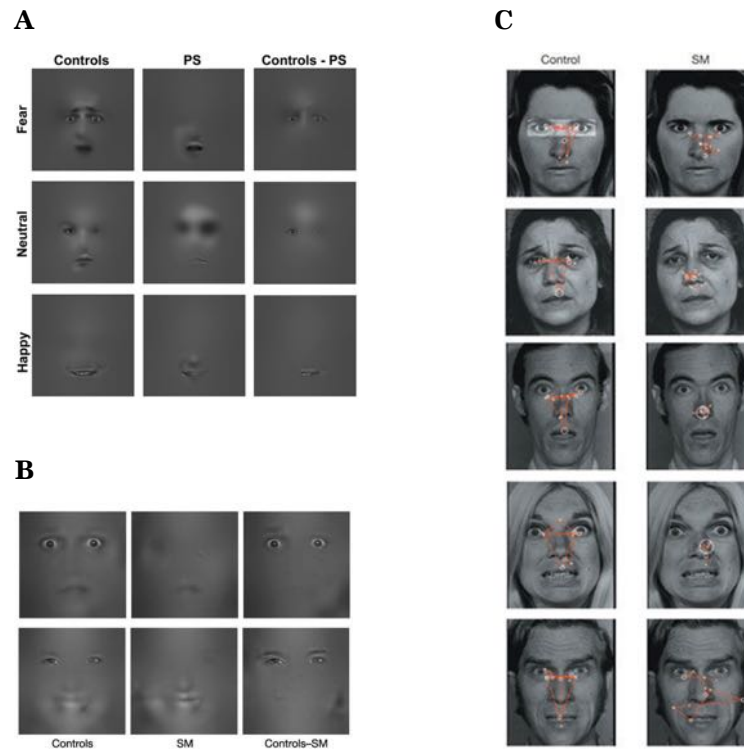


Figure 6.1. Suboptimal facial information use in two brain-damaged patients. (A) Compared to healthy controls (left panel), PS (center) fails to look at the eyes and rather relies on the lower part of the face, namely the mouth and the external contours to categorize the expressions of fear, neutral, and happiness. Reprinted from Fiset et al. (2017). (B) Similar findings were reported in patient SM when instructed to categorize fear (upper row) and happiness (lower row). (C) As demonstrated by the eye-tracking data, SM extracts facial information in a suboptimal way when presented with angry, sad, and fearful faces. Reprinted from Adolphs et al. (2005).

Another limitation pertaining to this second experimental contribution lies in the fact that we did not objectively assess PS's abilities to recognize the identity of moving faces. In order to ensure that PS's ability to recognize dynamic expressions is specific to facial expressions and not due to an overall advantage in the processing of dynamic cues, a dynamic task of identity would have been required. To address this issue, we will conduct a new study to investigate PS's visual information processing strategies when instructed to recognize the identity of moving faces. If PS similarly uses all facial features to recognize dynamic identities as she did with dynamic expressions, our findings will provide further evidence to substantiate the dissociation between the processing of static and dynamic face information.

6.3.3 Limitations to our third experimental contribution

The results of our third study (CHAPTER 5) revealed a decline in the recognition of facial expression with increase in age. This reduction was steeper for some expressions as compared to others and was also modulated by the temporal properties of the expression presented. We should acknowledge here that the cross-sectional nature of our study raises the possibility that our results were influenced by cohort effects. Differences in recognition abilities may stem from differences in the cohorts tested, as it has been previously demonstrated in studies assessing intelligence and other cognitive abilities among different age groups (Flynn, 1987; Gerstorf et al., 2011; Zelinski & Kennison, 2007). For instance, our young and middle-aged adults were all students or employees with higher level jobs and good quality of life, giving rise to the possibility that distinctive cultural or educational factors might have influenced their ability to discern emotions. Additionally, in the mid-40s, social and moral norms encouraged people to suppress some of their expressions (e.g., anger, sadness) in order to maintain socially appropriate interactions. The avoidance or suppression of emotional reactions in our elderly people might have impacted their recognition performance as it has been previously shown in collectivistic cultures (Zhao et al., 2016). Moreover, with the advent of social networking, our younger participants might have been more exposed to static emotional expressions as compared to older adults. Elderly people, in contrast, might be more likely to watch television and, therefore, be exposed to dynamic faces. To the best of our knowledge, the extent to which facial expression recognition is influenced by these cohort effects has never been investigated and should be clarified in future studies that investigate the developmental trajectory of emotion recognition across the adult life span.

Another limitation pertaining to this study lies in the fact that we only used stimuli displaying the emotional expressions of young actors, neither children nor elderly people. Whereas some previous studies did not observe an own-age bias in face recognition (Memon et al., 2003; Mondloch, Maurer, & Ahola, 2006; Wiese, Schweinberger, & Hansen, 2008; Wilcock, Bull, & Vrij, 2007), others reported that younger adults were more accurate in recognizing faces of young adults than the faces of children (Anastasi & Rhodes, 2005, 2006; Britton & Britton, 1969; Kuefner et al., 2008), infants (Chance, Goldstein, & Andersen, 1986; Kuefner et al., 2008), or older adults (He, Ebner, & Johnson, 2011). Other studies observed that both children and adults more readily recognized own-age faces (Anastasi & Rhodes, 2005; Hills & Lewis, 2011; Wright & Stroud, 2002). In light of these findings, it would be interesting to investigate whether elderly people also benefit from the presentation of own-age faces and whether the recognition profiles of the different expressions would be the same when using such a procedure.

6.4 Future directions

Are cross-cultural perceptual strategies used for facial expression recognition in infants influenced by the temporal properties of the expression presented?

A very recent study by Geangu et al. (2016) evidenced that cultural differences in visual perceptual strategies are present in 7-month-old infants, but not in 5-month-old ones, when exposed to static facial emotional expressions. Based on our work with dynamic emotional expressions ([CHAPTER 4](#) and [CHAPTER 5](#)), we will conduct a new study to address the question as to whether differences in perceptual strategies occur earlier than 7 months of age in the case of dynamic faces. To this aim, we will expose 5-month-old infants from Western Caucasian and East Asian cultures to the dynamic expressions of fear and happiness by using a familiarization-novelty preference paradigm (for more details on this paradigm, see [Box 5](#)). Infants will be familiarized with dynamic emotional expressions of fear or happiness across different identities (familiarization phase), before being exposed to the presentation of pairs of dynamic expressions displaying the familiarized emotion together with a novel one (test phase). Infants' perceptual strategies will be monitored by eye-tracking during both familiarization and test phases.

Moreover, to further our research, we also plan to investigate whether differences at the neural level are present in young infants from different cultures, or whether the same neural networks are involved despite differences in the perceptual strategies used. In adults, cross-cultural studies have evidenced differences in the neural substrates underlying the processing of facial emotional expressions. For instance, Chiao and colleagues (2008) have shown differences in the activation of the amygdala when observers were exposed to fear. Differences in the reward and affect system when participants were presented with positive emotions (Park et al., 2016), as well as differences in the anterior cingulate cortex when participants were exposed to the pain of others, were also observed (Xu et al., 2009). In infants, we will use fNIRS in order to examine whether cultural differences in emotion processing are already present at the neural level in infancy.

Do the dorsal face-selective areas of our prosopagous patient elicit stronger neural activation to the presentation of dynamic expressions?

As mentioned above, the results of the second experimental contribution of this thesis support the idea that dissociable cortical pathways underlie the processing of different facial aspects: a ventral stream tuned to the processing of static expressions and identity and a dorsal stream engaged in the processing of dynamic expressions. We are currently planning a future functional

neuroimaging study with fMRI to verify this assumption and to provide a more comprehensive view of the brain regions involved in the processing of static and dynamic faces in PS.

Could the pSTS, as a motion-based recognition system, support identity recognition in prosopagnosic patients with lesions in the ventral stream?

Another issue that we will further investigate is whether PS also benefits from the presentation of dynamic faces for identity recognition. Although in everyday life patient PS does not seem to be able to recover identity through dynamic cues, previous studies have suggested that dynamic faces facilitate learning and identity recognition (for reviews, see Butcher & Lander, 2017; Xiao et al., 2014). These studies were conducted using familiar faces (e.g., Lander & Bruce, 2000; O'Toole et al., 2002; O'Toole & Roark, 2010), unfamiliar faces (e.g., Butcher et al., 2011; Lander & Bruce, 2003), or suboptimal stimuli (e.g., Lander, Bruce, & Hill, 2001). Moreover, Anzellotti and Caramazza (2017) have recently provided converging evidence that the pSTS contains invariant representations of person identity that are critical for multimodal identity recognition. In addition to their sensitivity to motion information, dorsal-face selective brain regions may also be the neuroanatomical locus of the processing and integration of multisensory identity cues (see, Yovel & O'Toole, 2016).

Do other patients also benefit from the presentation of dynamic expressions?

Moving forward from our third study, we are currently investigating the ability of patients suffering from mild cognitive impairment (MCI) to recognize static, dynamic, and shuffled expressions. At the moment, the number of people affected by dementia is believed to be close to 50 million worldwide, a clinical condition that will affect more than 130,000 million people by 2050. A new case of dementia is expected to be diagnosed every three second, and in Switzerland, the number of people suffering from dementia, MCI, or Alzheimer's disease will dramatically increase. Taking these alarming statistics into consideration, a growing body of research is now focusing on this clinical condition. Yet, till date, none of these studies have investigated the recognition of static, dynamic, and shuffled expressions in MCI or Alzheimer's patients, using a database of stimuli controlled for the amount of information provided over time.

In order to address this gap in the cognitive aging literature, we decided to conduct this new research with the aim to answer the three following questions: Do MCI patients also benefit from the presentation of dynamic expressions? If so, is this dynamic advantage similarly driven by suboptimal performance with static expressions as observed in normal aging adults?

Finally, is the potential suboptimal performance with static expressions in MCI patients more pronounced than the one observed in normal aging people?

At the practical level, our results could have implications that go beyond our research, potentially constituting a future diagnostic tool. We could also use our findings in the case of elderly patients with affective decoding disorders, in cognitive interventions, as rehabilitation techniques when patients encounter social difficulties, or in combination with more global therapeutic approaches.

More generally, future experimental investigations could be carried out using the same stimuli, procedure, and statistical analyses as used in our third study, but with other clinical populations such as autistic children, children or adolescents suffering from PTSD, or adults with difficult life trajectories, those who have been abused, misemployed, maltreated or are migrants.

Optogenetics: a window to the future

On a broader note, besides fNIRS, other technologies have been recently developed to further dismantle brain circuits underlying cognition and behavior. Optogenetics, for instance, which combines optical and genetic methods to optically control the electrical activity of excitable neurons or cells, is one of the most powerful and promising technologies in the field of neuroscience (Fenno et al., 2011; Kalanithi & Henderson, 2012). This approach has been successfully used in rodents to potentiate treatment mechanisms or functional recovery after stroke (Cheng et al., 2014; Gradinaru et al., 2009), to deconstruct and map neural circuits involved in neurodegenerative diseases (e.g., Gradinaru et al., 2009), epilepsy (e.g., Kokaia, Andersson, & Ledri, 2013; Krook-Magnuson et al., 2013; Tonnesen et al., 2009), mood disorders (Lammel, Tye, & Warden, 2014), or schizophrenia (Cho & Sohal, 2014; Kellendonk et al., 2006). Optogenetics has also been used to modulate previously acquired fear associations, facilitating their extinction (Klavir et al., 2017). It is of major importance that this technology has been very recently applied for the first time to human brain tissues (Andersson et al., 2016), paving the way to promising new research avenues in the medical domain. Optogenetics may one day also have potential applications for brain-damaged patients or patients suffering from dementia or other pathologies associated with abnormal cognitive aging.

6.5 Conclusions

The overall aim of this thesis was to use dynamic faces in order to investigate different aspects of face processing, in different populations and age groups. In our first study, we used dynamic faces to investigate whether the multisensory integration of gender is influenced by the speech manner. Our results revealed an earlier emergence of the perception of multisensory gender coherence for adult-directed speech. These findings are of interest to those studying infant face perception, infant-directed speech, and learning. In our second and third study, we investigated whether dynamic cues contribute to the processing of facial expressions in a single-case of *pure* acquired prosopagnosia as well as in healthy participants across the life span. Our findings provide new evidence for a dynamic advantage for facial expression recognition in brain-damaged patients, as well as in elderly people.

I hope that my work will contribute – even modestly – to highlight the importance of moving towards a comprehensive understanding of dynamic human face processing. I am truly convinced that exploring face perception with multisensory dynamic faces will be a major step forward, as it will provide a tighter link between behavioral, neuropsychological, or functional neuroimaging findings and the way faces are processed by humans in everyday life. My future work will follow in this direction with the long term aim to move towards a genuine ecological approach to the study of face processing. I hope that one day, with the help of new (neuroimaging) technologies, it will be possible to precisely track online the cascade of events underlying the very complex information processing of faces, from the very entry level (the eye) into the fine-grained neural processing. Importantly, I also believe that this knowledge and techniques could be of use in the future in a variety of clinical or neuropsychological settings, as well as in the realm of patient care to tailor early interventions or rehabilitation training programs. I am fully aware that this is a utopian view. However, landing on the moon was also a utopian view in the beginning of the 19th century.

References

- Acerra, F., Burnod, Y., & de Schonen, S. (2002). Modelling aspects of face processing in early infancy. *Developmental Science*, 5(1), 98-117.
- Adolphs, R. (2013). The biology of fear. *Current Biology*, 23(2), R79-93.
- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, 433(7021), 68-72.
- Adolphs, R., & Tranel, D. (2004). Impaired judgments of sadness but not happiness following bilateral amygdala damage. *Journal of Cognitive Neuroscience*, 16(3), 453-462.
- Adolphs, R., Tranel, D., & Damasio, A. R. (2003). Dissociable neural systems for recognizing emotions. *Brain and Cognition*, 52(1), 61-69.
- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, 372(6507), 669-672.
- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1995). Fear and the human amygdala. *Journal of Neuroscience*, 15(9), 5879-5891.
- Adolphs, R., Tranel, D., Hamann, S., Young, A. W., Calder, A. J., Phelps, E. A., Anderson, A., Lee, G. P., & Damasio, A. R. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia*, 37(10), 1111-1117.
- Afraz, A., Boyden, E. S., & DiCarlo, J. J. (2015). Optogenetic and pharmacological suppression of spatial clusters of face neurons reveal their causal role in face gender discrimination. *Proceedings of the National Academy of Sciences, USA*, 112(21), 6730-6735.
- Alves, N. T. (2013). Recognition of static and dynamic facial expressions: A study review. *Estudos de Psicologia (Natal)*, 18(1), 125-130.
- Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, 16(5), 403-410.
- Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review*, 12(6), 1043-1047.

-
- Anastasi, J. S., & Rhodes, M. G. (2006). Evidence for an own-age bias in face recognition. *North American Journal of Psychology*, 8(2), 237-253.
- Andersen, S. L. (2003). Trajectories of brain development: Point of vulnerability or window of opportunity? *Neuroscience & Biobehavioral Reviews*, 27(1), 3-18.
- Anderson, J. D., & Fisher Anderson, B. (2007). *Moving image theory: Ecological considerations*. Carbondale: Southern Illinois University Press.
- Andersson, M., Avaliani, N., Svensson, A., Wickham, J., Pinborg, L. H., Jespersen, B., Christiansen, S. H., Bengzon, J., Woldbye, D. P., & Kokaia, M. (2016). Optogenetic control of human neurons in organotypic brain cultures. *Scientific Reports*, 6, 24818.
- Andruski, J., & Kuhl, P. (1997). *The acoustic structure of vowels in infant-and adult-directed speech*. Paper presented at the Biannual Meeting of the Society for Research in Child Development, Washington, DC (USA).
- Anzellotti, S., & Caramazza, A. (2017). Multimodal representations of person identity individuated with fMRI. *Cortex*, 89, 85-97.
- Anzellotti, S., Fairhall, S. L., & Caramazza, A. (2013). Decoding representations of face identity that are tolerant to rotation. *Cerebral Cortex*, 24(8), 1988-1995.
- Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R., & Livingstone, M. S. (2017). Seeing faces is necessary for face-domain formation. *Nature Neuroscience*, 20(10), 1404-1412.
- Arnold, M. B. (1960). *Emotion and personality*. New York: Columbia University Press.
- Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain Topography*, 24(2), 149-163.
- Aslin, R. (2000). *Interpretation of infant listening times using the headturn preference technique*. Paper presented at the International Conference on Infant Studies, Brighton (UK).
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6), 717-746.
- Avidan, G., Tanzer, M., Hadj-Bouziane, F., Liu, N., Ungerleider, L. G., & Behrmann, M. (2014). Selective dissociation between core and extended regions of the face processing network in congenital prosopagnosia. *Cerebral Cortex*, 24(6), 1565-1578.

-
- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, 19(7), 724-732.
- Axelrod, V., & Yovel, G. (2012). Hierarchical processing of face viewpoint in human visual cortex. *Journal of Neuroscience*, 32(7), 2442-2452.
- Back, E., Ropar, D., & Mitchell, P. (2007). Do the eyes have it? Inferring mental states from animated faces in autism. *Child Development*, 78(2), 397-411.
- Baker, C. I. (2008). Face to face with cortex. *Nature Neuroscience*, 11(8), 862-864.
- Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science*, 11(4), 267-273.
- Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science*, 17(2), 159-163.
- Barenholtz, E., Lewkowicz, D. J., Davidson, M., & Mavica, L. (2014). Categorical congruence facilitates multisensory associative learning. *Psychonomic Bulletin & Review*, 21(5), 1346-1352.
- Barkow, J. H., Cosmides, L., & Tooby, J. (1995). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.
- Barrett, L. F., & Kensinger, E. A. (2010). Context is routinely encoded during emotion perception. *Psychological Science*, 21(4), 595-599.
- Barton, J. J., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, 58(1), 71-78.
- Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance*, 4(3), 373-379.
- Bate, S., & Bennetts, R. (2015). The independence of expression and identity in face-processing: evidence from neuropsychological case studies. *Frontiers in Psychology*, 6, 1-7.
- Bate, S., Cook, S. J., Mole, J., & Cole, J. (2013). First report of generalized face processing difficulties in möbius sequence. *PLoS One*, 8(4), 1-9.
- Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, 23(2), 223-229.

- Bayet, L., Pascalis, O., & Gentaz, É. (2014). Le développement de la discrimination des expressions faciales émotionnelles chez les nourrissons dans la première année. *Année Psychologique*, 114(03), 469-500.
- Bayet, L., Quinn, P. C., Laboissière, R., Caldara, R., Lee, K., & Pascalis, O. (2017). Fearful but not happy expressions boost face detection in human infants. *Proceedings of the Royal Society of London B*, 284(1862), 1-9.
- Baylis, G. C., Rolls, E. T., & Leonard, C. M. (1987). Functional subdivisions of the temporal lobe neocortex. *Journal of Neuroscience*, 7(2), 330-342.
- Beaupré, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology*, 36(3), 355-370.
- Bell, A. H., Malecek, N. J., Morin, E. L., Hadj-Bouziane, F., Tootell, R. B., & Ungerleider, L. G. (2011). Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. *Journal of Neuroscience*, 31(34), 12229-12240.
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8(6), 551-565.
- Benton, A., & Van Allen, M. (1972). Prosopagnosia and facial discrimination. *Journal of the Neurological Sciences*, 15, 167-172.
- Bernstein, M., Erez, Y., Blank, I., & Yovel, G. (2017). The processing of dynamic faces in the human brain: Support for an integrated neural framework of face processing. *BioRxiv*, 140939.
- Bernstein, M., & Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. *Neuroscience & Biobehavioral Reviews*, 55, 536-546.
- Biele, C., & Grabowska, A. (2006). Sex differences in perception of emotion intensity in dynamic and static facial expressions. *Experimental Brain Research*, 171(1), 1-6.
- Bilalic, M., Langner, R., Ulrich, R., & Grodd, W. (2011). Many faces of expertise: Fusiform face area in chess experts and novices. *Journal of Neuroscience*, 31(28), 10206-10214.
- Blair, R. J., & Cipolotti, L. (2000). Impaired social response reversal: A case of acquired sociopathy. *Brain*, 123(6), 1122-1141.
- Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS One*, 3(8), 1-8.

- Blakemore, S.-J. (2012). Imaging brain development: The adolescent brain. *NeuroImage*, 61(2), 397-406.
- Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry*, 47(3-4), 296-312.
- Bodamer, J. (1947). Die Prosopagnosie. *Archive für Psychiatrie und Nervenkrankheiten*, 179, 6-54.
- Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer. *Ear & Hearing*, 32(2), 266.
- Bornstein, M. H., Arterberry, M. E., Mash, C., & Manian, N. (2011). Discrimination of facial expression by 5-month-old infants of nondepressed and clinically depressed mothers. *Infant Behavior and Development*, 34(1), 100-106.
- Bornstein, M. H., Mash, C., Arterberry, M. E., & Manian, N. (2012). Object perception in 5-month-old infants of clinically depressed and nondepressed mothers. *Infant Behavior and Development*, 35(1), 150-157.
- Borod, J. C., Yecker, S. A., Brickman, A. M., Moreno, C. R., Sliwinski, M., Foldi, N. S., Alpert, M., & Welkowitz, J. (2004). Changes in posed facial expression of emotion across the adult life span. *Experimental Aging Research*, 30(4), 305-331.
- Bould, E., & Morris, N. (2008). Role of motion signals in recognizing subtle facial expressions of emotion. *British Journal of Psychology*, 99(2), 167-189.
- Bould, E., Morris, N., & Wink, B. (2008). Recognising subtle emotional expressions: The role of facial movements. *Cognition and Emotion*, 22(8), 1569-1587.
- Bouvier, S. E., & Engel, S. A. (2004). Amblyopia & other visual disorders: Patterns of cortical damage in achromatopsia and prosopagnosia. *Journal of Vision*, 4(8), 183-191.
- Bowers, D., Bauer, R. M., Coslett, H. B., & Heilman, K. M. (1985). Processing of faces by patients with unilateral hemisphere lesions. I. Dissociation between judgments of facial affect and facial identity. *Brain and Cognition*, 4(3), 258-272.
- Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433-436.
- Britton, J. O., & Britton, J. H. (1969). Discrimination of age by preschool children. *Journal of Gerontology*, 24(4), 457-460.

- Bronshtein, A., Antonova, T., Kamenetskaya, A., Luppova, N., & Sytova, V. (1958). On the development of the functions of analyzers in infants and some animals at the early stage of ontogenesis. *Problems of Evolution of Physiological Functions*, 106-116.
- Brosigole, L., & Weisman, J. (1995). Mood recognition across the ages. *International Journal of Neuroscience*, 82(3-4), 169-189.
- Brown, E., & Perrett, D. I. (1993). What gives a face its gender? *Perception*, 22(7), 829-840.
- Bruce, V., Burton, A. M., Hanna, E., Healey, P., Mason, O., Coombes, A., Fright, R., & Linney, A. (1993). Sex discrimination: How do we tell the difference between male and female faces? *Perception*, 22(2), 131-152.
- Bruce, V., Campbell, R. N., Doherty-Shedden, G., Import, A., Langton, S., McAuley, S., & Wright, R. (2000). Testing face processing skills in children. *British Journal of Developmental Psychology*, 18(3), 319-333.
- Bruce, V., & Valentine, T. (1988). When a nod's as good as a wink: The role of dynamic information in facial recognition. In M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, pp. 169-174). New York, NY: John Wiley & Sons.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77(Pt 3), 305-327.
- Bruyer, R., Laterre, C., Seron, X., Feyereisen, P., Strypstein, E., Pierrard, E., & Rectem, D. (1983). A case of prosopagnosia with some preserved covert remembrance of familiar faces. *Brain and Cognition*, 2(3), 257-284.
- Bukach, C. M., Bub, D. N., Gauthier, I., & Tarr, M. J. (2006). Perceptual expertise effects are not all or none: Spatially limited perceptual expertise for faces in a case of prosopagnosia. *Journal of Cognitive Neuroscience*, 18(1), 48-63.
- Bukach, C. M., Grand, R., Kaiser, M. D., Bub, D. N., & Tanaka, J. W. (2008). Preservation of mouth region processing in two cases of prosopagnosia. *Journal of Neuropsychology*, 2(1), 227-244.
- Bukowski, H., Dricot, L., Hanseeuw, B., & Rossion, B. (2013). Cerebral lateralization of face-sensitive areas in left-handers: Only the FFA does not get it right. *Cortex*, 49(9), 2583-2589.
- Bushneil, I., Sai, F., & Mullin, J. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, 7(1), 3-15.

-
- Busigny, T., Graf, M., Mayer, E., & Rossion, B. (2010a). Acquired prosopagnosia as a face-specific disorder: Ruling out the general visual similarity account. *Neuropsychologia*, 48(7), 2051-2067.
- Busigny, T., Joubert, S., Felician, O., Ceccaldi, M., & Rossion, B. (2010b). Holistic perception of the individual face is specific and necessary: Evidence from an extensive case study of acquired prosopagnosia. *Neuropsychologia*, 48(14), 4057-4092.
- Busigny, T., & Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, 46(8), 965-981.
- Butcher, N., & Lander, K. (2017). Exploring the motion advantage: Evaluating the contribution of familiarity and differences in facial motion. *The Quarterly Journal of Experimental Psychology*, 70(5), 919-929.
- Butcher, N., Lander, K., Fang, H., & Costen, N. (2011). The effect of motion at encoding and retrieval for same-and other-race face recognition. *British Journal of Psychology*, 102(4), 931-942.
- Caldara, R. (2017). Culture Reveals a Flexible System for Face Processing. *Current Directions in Psychological Science*, 26(3), 249-255.
- Caldara, R., & Abdi, H. (2006). Simulating the 'other-race' effect with autoassociative neural networks: Further evidence in favor of the face-space model. *Perception*, 35(5), 659-670.
- Caldara, R., Rossion, B., Bovet, P., & Hauert, C. A. (2004). Event-related potentials and time course of the "other-race" face classification advantage. *Neuroreport*, 15(5), 905-910.
- Caldara, R., Schyns, P., Mayer, E., Smith, M. L., Gosselin, F., & Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of Cognitive Neuroscience*, 17(10), 1652-1666.
- Caldara, R., & Seghier, M. L. (2009). The fusiform face area responds automatically to statistical regularities optimal for face categorization. *Human Brain Mapping*, 30(5), 1615-1625.
- Caldara, R., Seghier, M. L., Rossion, B., Lazeyras, F., Michel, C., & Hauert, C. A. (2006). The fusiform face area is tuned for curvilinear patterns with more high-contrasted elements in the upper part. *NeuroImage*, 31(1), 313-319.
- Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the 'spotlight' reveals universal information use for face recognition. *PLoS One*, 5(3), 1-12.

-
- Calder, A. J. (1996). Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cognitive Neuropsychology*, 13(5), 699-745.
- Calder, A. J. (2011). Does facial identity and facial expression recognition involve separate visual routes? In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 427-448). Oxford: Oxford University Press.
- Calder, A. J., Burton, A. M., Miller, P., Young, A. W., & Akamatsu, S. (2001a). A principal component analysis of facial expressions. *Vision Research*, 41(9), 1179-1208.
- Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, 3(11), 1077-1078.
- Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., & Young, A. W. (2003). Facial expression recognition across the adult life span. *Neuropsychologia*, 41(2), 195-202.
- Calder, A. J., Lawrence, A. D., & Young, A. W. (2001b). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience*, 2(5), 352-363.
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, 6(8), 641-651.
- Campbell, J., & Burke, D. (2009). Evidence that identity-dependent and identity-independent neural populations are recruited in the perception of five basic emotional facial expressions. *Vision Research*, 49(12), 1532-1540.
- Campbell, R., Landis, T., & Regard, M. (1986). Face recognition and lipreading. A neurological dissociation. *Brain*, 109(Pt 3), 509-521.
- Camras, L. A., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior*, 9(2), 84-94.
- Canestrini, S. (1913). *Über das Sinnesleben des Neugeborenen nach physiologischen Experimenten*. Berlin: Springer.
- Carlson, J. M., & Mujica-Parodi, L. R. (2015). Facilitated attentional orienting and delayed disengagement to conscious and nonconscious fearful faces. *Journal of Nonverbal Behavior*, 39(1), 69-77.
- Caron, R. F., Caron, A. J., & Myers, R. S. (1982). Abstraction of invariant face expressions in infancy. *Child Development*, 1008-1015.

-
- Carstensen, L. L., Fung, H. H., & Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and Emotion*, 27(2), 103-123.
- Carstensen, L. L., & Mikels, J. A. (2005). At the intersection of emotion and cognition: Aging and the positivity effect. *Current Directions in Psychological Science*, 14(3), 117-121.
- Casey, B., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: What have we learned about cognitive development? *Trends in Cognitive Sciences*, 9(3), 104-110.
- Cavedo, E., Boccardi, M., Ganzola, R., Canu, E., Beltramello, A., Caltagirone, C., Thompson, P., & Frisoni, G. (2011). Local amygdala structural differences with 3T MRI in patients with Alzheimer disease. *Neurology*, 76(8), 727-733.
- Chance, J. E., Goldstein, A. G., & Andersen, B. (1986). Recognition memory for infant faces: An analog of the other-race effect. *Bulletin of the Psychonomic Society*, 24(4), 257-260.
- Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General*, 132(2), 310-324.
- Chen, J. J., Rosas, H. D., & Salat, D. H. (2011). Age-associated reductions in cerebral blood flow are independent from regional atrophy. *NeuroImage*, 55(2), 468-478.
- Cheng, M., Wang, E., Woodson, W., Wang, S., Sun, G., Lee, A., Arac, A., Fenno, L., Deisseroth, K., & Steinberg, G. (2014). Optogenetic neuronal stimulation promotes functional recovery after stroke. *Proceedings of the National Academy of Sciences, USA*, 111(35), 12913-12918.
- Chiao, J. Y., Iidaka, T., Gordon, H. L., Nogawa, J., Bar, M., Aminoff, E., Sadato, N., & Ambady, N. (2008). Cultural specificity in amygdala response to fear faces. *Journal of Cognitive Neuroscience*, 20(12), 2167-2174.
- Cho, K. K., & Sohal, V. S. (2014). Optogenetic approaches for investigating neural pathways implicated in schizophrenia and related disorders. *Hum Mol Genet*, 23(R1), R64-68.
- Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory & Cognition*, 26(4), 780-790.
- Clemente, C. D. (1997). *Anatomy: A regional atlas of the human body* (4th ed.). Baltimore: Williams & Wilkins.
- Cohen, L. B., & Strauss, M. S. (1979). Concept acquisition in the human infant. *Child Development*, 419-424.

- Cohen, P. (2012). *Applied data analytic techniques for turning points research*. New York: Routledge.
- Cohen-Kadosh, K. (2011). What can emerging cortical face networks tell us about mature brain organisation? *Developmental Cognitive Neuroscience*, 1(3), 246-255.
- Cohen-Kadosh, K., Henson, R. N., Cohen Kadosh, R., Johnson, M. H., & Dick, F. (2010). Task-dependent activation of face-sensitive cortex: An fMRI adaptation study. *Journal of Cognitive Neuroscience*, 22(5), 903-917.
- Cole, M., & Perez-Cruet, J. (1964). Prosopagnosia. *Neuropsychologia*, 2, 237-246.
- Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61(5), 1584-1595.
- Cornell, E. H. (1974). Infants' discrimination of photographs of faces following redundant presentations. *Journal of Experimental Child Psychology*, 18(1), 98-106.
- Coulon, M., Guellai, B., & Streri, A. (2011). Recognition of unfamiliar talking faces at birth. *International Journal of Behavioral Development*, 35(3), 282-287.
- Craig, K. D., Hyde, S. A., & Patrick, C. J. (1991). Genuine, suppressed and faked facial behavior during exacerbation of chronic low back pain. *Pain*, 46(2), 161-171.
- Crawford, J. R., & Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia*, 40(8), 1196-1208.
- Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. *Processes of Change in Brain and Cognitive Development. Attention and performance XXI*, 21, 249-274.
- Cudeck, R., & Klebe, K. J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods*, 7(1), 41-63.
- Cunningham, D. W., & Wallraven, C. (2009a). Dynamic information for the recognition of conversational expressions. *Journal of Vision*, 9(13), 1-17.
- Cunningham, D. W., & Wallraven, C. (2009b). *The interaction between motion and form in expression recognition*. Paper presented at the Proceedings of the 6th symposium on applied perception in graphics and visualization (pp. 41-44), New York (USA).
- Dalrymple, K. A., Oruc, I., Duchaine, B., Pancaroglu, R., Fox, C. J., Iaria, G., Handy, T. C., & Barton, J. J. (2011). The anatomic basis of the right face-selective N170 in acquired prosopagnosia: A combined ERP/fMRI study. *Neuropsychologia*, 49(9), 2553-2563.

- Damasio, A. R. (1985). Disorders of complex visual processing: Agnosia, achromatopsia, Balint's syndrome, and related difficulties of orientation and construction. In M. M. Mesulam (Ed.), *Principles of behavioral neurology* (pp. 259-288). Philadelphia: F.A. Davis.
- Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982). Prosopagnosia: Anatomic basis and behavioral mechanisms. *Neurology*, 32(4), 331-341.
- Darwin, C. (1999/1872). *The Expression of the emotions in man and animals* (3 ed.). London: Fontana Press.
- De Gelder, B., Pourtois, G., Vroomen, J., & Bachoud-Levi, A. C. (2000). Covert processing of faces in prosopagnosia is restricted to facial expressions: Evidence from cross-modal bias. *Brain and Cognition*, 44, 425-444.
- De Renzi, E., & Di Pellegrino, G. (1998). Prosopagnosia and alexia without object agnosia. *Cortex*, 34(3), 403-415.
- De Sonneville, L., Verschoor, C., Njikiktijen, 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.
- De Winter, F.-L., Van den Stock, J., de Gelder, B., Peeters, R., Jastorff, J., Sunaert, S., Vanduffel, W., Vandenberghe, R., & Vandenbulcke, M. (2016). Amygdala atrophy affects emotion-related activity in face-responsive regions in frontotemporal degeneration. *Cortex*, 82, 179-191.
- Del Giudice, M., & Colle, L. (2007). Differences between children and adults in the recognition of enjoyment smiles. *Developmental Psychology*, 43(3), 796-803.
- Delis, I., Chen, C., Jack, R. E., Garrod, O. G., Panzeri, S., & Schyns, P. G. (2016). Space-by-time manifold representation of dynamic facial expressions for emotion categorization. *Journal of Vision*, 16(8), 14.
- Delvenne, J.-F., Seron, X., Coyette, F., & Rossion, B. (2004). Evidence for perceptual deficits in associative visual (prosop) agnosia: A single-case study. *Neuropsychologia*, 42(5), 597-612.
- Deruelle, C., & de Schonen, S. (1998). Do the right and left hemispheres attend to the same visuospatial information within a face in infancy? *Developmental Neuropsychology*, 14(4), 535-554.
- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, 4(8), 2051-2062.

- Dohen, M., Loevenbruck, H., & Hill, H. (2009). Recognizing prosody from the lips. In A. Wee-Chung-Liew & S. Wang (Eds.), *Visual speech recognition: Lip segmentation and mapping* (pp. 416-438). Hershey, PA: Medical Information Science Reference.
- Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, 1, 393-416.
- Duncan, T.E, Duncan, S, & Strycker, L.A. (2013). *An introduction to latent variable growth curve modeling: Concepts, issues, and application*. New York: Routledge Academic.
- Dunst, C. J., & Hamby, D. W. (2012). Guide for calculating and interpreting effect sizes and confidence intervals in intellectual and developmental disability research studies. *Journal of Intellectual and Developmental Disability*, 37(2), 89-99.
- Durand, K., Gallay, M., Seigneuric, A., Robichon, F., & Baudouin, J.-Y. (2007). The development of facial emotion recognition: The role of configural information. *Journal of Experimental Child Psychology*, 97(1), 14-27.
- Durston, S., Pol, H. E. H., Casey, B., Giedd, J. N., Buitelaar, J. K., & Van Engeland, H. (2001). Anatomical MRI of the developing human brain: What have we learned? *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(9), 1012-1020.
- Ehrlich, S. M., Schiano, D. J., & Sheridan, K. (2000). *Communicating facial affect: It's not the realism, it's the motion*. Paper presented at the CHI'00 Extended Abstracts on Human Factors in Computing Systems (pp. 252-253), New York (USA).
- Ekman, P. (1992a). Are there basic emotions? *Psychological Review*, 99(3), 550-553.
- Ekman, P. (1992b). An argument for basic emotions. *Cognition & Emotion*, 6(3-4), 169-200.
- Ekman, P. (2003). *Emotions revealed. Understanding faces and feelings*. London: The Orion Publishing Group Ltd.
- Ekman, P., & Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*. New York: Oxford University Press.
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 124-129.
- Ekman, P., & Friesen, W. V. (1976a). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, 1(1), 56-75.
- Ekman, P., & Friesen, W. V. (1976b). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

-
- Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (1982). What emotion categories or dimensions can observers judge from facial behavior? In P. Ekman (Ed.), *Emotion in the human face* (pp. 39-55). New York: Cambridge University Press.
- Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221, 1208-1210.
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, 164(3875), 86-88.
- Fagan, J. F., & Singer, L. T. (1979). The role of simple feature differences in infants' recognition of faces. *Infant Behavior and Development*, 2, 39-45.
- Fantz, R. L. (1961). The origin of form perception. *Scientific American*, 204(5), 66-73.
- Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, 146(3644), 668-670.
- Fantz, R. L. (1965). Visual perception from birth as shown by pattern selectivity. *Annals of the New York Academy of Sciences*, 118(1), 793-814.
- Farah, M. J. (1990). *Visual agnosia*. Cambridge: MIT Press.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review*, 105(3), 482-498.
- Farroni, T., Chiarelli, A. M., Lloyd-Fox, S., Massaccesi, S., Merla, A., Di Gangi, V., Mattarello, T., Faraguna, D., & Johnson, M. H. (2013). Infant cortex responds to other humans from shortly after birth. *Scientific Reports*, 3(2851), 1-5.
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, USA*, 99(14), 9602-9605.
- Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings of the National Academy of Sciences, USA*, 102(47), 17245-17250.
- Farroni, T., Menon, E., Rigato, S., & Johnson, M. H. (2007). The perception of facial expressions in newborns. *European Journal of Developmental Psychology*, 4(1), 2-13.

-
- Fenno, L., Yizhar, O., & Deisseroth, K. (2011). The development and application of optogenetics. *Annual Review of Neuroscience*, 34, 389-412.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8(2), 181-195.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10(3), 279-293.
- Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9(3), 228-231.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20(1), 104-113.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16(3), 477-501.
- Field, T. M., Woodson, R., Cohen, D., Greenberg, R., Garcia, R., & Collins, K. (1983). Discrimination and imitation of facial expressions by term and preterm neonates. *Infant Behavior and Development*, 6(4), 485-489.
- Fiorentini, C., & Viviani, P. (2011). Is there a dynamic advantage for facial expressions? *Journal of Vision*, 11(3), 1-15.
- Fischer, H., Nyberg, L., & Bäckman, L. (2010). Age-related differences in brain regions supporting successful encoding of emotional faces. *Cortex*, 46(4), 490-497.
- Fischer, H., Sandblom, J., Gavazzini, J., Fransson, P., Wright, C. I., & Bäckman, L. (2005). Age-differential patterns of brain activation during perception of angry faces. *Neuroscience Letters*, 386(2), 99-104.
- Fiset, D., Blais, C., Royer, J., Richoz, A.-R., Dugas, G., & Caldara, R. (2017). Mapping the impairment in decoding static facial expressions of emotion in prosopagnosia. *Social Cognitive and Affective Neuroscience*, 12(8), 1334-1341.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101(2), 171-191.
- Folstein, M. F. (1975). A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.

- Fox, C. J., & Barton, J. J. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research*, 1127(1), 80-89.
- Fox, C. J., Iaria, G., & Barton, J. J. (2009a). Defining the face processing network: Optimization of the functional localizer in fMRI. *Human Brain Mapping*, 30(5), 1637-1651.
- Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. (2009b). The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *NeuroImage*, 44(2), 569-580.
- Friendly, R. H., Rendall, D., & Trainor, L. J. (2013). Plasticity after perceptual narrowing for voice perception: Reinstating the ability to discriminate monkeys by their voices at 12 months of age. *Frontiers in Psychology*, 4, 1-8.
- Frijda, N. H. (2008). The psychologists' point of view. In M. Lewis, J. M. Haviland-Jones & L. Feldman Barrett (Eds.), *Handbook of emotions* (pp. 68-87). New York: The Guilford Press.
- Ganel, T., Goshen-Gottstein, Y., & Ganel, T. (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.
- Ganel, T., Valyear, K. F., Goshen-Gottstein, Y., & Goodale, M. A. (2005). The involvement of the "fusiform face area" in processing facial expression. *Neuropsychologia*, 43(11), 1645-1654.
- Gao, X., & Maurer, D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of Experimental Child Psychology*, 102(4), 503-521.
- 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.
- Gao, X., Maurer, D., & Nishimura, M. (2013). Altered representation of facial expressions after early visual deprivation. *Frontiers in Psychology*, 4, 1-11.
- Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology*, 11(2), 219-224.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000a). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191.
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience*, 2(6), 568-573.

- Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000b). The fusiform “face area” is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, 12(3), 495-504.
- Geangu, E., Ichikawa, H., Lao, J., Kanazawa, S., Yamaguchi, M. K., Caldara, R., & Turati, C. (2016). Culture shapes 7-month-olds’ perceptual strategies in discriminating facial expressions of emotion. *Current Biology*, 26(14), R663-R664.
- Geinisman, Y., Detoledo-Morrell, L., Morrell, F., & Heller, R. E. (1995). Hippocampal markers of age-related memory dysfunction: Behavioral, electrophysiological and morphological perspectives. *Progress in Neurobiology*, 45(3), 223-252.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 457-472.
- George, M. S., Ketter, T. A., Gill, D. S., Haxby, J. V., Ungerleider, L. G., Herscovitch, P., & Post, R. M. (1993). Brain regions involved in recognizing facial emotion or identity: An oxygen-15 PET study. *The Journal of Neuropsychiatry and Clinical Neuroscience*, 5(4), 384-394.
- George, P. A., & Hole, G. J. (1995). Factors influencing the accuracy of age estimates of unfamiliar faces. *Perception*, 24(9), 1059-1073.
- Gepner, B., Deruelle, C., & Grynfeldt, S. (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. *Journal of Autism and Developmental Disorders*, 31(1), 37-45.
- Gerstorf, D., Ram, N., Hoppmann, C., Willis, S. L., & Schaie, K. W. (2011). Cohort differences in cognitive aging and terminal decline in the Seattle Longitudinal Study. *Developmental Psychology*, 47(4), 1026-1041.
- Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11(5), 473-490.
- Giedd, J. N. (2008). The teen brain: Insights from neuroimaging. *Journal of Adolescent Health*, 42(4), 335-343.
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C., & Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, 2(10), 861-863.
- Giedd, J. N., Vaituzis, A. C., Hamburger, S. D., Lange, N., Rajapakse, J. C., Kaysen, D., Vauss, Y. C., & Rapoport, J. L. (1996). Quantitative MRI of the temporal lobe, amygdala, and hippocampus in normal human development: Ages 4-18 years. *Journal of Comparative Neurology*, 366(2), 223-230.

- Gill, D., Garrod, O. G., Jack, R. E., & Schyns, P. G. (2014). Facial movements strategically camouflage involuntary social signals of face morphology. *Psychological Science*, 25(5), 1079-1086.
- Goffaux, V., Peters, J., Haubrechts, J., Schiltz, C., Jansma, B., & Goebel, R. (2011). From coarse to fine? Spatial and temporal dynamics of cortical face processing. *Cerebral Cortex*, 21(2), 467-476.
- Gold, J. M., Barker, J. D., Barr, S., Bittner, J. L., Bromfield, W. D., Chu, N., Goode, R. A., Lee, D., Simmons, M., & Srinath, A. (2013). The efficiency of dynamic and static facial expression recognition. *Journal of Vision*, 13(5), 1-12.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20-25.
- Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, 56(4), 544-549.
- Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, 41(17), 2261-2271.
- Gradinaru, V., Mogri, M., Thompson, K. R., Henderson, J. M., & Deisseroth, K. (2009). Optical deconstruction of parkinsonian neural circuitry. *Science*, 324(5925), 354-359.
- Grainger, S. A., Henry, J. D., Phillips, L. H., Vanman, E. J., & Allen, R. (2015). Age deficits in facial affect recognition: The influence of dynamic cues. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 72(4), 622-632.
- Gray, J. A. (1982). *The neuropsychology of anxiety*. Oxford: Oxford University Press.
- Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, 7(5), 555-562.
- Gross, A. L., & Ballif, B. (1991). Children's understanding of emotion from facial expressions and situations: A review. *Developmental Review*, 11(4), 368-398.
- Gross, C. G., Bender, D. B., & Rocha-Miranda, C. E. d. (1969). Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science*, 1303-1306.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *Journal of Neurophysiology*, 35(1), 96-111.
- Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., & Vuilleumier, P. (2012). White-matter connectivity between face-responsive regions in the human brain. *Cerebral Cortex*, 22(7), 1564-1576.

- Guellaï, B., Coulon, M., & Streri, A. (2011). The role of motion and speech in face recognition at birth. *Visual Cognition*, 19(9), 1212-1233.
- Gunning-Dixon, F. M., Gur, R. C., Perkins, A. C., Schroeder, L., Turner, T., Turetsky, B. I., Chan, R. M., Loughhead, J. W., Alsop, D. C., & Maldjian, J. (2003). Age-related differences in brain activation during emotional face processing. *Neurobiology of Aging*, 24(2), 285-295.
- Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., Turner, T., Bajcsy, R., Posner, A., & Gur, R. E. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods*, 115(2), 137-143.
- Hamann, S. B., Ely, T. D., Hoffman, J. M., & Kilts, C. D. (2002). Ecstasy and agony: Activation of the human amygdala in positive and negative emotion. *Psychological Science*, 13(2), 135-141.
- Hamm, J., Kohler, C. G., Gur, R. C., & Verma, R. (2011). Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders. *Journal of Neuroscience Methods*, 200(2), 237-256.
- Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology*, 50(4), 354-377.
- Harwood, N. K., Hall, L. J., & Shinkfield, A. J. (1999). Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. *American Journal on Mental Retardation*, 104(3), 270-278.
- Hasan, B. A. S., Valdes-Sosa, M., Gross, J., & Belin, P. (2016). "Hearing faces and seeing voices": Amodal coding of person identity in the human brain. *Scientific Reports*, 6, 1-8.
- Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioral Brain Research*, 32, 203-218.
- Hauser, M. D. (1993). Right hemisphere dominance for the production of facial expression in monkeys. *Science*, 261(5120), 475-477.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223-233.
- Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, 22(1), 189-199.

- He, Y., Ebner, N. C., & Johnson, M. K. (2011). What predicts the own-age bias in face recognition memory? *Social Cognition*, 29(1), 97-109.
- Hecaen, H., & Angelergues, R. (1962). Agnosia for faces (prosopagnosia). *Archives of Neurology*, 7, 92-100.
- Hemond, C. C., Kanwisher, N. G., & De Beeck, H. P. O. (2007). A preference for contralateral stimuli in human object-and face-selective cortex. *PLoS One*, 2(6), 1-5.
- Herba, C., Landau, S., Russell, T., Ecker, C., & Phillips, M. L. (2006). The development of emotion-processing in children: Effects of age, emotion, and intensity. *Journal of Child Psychology and Psychiatry*, 47(11), 1098-1106.
- Herba, C., & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: Behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry*, 45(7), 1185-1198.
- Hillairet de Boisferon, A., Dupierriex, E., Quinn, P. C., Løevenbruck, H., Lewkowicz, D. J., Lee, K., & Pascalis, O. (2015). Perception of multisensory gender coherence in 6-and 9-month-old infants. *Infancy*, 20(6), 661-674.
- Hilliard, R. D. (1973). Hemispheric laterality effects on a facial recognition task in normal subjects. *Cortex*, 9(3), 246-258.
- Hills, P. J., & Lewis, M. B. (2011). The own-age face recognition bias in children and adults. *The Quarterly Journal of Experimental Psychology*, 64(1), 17-23.
- Hobson, R. P., Ouston, J., & Lee, A. (1988). What's in a face? The case of autism. *British Journal of Psychology*, 79(4), 441-453.
- Hof, P. R., & Morrison, J. H. (2004). The aging brain: Morphomolecular senescence of cortical circuits. *Trends in Neurosciences*, 27(10), 607-613.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, 3(1), 80-84.
- Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory*, 7(6), 418-439.
- Horning, S. M., Cornwell, R. E., & Davis, H. P. (2012). The recognition of facial expressions: An investigation of the influence of age and cognition. *Aging, Neuropsychology, and Cognition*, 19(6), 657-676.

- Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, 13(4), 341-348.
- Humphreys, G. W., Donnelly, N., & Riddoch, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces: Neuropsychological evidence. *Neuropsychologia*, 31(2), 173-181.
- Humphreys, K., Gosselin, F., Schyns, P. G., & Johnson, M. H. (2006). Using “Bubbles” with babies: A new technique for investigating the informational basis of infant perception. *Infant Behavior and Development*, 29(3), 471-475.
- Isaacowitz, D. M., Löckenhoff, C. E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging*, 22(1), 147-159.
- Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- Jack, C. R., Petersen, R. C., Xu, Y. C., Waring, S. C., O'brien, P. C., Tangalos, E. G., Smith, G. E., Ivnik, R. J., & Kokmen, E. (1997). Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology*, 49(3), 786-794.
- Jack, R. E. (2013). Culture and facial expressions of emotion. *Visual Cognition*, 21(9-10), 1248-1286.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology*, 19(18), 1543-1548.
- Jack, R. E., Caldara, R., & Schyns, P. G. (2012a). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology General*, 141(1), 19-25.
- Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012b). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences, USA*, 109(19), 7241-7244.
- Jack, R. E., Garrod, O. G., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, 24(2), 187-192.
- Jack, R. E., & Schyns, P. G. (2015). The human face as a dynamic tool for social communication. *Current Biology*, 25(14), R621-R634.
- James, W. (1884). What is an emotion? *Mind*, 9, 188-205.
- Jiang, F., Dricot, L., Weber, J., Righi, G., Tarr, M. J., Goebel, R., & Rossion, B. (2011). Face categorization in visual scenes may start in a higher order area of the right fusiform

- gyrus: Evidence from dynamic visual stimulation in neuroimaging. *Journal of Neurophysiology*, 106(5), 2720-2736.
- Jiang, Z., Li, W., Recio, G., Liu, Y., Luo, W., Zhang, D., & Sun, D. (2014). Time pressure inhibits dynamic advantage in the classification of facial expressions of emotion. *PLoS One*, 9(6), 1-7.
- Johnson, M. H. (2011). Face perception: A developmental perspective. In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 3-14). Oxford: Oxford University Press.
- Johnson, M. H., Senju, A., & Tomalski, P. (2015). The two-process theory of face processing: Modifications based on two decades of data from infants and adults. *Neuroscience & Biobehavioral Reviews*, 50, 169-179.
- Johnston, P., Mayes, A., Hughes, M., & Young, A. W. (2013). Brain networks subserving the evaluation of static and dynamic facial expressions. *Cortex*, 49(9), 2462-2472.
- Jonas, J., Rossion, B., Brissart, H., Frismand, S., Jacques, C., Hossu, G., Colnat-Coulbois, S., Vespignani, H., Vignal, J.-P., & Maillard, L. (2015). Beyond the core face-processing network: Intracerebral stimulation of a face-selective area in the right anterior fusiform gyrus elicits transient prosopagnosia. *Cortex*, 72, 140-155.
- Kalanithi, P. S., & Henderson, J. M. (2012). Optogenetic neuromodulation. *International Review of Neurobiology*, 107, 185-205.
- Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, 30(7), 875-887.
- Kan, Y., Mimura, M., Kamijima, K., & Kawamura, M. (2004). Recognition of emotion from moving facial and prosodic stimuli in depressed patients. *Journal of Neurology Neurosurgery and Psychiatry*, 75(12), 1667-1671.
- Kanade, T., Cohn, J. F., & Tian, Y. (2000). *Comprehensive database for facial expression analysis*. Paper presented at Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (pp. 110-115), Grenoble (France).
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11), 4302-4311.
- Kaplan, P. S., Bachorowski, J.-A., Smoski, M. J., & Hudenko, W. J. (2002). Infants of depressed mothers, although competent learners, fail to learn in response to their own mothers' infant-directed speech. *Psychological Science*, 13(3), 268-271.

- Kaplan, P. S., Bachorowski, J. A., & Zarlengo-Strouse, P. (1999). Child-directed speech produced by mothers with symptoms of depression fails to promote associative learning in 4-month-old infants. *Child Development*, 70(3), 560-570.
- Kätsyri, J. (2006). *Human recognition of basic emotions from posed and animated dynamic facial expressions*. Unpublished doctoral dissertation. Helsinki University of Technology. Pdf found online:
<https://aaltodoc.aalto.fi/bitstream/handle/123456789/2807/isbn951228538X.pdf?sequence=1&isAllowed=y>.
- Kätsyri, J., Saalasti, S., Tiippana, K., von Wendt, L., & Sams, M. (2008). Impaired recognition of facial emotions from low-spatial frequencies in Asperger syndrome. *Neuropsychologia*, 46(7), 1888-1897.
- Kätsyri, J., & Sams, M. (2008). The effect of dynamics on identifying basic emotions from synthetic and natural faces. *International Journal of Human-Computer Studies*, 66(4), 233-242.
- Keating, P. A., Lindblom, B., Lubker, J., & Kreiman, J. (1994). Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics*, 22(4), 407-422.
- Keightley, M. L., Chiew, K. S., Winocur, G., & Grady, C. L. (2007). Age-related differences in brain activity underlying identification of emotional expressions in faces. *Social Cognitive and Affective Neuroscience*, 2(4), 292-302.
- Kellendonk, C., Simpson, E. H., Polan, H. J., Malleret, G., Vronskaya, S., Winiger, V., Moore, H., & Kandel, E. R. (2006). Transient and selective overexpression of dopamine D2 receptors in the striatum causes persistent abnormalities in prefrontal cortex functioning. *Neuron*, 49(4), 603-615.
- Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality? *Journal of Experimental Child Psychology*, 104(1), 105-114.
- Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science*, 18(12), 1084-1089.
- Kelso, J. s., Vatikiotis-Bateson, E., Saltzman, E. L., & Kay, B. (1985). A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics, and dynamic modeling. *The Journal of the Acoustical Society of America*, 77(1), 266-280.
- Kessler, H., Doyen-Waldecker, C., Hofer, C., Hoffmann, H., Traue, H. C., & Abler, B. (2011). Neural correlates of the perception of dynamic versus static facial expressions of emotion. *Psychosocial Medicine*, 8, 1-8.

- Killgore, W. D., & Yurgelun-Todd, D. A. (2004). Activation of the amygdala and anterior cingulate during nonconscious processing of sad versus happy faces. *NeuroImage*, 21(4), 1215-1223.
- Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage*, 18(1), 156-168.
- Kim, H. I., & Johnson, S. P. (2013). Do young infants prefer an infant-directed face or a happy face? *International Journal of Behavioral Development*, 37(2), 125-130.
- Kim, H. I., & Johnson, S. P. (2014). Detecting 'infant-directedness' in face and voice. *Developmental Science*, 17(4), 621-627.
- Kim, M., Ducros, M., Carlson, T., Ronen, I., He, S., Ugurbil, K., & Kim, D.-S. (2006). Anatomical correlates of the functional organization in the human occipitotemporal cortex. *Magnetic Resonance Imaging*, 24(5), 583-590.
- Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, 87(2), 820-857.
- Klavir, O., Prigge, M., Sarel, A., Paz, R., & Yizhar, O. (2017). Manipulating fear associations via optogenetic modulation of amygdala inputs to prefrontal cortex. *Nature Neuroscience*, 20(6), 836-844.
- Kleiner, K. A. (1987). Amplitude and phase spectra as indices of infants' pattern preferences. *Infant Behavior and Development*, 10(1), 49-59.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1.
- Knappmeyer, B., Thornton, I. M., & Bulthoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. *Vision Research*, 43(18), 1921-1936.
- Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition*, 4(3), 265-273.
- Kohler, C. G., Bilker, W., Hagendoorn, M., Gur, R. E., & Gur, R. C. (2000). Emotion recognition deficit in schizophrenia: Association with symptomatology and cognition. *Biological Psychiatry*, 48(2), 127-136.

- Kohler, C. G., Martin, E. A., Stolar, N., Barrett, F. S., Verma, R., Brensinger, C., Bilker, W., Gur, R. E., & Gur, R. C. (2008). Static posed and evoked facial expressions of emotions in schizophrenia. *Schizophrenia Research*, 105(1-3), 49-60.
- Kohler, C. G., Turner, T., Stolar, N. M., Bilker, W. B., Brensinger, C. M., Gur, R. E., & Gur, R. C. (2004a). Differences in facial expressions of four universal emotions. *Psychiatry Research*, 128(3), 235-244.
- Kohler, C. G., Turner, T. H., Bilker, W. B., Brensinger, C. M., Siegel, S. J., Kanes, S. J., Gur, R. E., & Gur, R. C. (2003). Facial emotion recognition in schizophrenia: Intensity effects and error pattern. *The American Journal of Psychiatry*, 160(10), 1768-1774.
- Kohler, C. G., Turner, T. H., Gur, R. E., & Gur, R. C. (2004b). Recognition of facial emotions in neuropsychiatric disorders. *CNS Spectrums*, 9(4), 267-274.
- Kokaia, M., Andersson, M., & Ledri, M. (2013). An optogenetic approach in epilepsy. *Neuropharmacology*, 69, 89-95.
- Kolb, B., Wilson, B., & Taylor, L. (1992). Developmental changes in the recognition and comprehension of facial expression: Implications for frontal lobe function. *Brain and Cognition*, 20(1), 74-84.
- Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, 30(9), 1115-1125.
- Krendl, A. C., & Ambady, N. (2010). Older adults' decoding of emotions: Role of dynamic versus static cues and age-related cognitive decline. *Psychology and Aging*, 25(4), 788-793.
- Krone, T., Albers, C. J., Kuppens, P., & Timmerman, M. E. (2017). A model for emotion dynamics: A multivariate statistical model for emotion dynamics. *Emotion*.
- Krook-Magnuson, E., Armstrong, C., Oijala, M., & Soltesz, I. (2013). On-demand optogenetic control of spontaneous seizures in temporal lobe epilepsy. *Nature Communications*, 4, 1376.
- Krumhuber, E. G., Kappas, A., & Manstead, A. S. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review*, 5(1), 41-46.
- Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan* (2nd). New York: Academic Press.
- Kubicek, C., De Boisferon, A. H., Dupierrix, E., Pascalis, O., Lœvenbruck, H., Gervain, J., & Schwarzer, G. (2014). Cross-modal matching of audio-visual German and French fluent speech in infancy. *PLoS One*, 9(2), 1-11.

- Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of Experimental Psychology: Human Perception and Performance*, 34(4), 811-817.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831-843.
- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, 218(4577), 1138-1141.
- Kuhl, P. K., & Meltzoff, A. N. (1988). Speech as an intermodal object of perception. In A. Yonas (Ed.), *Perceptual development in infancy: The Minnesota symposia on child psychology* (pp. 235-266). Hillsdale, NJ: Erlbaum.
- Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America*, 100(4), 2425-2438.
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9(2), F13-F21.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 606-608.
- LaBar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, 13(10), 1023-1033.
- LaBarbera, J., Izard, C., Vietze, P., & Parisi, S. (1976). Four-and six-month-old infants' visual responses to joy, anger, and neutral expressions. *Child Development*, 535-538.
- Lai, J., Pancaroglu, R., Oruc, I., Barton, J. J., & Davies-Thompson, J. (2014). Neuroanatomic correlates of the feature-salience hierarchy in face processing: An fMRI-adaptation study. *Neuropsychologia*, 53, 274-283.
- Lammel, S., Tye, K. M., & Warden, M. R. (2014). Progress in understanding mood disorders: Optogenetic dissection of neural circuits. *Genes Brain and Behavior*, 13(1), 38-51.
- Lander, K., & Bruce, V. (2000). Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology*, 12(4), 259-272.
- Lander, K., & Bruce, V. (2003). The role of motion in learning new faces. *Visual Cognition*, 10(8), 897-912.

- Lander, K., Bruce, V., & Hill, H. (2001). Evaluating the effectiveness of pixelation and blurring on masking the identity of familiar faces. *Applied Cognitive Psychology*, 15(1), 101-116.
- Lander, K., & Butcher, N. (2015). Independence of face identity and expression processing: exploring the role of motion. *Frontiers in Psychology*, 6, 1-6.
- Landis, T., Regard, M., Bliestle, A., & Kleihues, P. (1988). Prosopagnosia and agnosia for noncanonical views. An autopsied case. *Brain*, 111(Pt 6), 1287-1297.
- Langlois, J. H., Roggman, L. A., Casey, R. J., Ritter, J. M., Rieser-Danner, L. A., & Jenkins, V. Y. (1987). Infant preferences for attractive faces: Rudiments of a stereotype. *Developmental Psychology*, 23, 363-369.
- Larochette, A. C., Chambers, C. T., & Craig, K. D. (2006). Genuine, suppressed and faked facial expressions of pain in children. *Pain*, 126(1-3), 64-71.
- Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2003). Expert face processing requires visual input to the right hemisphere during infancy. *Nature Neuroscience*, 6(10), 1108-1112.
- Lee, K., Anzures, G., Quinn, P. C., Pascalis, O., & Slater, A. (2011). Development of face processing expertise. In G. Rhodes, A. Calder, M. Johnson & J. V. Haxby (Eds.), *Oxford handbook of face perception* (pp. 753-778). New York: Oxford University Press.
- Legerstee, M. (1990). Infants use multimodal information to imitate speech sounds. *Infant Behavior and Development*, 13(3), 343-354.
- Leinbach, M. D., & Fagot, B. I. (1993). Categorical habituation to male and female faces: Gender schematic processing in infancy. *Infant Behavior and Development*, 16(3), 317-332.
- Leitzke, B. T., & Pollak, S. D. (2016). Developmental changes in the primacy of facial cues for emotion recognition. *Developmental Psychology*, 52(4), 572-581.
- Lennox, B., Jacob, R., Calder, A., Lupson, V., & Bullmore, E. (2004). Behavioural and neurocognitive responses to sad facial affect are attenuated in patients with mania. *Psychological Medicine*, 34(5), 795-802.
- Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience*, 10(1), 37-47.
- Leroy, F., Glasel, H., Dubois, J., Hertz-Pannier, L., Thirion, B., Mangin, J. F., & Dehaene-Lambertz, G. (2011). Early maturation of the linguistic dorsal pathway in human infants. *Journal of Neuroscience*, 31(4), 1500-1506.

- Lewkowicz, D. J. (1998). Infants' response to the audible and visible properties of the human face: II. Discrimination of differences between singing and adult-directed speech. *Developmental Psychobiology*, 32(4), 261-274.
- Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin*, 126(2), 281-308.
- Lewkowicz, D. J. (2010). Infant perception of audio-visual speech synchrony. *Developmental Psychology*, 46(1), 66-77.
- Lewkowicz, D. J., Minar, N. J., Tift, A. H., & Brandon, M. (2015). Perception of the multisensory coherence of fluent audiovisual speech in infancy: Its emergence and the role of experience. *Journal of Experimental Child Psychology*, 130, 147-162.
- Lewkowicz, D. J., & Pons, F. (2013). Recognition of amodal language identity emerges in infancy. *International Journal of Behavioral Development*, 37(2), 90-94.
- Liang, Y., Liu, B., Xu, J., Zhang, G., Li, X., Wang, P., & Wang, B. (2017). Decoding facial expressions based on face-selective and motion-sensitive areas. *Human Brain Mapping*, 38(6), 3113-3125.
- Lints-Martindale, A. C., Hadjistavropoulos, T., Barber, B., & Gibson, S. J. (2007). A psychophysical investigation of the facial action coding system as an index of pain variability among older adults with and without Alzheimer's disease. *Pain Med*, 8(8), 678-689.
- Liu, H. M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, 6(3), F1-F10.
- Liu, S., Xiao, N. G., Quinn, P. C., Zhu, D., Ge, L., Pascalis, O., & Lee, K. (2015). Asian infants show preference for own-race but not other-race female faces: The role of infant caregiving arrangements. *Frontiers in Psychology*, 6, 1-8.
- LoBue, V. (2010). What's so scary about needles and knives? Examining the role of experience in threat detection. *Cognition and Emotion*, 24(1), 180-187.
- Lœvenbruck, H. (1999). *An investigation of articulatory correlates of the accentual phrase in French*. Paper presented at the Proceedings of the 14th International Congress of Phonetic Sciences (pp. 667-670), San Francisco (USA).
- Lœvenbruck, H. (2000). *Effets articulatoires de l'emphase contrastive sur la phrase accentuelle en français*. Paper presented at the Actes des XXIIIèmes JEP (pp. 165-168), Aussois (France).

- Macchi Cassia, V., Simion, F., & Umiltà, C. (2001). Face preference at birth: The role of an orienting mechanism. *Developmental Science*, 4(1), 101-108.
- Macchi Cassia, V., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science*, 15(6), 379-383.
- MacKain, K., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983). Infant intermodal speech perception is a left-hemisphere function. *Science*, 219(4590), 1347-1349.
- MacPherson, S. E., Phillips, L. H., & Della Sala, S. (2002). Age, executive function and social decision making: A dorsolateral prefrontal theory of cognitive aging. *Psychology and Aging*, 17(4), 598-609.
- Malatesta, C. Z., Izard, C. E., Culver, C., & Nicolich, M. (1987). Emotion communication skills in young, middle-aged, and older women. *Psychology and Aging*, 2(2), 193-203.
- Malone, D. R., Morris, H. H., Kay, M. C., & Levin, H. S. (1982). Prosopagnosia: A double dissociation between the recognition of familiar and unfamiliar faces. *Journal of Neurology Neurosurgery and Psychiatry*, 45(9), 820-822.
- Mancini, G., Agnoli, S., Baldaro, B., Bitti, P. E., & Surcinelli, P. (2013). Facial expressions of emotions: Recognition accuracy and affective reactions during late childhood. *Journal of Psychology*, 147(6), 599-617.
- Martin, F., Baudouin, J. Y., Tiberghien, G., & Franck, N. (2005). Processing emotional expression and facial identity in Schizophrenia. *Psychiatry Research*, 134, 43-53.
- Mather, M., Canli, T., English, T., Whitfield, S., Wais, P., Ochsner, K., John, D. G., & Carstensen, L. L. (2004). Amygdala responses to emotionally valenced stimuli in older and younger adults. *Psychological Science*, 15(4), 259-263.
- Matsumoto, D., & Ekman, P. (1988). Japanese and caucasian facial expressions of emotion (jafec) and neutral faces. San Francisco, CA: Intercultural and Emotion Research Laboratory, Department of Psychology, San Francisco State University.
- Mattson, A. J., Levin, H. S., & Grafman, J. (2000). A case of prosopagnosia following moderate closed head injury with left hemisphere focal lesion. *Cortex*, 36(1), 125-137.
- Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, 6(6), 255-260.
- McCall, R. B., & Kagan, J. (1970). Individual differences in the infant's distribution of attention to stimulus discrepancy. *Developmental Psychology*, 2(1), 90-98.

- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, 9(5), 605-610.
- McCotter, M., Gosselin, F., Sowden, P., & Schyns, P. G. (2005). The use of visual information in natural scenes. *Vision Cognition*, 12(6), 938-953.
- McDougall, W. (1926). *An introduction to social psychology*. Boston: Luce.
- McDowell, C. L., Harrison, D. W., & Demaree, H. A. (1994). Is right hemisphere decline in the perception of emotion a function of aging? *International Journal of Neuroscience*, 79(1-2), 1-11.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746-748.
- McKone, E., & Robbins, R. (2011). Are faces special? In G. Rhodes, A. Calder, M. Johnson & J. V. Haxby (Eds.), *Oxford handbook of face perception* (pp. 149-176). New York: Oxford University Press.
- Meadows, J. (1974). The anatomical basis of prosopagnosia. *Journal of Neurology, Neurosurgery & Psychiatry*, 37(5), 489-501.
- Memon, A., Bartlett, J., Rose, R., & Gray, C. (2003). The aging eyewitness: Effects of age on face, delay, and source-memory ability. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 58(6), P338-P345.
- Mendoza, R., Cabral-Calderin, Y., Dominguez, M., Garcia, A., Borrego, M., Caballero, A., Guerra, S., & Reyes, M. M. (2011). Impairment of emotional expression recognition in schizophrenia: A Cuban familial association study. *Psychiatry Research*, 185(1-2), 44-48.
- Miellet, S., Caldara, R., & Schyns, P. G. (2011). Local Jekyll and global Hyde: The dual identity of face identification. *Psychological Science*, 22(12), 1518-1526.
- Miellet, S., He, L., Zhou, X., Lao, J., & Caldara, R. (2012). When East meets West: Gaze-contingent blindspots abolish cultural diversity in eye movements for faces. *Journal of Eye Movement Research*, 5(2), 1-12.
- Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion*, 9(5), 619-630.
- Miller, M. I., Younes, L., Ratnanather, J. T., Brown, T., Trinh, H., Lee, D. S., Tward, D., Mahon, P. B., Mori, S., & Albert, M. (2015). Amygdalar atrophy in symptomatic Alzheimer's disease based on diffeomorphometry: The BIOCARD cohort. *Neurobiology of Aging*, 36, S3-S10.

-
- Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, 46(3), 774-785.
- Mitchell, A. J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research*, 43(4), 411-431.
- Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., & Kleiner-Gathercoal, K. A. (1999). Face perception during early infancy. *Psychological Science*, 10(5), 419-422.
- Mondloch, C. J., Maurer, D., & Ahola, S. (2006). Becoming a face expert. *Psychological Science*, 17(11), 930-934.
- Montague, D. P., & Walker-Andrews, A. S. (2001). Peekaboo: A new look at infants' perception of emotion expressions. *Developmental Psychology*, 37(6), 826-838.
- Montague, D. P., & Walker-Andrews, A. S. (2002). Mothers, fathers, and infants: The role of person familiarity and parental involvement in infants' perception of emotion expressions. *Child Development*, 73(5), 1339-1352.
- Montiroso, R., Peverelli, M., Frigerio, E., Crespi, M., & Borgatti, R. (2010). The development of dynamic facial expression recognition at different intensities in 4- to 18-year-olds. *Social Development*, 19(1), 71-92.
- Moreno, C., Borod, J. C., Welkowitz, J., & Alpert, M. (1993). The perception of facial emotion across the adult life span. *Developmental Neuropsychology*, 9(3-4), 305-314.
- Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, 98(2), 164-181.
- Mowrer, O. H. (1960). *Learning theory and behavior*. New York: Wiley.
- Murphy, F. C., Nimmo-Smith, I., & Lawrence, A. D. (2003). Functional neuroanatomy of emotions: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, 3(3), 207-233.
- Nakano, T., & Nakatani, K. (2014). Cortical networks for face perception in two-month-old infants. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1793), 1-5.
- Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development*, 27(4), 455-464.
- Nakato, E., Otsuka, Y., Kanazawa, S., Yamaguchi, M. K., & Kakigi, R. (2011). Distinct differences in the pattern of hemodynamic response to happy and angry facial

- expressions in infants — A near-infrared spectroscopic study. *NeuroImage*, 54(2), 1600-1606.
- Naoi, N., Minagawa-Kawai, Y., Kobayashi, A., Takeuchi, K., Nakamura, K., Yamamoto, J., & Kojima, S. (2012). Cerebral responses to infant-directed speech and the effect of talker familiarity. *NeuroImage*, 59(2), 1735-1744.
- Narumoto, J., Okada, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Brain Research Cognitive Brain Research*, 12(2), 225-231.
- Navarro, C., & Gonzalo, L. M. (2017). Cambios en el complejo amigdalino humano debidos a la edad. *Revista de Medicina de la Universidad de Navarra*, 7-12.
- Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development*, 10(1-2), 3-18.
- Nelson, C. A., & De Haan, M. (1996). Neural correlates of infants' visual responsiveness to facial expressions of emotion. *Developmental Psychobiology*, 29(7), 577-595.
- Nelson, N. L., Hudspeth, K., & Russell, J. A. (2013). A story superiority effect for disgust, fear, embarrassment, and pride. *British Journal of Developmental Psychology*, 31(3), 334-348.
- Nelson, N. L., & Russell, J. A. (2011). Putting motion in emotion: Do dynamic presentations increase preschooler's recognition of emotion? *Cognitive Development*, 26(3), 248-259.
- Nesse, R. M., & Ellsworth, P. C. (2009). Evolution, emotions, and emotional disorders. *American Psychologist*, 64(2), 129-139.
- Newman, R. S., & Hussain, I. (2006). Changes in Preference for Infant-Directed Speech in Low and Moderate Noise by 4.5-to 13-Month-Olds. *Infancy*, 10(1), 61-76.
- O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences*, 6(6), 261-266.
- O'Toole, D., & Roark, D. A. (2010). Memory for moving faces: The interplay of two recognition systems. In C. Curio, H. H. Bülthoff & M. Giese (Eds.), *Dynamic faces: Insights from experiments and computation*. Cambridge, MA: MIT Press.
- O'Toole, A. J., Deffenbacher, K. A., Valentin, D., McKee, K., Huff, D., & Abdi, H. (1998). The perception of face gender: The role of stimulus structure in recognition and classification. *Memory & Cognition*, 26(1), 146-160.

-
- Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion*, 1(1), 29-50.
- Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., & Gross, J. J. (2004). For better or for worse: Neural systems supporting the cognitive down-and up-regulation of negative emotion. *NeuroImage*, 23(2), 483-499.
- Orban de Xivry, J. J., Ramon, M., Lefevre, P., & Rossion, B. (2008). Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *Journal of Neuropsychology*, 2(1), 245-268.
- Orgeta, V., & Phillips, L. H. (2007). Effects of age and emotional intensity on the recognition of facial emotion. *Experimental Aging Research*, 34(1), 63-79.
- Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97(3), 315-331.
- Otsuka, Y., Konishi, Y., Kanazawa, S., Yamaguchi, M. K., Abdi, H., & O'Toole, A. J. (2009). Recognition of moving and static faces by young infants. *Child Development*, 80(4), 1259-1271.
- Palermo, R., O'Connor, K. B., Davis, J. M., Irons, J., & McKone, E. (2013). New tests to measure individual differences in matching and labelling facial expressions of emotion, and their association with ability to recognise vocal emotions and facial identity. *PLoS One*, 8(6), 1-16.
- Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia*, 49(5), 1226-1235.
- Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *Behavioral and Brain Sciences*, 5(3), 407-422.
- Papoušek, M., Papoušek, H., & Haekel, M. (1987). Didactic adjustments in fathers' and mothers' speech to their 3-month-old infants. *Journal of Psycholinguistic Research*, 16(5), 491-516.
- Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behavior and Development*, 14(4), 415-440.
- Park, B., Tsai, J. L., Chim, L., Blevins, E., & Knutson, B. (2016). Neural evidence for cultural differences in the valuation of positive facial expressions. *Social Cognitive and Affective Neuroscience*, 11(2), 243-252.
- Pascalis, O., & Bachevalier, J. (1998). Face recognition in primates: A cross-species study. *Behavioural Processes*, 43(1), 87-96.

- Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, 296(5571), 1321-1323.
- Pascalis, O., De Haan, M., Nelson, C. A., & De Schonen, S. (1998). Long-term recognition memory for faces assessed by visual paired comparison in 3-and 6-month-old infants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1), 249-260.
- Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and an extension. *Infant Behavior and Development*, 18(1), 79-85.
- Pascalis, O., Scott, L. S., Kelly, D., Shannon, R., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences, USA*, 102(14), 5297-5300.
- Patterson, M. L., & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior and Development*, 22(2), 237-247.
- Patterson, M. L., & Werker, J. F. (2002). Infants' ability to match dynamic phonetic and gender information in the face and voice. *Journal of Experimental Child Psychology*, 81(1), 93-115.
- Paulmann, S., Jessen, S., & Kotz, S. A. (2009). Investigating the multimodal nature of human communication: An ERP study. *Journal of Psychophysiology*, 23(2), 63-76.
- Peelen, M. V., Lucas, N., Mayer, E., & Vuilleumier, P. (2009). Emotional attention in acquired prosopagnosia. *Social Cognitive and Affective Neuroscience*, 4(3), 268-277.
- Peelen, M. V., Wiggett, A. J., & Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron*, 49(6), 815-822.
- Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development*, 15(3), 325-345.
- Peltola, M. J., Leppänen, J. M., Mäki, S., & Hietanen, J. K. (2009). Emergence of enhanced attention to fearful faces between 5 and 7 months of age. *Social Cognitive and Affective Neuroscience*, 4(2), 134-142.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47(3), 329-342.
- Pesenti, M., Thioux, M., Samson, D., Bruyer, R., & Seron, X. (2000). Number processing and calculation in a case of visual agnosia. *Cortex*, 36(3), 377-400.

- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9(2), 148-158.
- Peter, V., Kalashnikova, M., Santos, A., & Burnham, D. (2016). Mature neural responses to infant-directed speech but not adult-directed speech in pre-verbal infants. *Scientific Reports*, 6, 1-14.
- Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, 16(2), 331-348.
- Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, 17(4), 292-299.
- Phillips, L. H., MacLean, R. D., & Allen, R. (2002). Age and the understanding of emotions: Neuropsychological and sociocognitive perspectives. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(6), P526-P530.
- Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biological Psychiatry*, 54(5), 504-514.
- Pinel, P., Lalanne, C., Bourgeron, T., Fauchereau, F., Poupon, C., Artiges, E., Le Bihan, D., Dehaene-Lambertz, G., & Dehaene, S. (2014). Genetic and environmental influences on the visual word form and fusiform face areas. *Cerebral Cortex*, 25(9), 2478-2493.
- Pinsk, M. A., Arcaro, M., Weiner, K. S., Kalkus, J. F., Inati, S. J., Gross, C. G., & Kastner, S. (2009). Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. *Journal of Neurophysiology*, 101(5), 2581-2600.
- Pitcher, D. (2014). Facial expression recognition takes longer in the posterior superior temporal sulcus than in the occipital face area. *Journal of Neuroscience*, 34(27), 9173-9177.
- Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage*, 56(4), 2356-2363.
- Pitcher, D., Duchaine, B., & Walsh, V. (2014). Combined TMS and fMRI reveal dissociable cortical pathways for dynamic and static face perception. *Current Biology*, 24(17), 2066-2070.
- Pitcher, D., Garrido, L., Walsh, V., & Duchaine, B. C. (2008). TMS disrupts the perception and embodiment of facial expressions. *Journal of Neuroscience*, 28, 8929-8933.

- Pitcher, D., Goldhaber, T., Duchaine, B., Walsh, V., & Kanwisher, N. (2012). Two critical and functionally distinct stages of face and body perception. *Journal of Neuroscience*, 32(45), 15877-15885.
- Pitcher, D., Japee, S., Rauth, L., & Ungerleider, L. G. (2017). The superior temporal sulcus is causally connected to the amygdala: A combined TBS-fMRI study. *Journal of Neuroscience*, 37(5), 1156-1161.
- Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology*, 17(18), 1568-1573.
- Plutchik, R. (1980a). *Emotion: A psychoevolutionary synthesis*. New York: Harper & Row.
- Plutchik, R. (1980b). The emotions: Facts, theories, and a new model. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research and experience: Vol. 1. Theories of emotion* (pp. 3-31). New York: Academic Press.
- Plutchik, R. (2003). *Emotions and life: Perspectives from psychology, biology, and evolution*. Washington, DC: American Psychological Association.
- Pochedly, J. T., Widen, S. C., & Russell, J. A. (2012). What emotion does the “facial expression of disgust” express? *Emotion*, 12(6), 1315-1319.
- Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences, USA*, 99(13), 9072-9076.
- Pollak, S. D., & Sinha, P. (2002). Effects of early experience on children's recognition of facial displays of emotion. *Developmental Psychology*, 38(5), 784-791.
- Poulin-Dubois, D., Serbin, L. A., & Derbyshire, A. (1998). Toddlers' intermodal and verbal knowledge about gender. *Merrill-Palmer Quarterly*, 44, 338-354.
- Poulin-Dubois, D., Serbin, L. A., Kenyon, B., & Derbyshire, A. (1994). Infants' intermodal knowledge about gender. *Developmental Psychology*, 30(3), 436-442.
- Pourtois, G., de Gelder, B., Bol, A., & Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex*, 41(1), 49-59.
- Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience*, 16(16), 5205-5215.

- Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, 18(6), 2188-2199.
- Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 358(1431), 435-445.
- Pyles, J. A., Verstynen, T. D., Schneider, W., & Tarr, M. J. (2013). Explicating the face perception network with white matter connectivity. *PLoS One*, 8(4), 1-12.
- Quaglino, A. (1867). Emiplegia sinistra con amaurosi— Guarigione—Perdita totale della percezione dei colori e della memoria della configurazione degli oggetti. *Giornale d'Oftalmologia Italiano*, 10, 106-112.
- Quinn, P. C., & Eimas, P. D. (1986). On categorization in early infancy. *Merrill-Palmer Quarterly*, 32, 331-363.
- Quinn, P. C., Kelly, D. J., Lee, K., Pascalis, O., & Slater, A. M. (2008a). Preference for attractive faces in human infants extends beyond conspecifics. *Developmental Science*, 11(1), 76-83.
- Quinn, P. C., Uttley, L., Lee, K., Gibson, A., Smith, M., Slater, A. M., & Pascalis, O. (2008b). Infant preference for female faces occurs for same-but not other-race faces. *Journal of Neuropsychology*, 2(1), 15-26.
- Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, 31(9), 1109-1121.
- Ramon, M., Busigny, T., & Rossion, B. (2010). Impaired holistic processing of unfamiliar individual faces in acquired prosopagnosia. *Neuropsychologia*, 48(4), 933-944.
- Ramsey, J. L., Langlois, J. H., & Marti, N. C. (2005). Infant categorization of faces: Ladies first. *Developmental Review*, 25(2), 212-246.
- Ramsey-Rennels, J. L., & Langlois, J. H. (2006). Infants' differential processing of female and male faces. *Current Directions in Psychological Science*, 15(2), 59-62.
- Rangarajan, V., Hermes, D., Foster, B. L., Weiner, K. S., Jacques, C., Grill-Spector, K., & Parvizi, J. (2014). Electrical stimulation of the left and right human fusiform gyrus causes different effects in conscious face perception. *Journal of Neuroscience*, 34(38), 12828-12836.
- Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 1-90). Mahwah, NJ: Lawrence Erlbaum Associates.

- Recio, G., Schacht, A., & Sommer, W. (2013). Classification of dynamic facial expressions of emotion presented briefly. *Cognition & Emotion*, 27(8), 1486-1494.
- Recio, G., Sommer, W., & Schacht, A. (2011). Electrophysiological correlates of perceiving and evaluating static and dynamic facial emotional expressions. *Brain Research*, 1376, 66-75.
- Reed, L. I., Sayette, M. A., & Cohn, J. F. (2007). Impact of depression on response to comedy: A dynamic facial coding analysis. *Journal of Abnormal Psychology*, 116(4), 804-809.
- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The human fetus preferentially engages with face-like visual stimuli. *Current Biology*, 27(13), 1825-1828.
- Rennels, J. L., & Davis, R. E. (2008). Facial experience during the first year. *Infant Behavior and Development*, 31(4), 665-678.
- Richoz, A.-R., Jack, R. E., Garrod, O. G., Schyns, P. G., & Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. *Cortex*, 65, 50-64.
- Richoz, A.-R., Quinn, P. C., de Boisferon, A. H., Berger, C., Loevenbruck, H., Lewkowicz, D. J., Lee, K., Dole, M., Caldara, R., & Pascalis, O. (2017). Audio-visual perception of gender by infants emerges earlier for adult-directed speech. *PLoS One*, 12(1), 1-15.
- Rochat, P., Striano, T., & Blatt, L. (2002). Differential effects of happy, neutral, and sad still-faces on 2-, 4-and 6-month-old infants. *Infant and Child Development*, 11(4), 289-303.
- Rodger, H., Vizioli, L., Ouyang, X., & Caldara, R. (2015). Mapping the development of facial expression recognition. *Developmental Science*, 18(6), 926-939.
- Roesch, E. B., Tamarit, L., Reveret, L., Grandjean, D., Sander, D., & Scherer, K. R. (2011). FACSGen: A tool to synthesize emotional facial expressions through systematic manipulation of facial action units. *Journal of Nonverbal Behavior*, 35(1), 1-16.
- Roh, J. H., Qiu, A., Seo, S. W., Soon, H. W., Kim, J. H., Kim, G. H., Kim, M.-J., Lee, J.-M., & Na, D. L. (2011). Volume reduction in subcortical regions according to severity of Alzheimer's disease. *Journal of Neurology*, 258(6), 1013-1020.
- Rolls, E. T. (1984). Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Human Neurobiology*, 3(4), 209-222.
- Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (2004). Basic objects in natural categories. In D. Balota & E. Marsh (Eds.). *Cognitive psychology: Key readings* (pp. 448-471). New York: Psychology Press.

- Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2003). Infant visual recognition memory: Independent contributions of speed and attention. *Developmental Psychology*, 39(3), 563-571.
- Rosenblum, L. D. (2008). Speech perception as a multimodal phenomenon. *Current Directions in Psychological Science*, 17(6), 405-409.
- Rossini, P. M., Rossi, S., Babiloni, C., & Polich, J. (2007). Clinical neurophysiology of aging brain: From normal aging to neurodegeneration. *Progress in Neurobiology*, 83(6), 375-400.
- Rossion, B. (2008). Constraining the cortical face network by neuroimaging studies of acquired prosopagnosia. *NeuroImage*, 40(2), 423-426.
- Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, 21(2), 139-253.
- Rossion, B. (2014). Understanding face perception by means of prosopagnosia and neuroimaging. *Frontiers in Bioscience (Elite Ed)*, 6, 258-307.
- Rossion, B. (2015). Face perception. In A. W. Toga (Ed.), *Brain mapping: An encyclopedic reference*, (Vol. 2, pp. 515-522). Cambridge, MA: Academic Press Elsevier.
- Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, 126(Pt 11), 2381-2395.
- Rossion, B., Dricot, L., Goebel, R., & Busigny, T. (2011). Holistic face categorization in higher order visual areas of the normal and prosopagnosic brain: Toward a non-hierarchical view of face perception. *Frontiers in Human Neuroscience*, 4, 1-30.
- Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, 1(1), 63-75.
- Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: A large-scale factorial fMRI face localizer analysis. *Brain and Cognition*, 79(2), 138-157.
- Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews*, 32(4), 863-881.
- Sachs, G., Steger-Wuchse, D., Kryspin-Exner, I., Gur, R. C., & Katschnig, H. (2004). Facial recognition deficits and cognition in schizophrenia. *Schizophrenia Research*, 68(1), 27-35.

- Sadeh, B., Podlipsky, I., Zhdanov, A., & Yovel, G. (2010). Event-related potential and functional MRI measures of face-selectivity are highly correlated: A simultaneous ERP-fMRI investigation. *Human Brain Mapping, 31*(10), 1490-1501.
- Sai, F. (2005). The role of the mother's voice in developing mother's face preference: Evidence for intermodal perception at birth. *Infant and Child Development, 14*(1), 29-50.
- Salloum, J. B., Ramchandani, V. A., Bodurka, J., Rawlings, R., Momenan, R., George, D., & Hommer, D. W. (2007). Blunted rostral anterior cingulate response during a simplified decoding task of negative emotional facial expressions in alcoholic patients. *Alcoholism: Clinical and Experimental Research, 31*(9), 1490-1504.
- Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science, 13*(4), 140-144.
- Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132*(4), 566-594.
- Sato, W., Kochiyama, T., Uono, S., & Yoshikawa, S. (2010). Amygdala integrates emotional expression and gaze direction in response to dynamic facial expressions. *NeuroImage, 50*(4), 1658-1665.
- Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Brain Research Cognitive Brain Research, 20*(1), 81-91.
- Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain, 130*(9), 2452-2461.
- Schachner, A., & Hannon, E. E. (2011). Infant-directed speech drives social preferences in 5-month-old infants. *Developmental Psychology, 47*(1), 19-25.
- Schaefer, K. L., Baumann, J., Rich, B. A., Luckenbaugh, D. A., & Zarate, C. A., Jr. (2010). Perception of facial emotion in adults with bipolar or unipolar depression and controls. *Journal of Psychiatric Research, 44*(16), 1229-1235.
- Scherer, K. (2005). "What are emotions? and how can they be measured?" *Social Science Information, 44*(4), 695-729.
- Scherer, K. R. (2000). Psychological models of emotion. In J. C. Borod (Ed.), *The neuropsychology of emotion*. New York: Oxford University Press.
- Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., & Rossion, B. (2006). Impaired face discrimination in acquired prosopagnosia is associated with abnormal

- response to individual faces in the right middle fusiform gyrus. *Cerebral Cortex*, 16(4), 574-586.
- Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research*, 194(3), 465-475.
- Schumann, C. M., Hamstra, J., Goodlin-Jones, B. L., Lotspeich, L. J., Kwon, H., Buonocore, M. H., Lammers, C. R., Reiss, A. L., & Amaral, D. G. (2004). The amygdala is enlarged in children but not adolescents with autism; the hippocampus is enlarged at all ages. *Journal of Neuroscience*, 24(28), 6392-6401.
- Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, 13(5), 402-409.
- Schyns, P. G., Petro, L. S., & Smith, M. L. (2009). Transmission of facial expressions of emotion co-evolved with their efficient decoding in the brain: Behavioral and brain evidence. *PLoS One*, 4(5), 1-16.
- Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science*, 16(4), 197-201.
- Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, 75(2), 221-242.
- Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain*, 115(1), 15-36.
- Sergent, J., & Signoret, J. L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex*, 2(5), 375-388.
- Sergent, J., & Villemure, J. G. (1989). Prosopagnosia in a right hemispherectomized patient. *Brain*, 112(Pt 4), 975-995.
- Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., Clasen, L., Evans, A., & Rapoport, J. L. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *Journal of Neuroscience*, 28(14), 3586-3594.
- Shuttleworth, E. C., Jr., Syring, V., & Allen, N. (1982). Further observations on the nature of prosopagnosia. *Brain and Cognition*, 1(3), 307-322.
- Simonsohn, U. (2017). Two-lines: The first valid test of u-shaped relationships with quadratic regressions.

- Simpson, E. A., Varga, K., Frick, J. E., & Frigaszy, D. (2000). Infants experience perceptual narrowing for nonprimate faces. *Infancy*, 16(3), 318-328.
- Skerry, A. E., & Spelke, E. S. (2014). Preverbal infants identify emotional reactions that are incongruent with goal outcomes. *Cognition*, 130(2), 204-216.
- Slater, A., Bremner, G., Johnson, S. P., Sherwood, P., Hayes, R., & Brown, E. (2000). Newborn infants' preference for attractive faces: The role of internal and external facial features. *Infancy*, 1(2), 265-274.
- Slater, A., Von der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S., & Samuels, C. (1998). Newborn infants prefer attractive faces. *Infant Behavior and Development*, 21(2), 345-354.
- Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, 16(3), 184-189.
- Solomon, R. C. (2008). The philosophy of emotions. In M. Lewis, J. M. Haviland-Jones & L. Feldman Barrett (Eds.), *Handbook of emotions* (pp. 3-16). New York: The Guilford Press.
- Somerville, L. H., Fani, N., & McClure-Tone, E. B. (2011). Behavioral and neural representation of emotional facial expressions across the lifespan. *Developmental Neuropsychology*, 36(4), 408-428.
- Sorger, B., Goebel, R., Schiltz, C., & Rossion, B. (2007). Understanding the functional neuroanatomy of acquired prosopagnosia. *NeuroImage*, 35(2), 836-852.
- Steeves, J. K., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2006). The fusiform face area is not sufficient for face recognition: Evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia*, 44(4), 594-609.
- Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: Typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology*, 56(2), 249-261.
- Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences, USA*, 105(1), 394-398.
- Sullivan, S., & Ruffman, T. (2004a). Emotion recognition deficits in the elderly. *International Journal of Neuroscience*, 114(3), 403-432.
- Sullivan, S., & Ruffman, T. (2004b). Social understanding: How does it fare with advancing years? *British Journal of Psychology*, 95(1), 1-18.

- Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 62(1), P53-P60.
- Surguladze, S., Brammer, M. J., Keedwell, P., Giampietro, V., Young, A. W., Travis, M. J., Williams, S. C., & Phillips, M. L. (2005). A differential pattern of neural response toward sad versus happy facial expressions in major depressive disorder. *Biological Psychiatry*, 57(3), 201-209.
- Surguladze, S. A., Brammer, M. J., Young, A. W., Andrew, C., Travis, M. J., Williams, S. C., & Phillips, M. L. (2003). A preferential increase in the extrastriate response to signals of danger. *NeuroImage*, 19(4), 1317-1328.
- Suzuki, A., & Akiyama, H. (2013). Cognitive aging explains age-related differences in face-based recognition of basic emotions except for anger and disgust. *Aging, Neuropsychology, and Cognition*, 20(3), 253-270.
- Suzuki, A., Hoshino, T., Shigemasu, K., & Kawamura, M. (2006). Disgust-specific impairment of facial expression recognition in Parkinson's disease. *Brain*, 129(3), 707-717.
- Sze, J. A., Goodkind, M. S., Gyurak, A., & Levenson, R. W. (2012). Aging and emotion recognition: not just a losing matter. *Psychology and Aging*, 27(4), 940.
- Talairach, P., & Tournoux, J. (1988). A stereotactic coplanar atlas of the human brain. Stuttgart: Thieme.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, 46(2), 225-245.
- Tanaka, J. W., & Simonyi, D. (2016). The “parts and wholes” of face recognition: A review of the literature. *The Quarterly Journal of Experimental Psychology*, 69(10), 1876-1889.
- Tardif, C., Lainé, F., Rodriguez, M., & Gepner, B. (2007). Slowing down presentation of facial movements and vocal sounds enhances facial expression recognition and induces facial-vocal imitation in children with autism. *Journal of Autism and Developmental Disorders*, 37(8), 1469-1484.
- Tessitore, A., Hariri, A. R., Fera, F., Smith, W. G., Das, S., Weinberger, D. R., & Mattay, V. S. (2005). Functional changes in the activity of brain regions underlying emotion processing in the elderly. *Psychiatry Research: Neuroimaging*, 139(1), 9-18.
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7(1), 53-71.

- Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science*, 10(5), 547-558.
- Tomlinson, E. K., Jones, C. A., Johnston, R. A., Meaden, A., & Wink, B. (2006). Facial emotion recognition from moving and static point-light images in schizophrenia. *Schizophrenia Research*, 85(1-3), 96-105.
- Tonnesen, J., Sorensen, A. T., Deisseroth, K., Lundberg, C., & Kokaia, M. (2009). Optogenetic control of epileptiform activity. *Proceedings of the National Academy of Sciences, USA*, 106(29), 12162-12167.
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, 168(3), 242-249.
- Trainor, L. J. (2005). Are there critical periods for musical development? *Developmental Psychobiology*, 46(3), 262-278.
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, 11(3), 188-195.
- Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, 20(3), 383-396.
- Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review*, 9(2), 335-340.
- Tranel, D., Damasio, A. R., & Damasio, H. (1988). Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. *Neurology*, 38(5), 690-696.
- Trautmann, S. A., Domínguez-Borràs, J., Escera, C., Herrmann, M., & Fehr, T. (2013). The perception of dynamic and static facial expressions of happiness and disgust investigated by ERPs and fMRI constrained source analysis. *PLoS One*, 8(6), 1-18.
- Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion: dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research*, 1284, 100-115.
- Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, 100(1), 73-99.
- Trehub, S. E., & Nakata, T. (2001). Emotion and music in infancy. *Musicae Scientiae*, 5, 37-61.

-
- Trehub, S. E., Plantinga, J., & Brcic, J. (2009). Infants detect cross-modal cues to identity in speech and singing. *Annals of the New York Academy of Sciences*, 1169(1), 508-511.
- Trojano, L., Conson, M., Salzano, S., Manzo, V., & Grossi, D. (2009). Unilateral left prosopometamorphopsia: A neuropsychological case study. *Neuropsychologia*, 47(3), 942-948.
- Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. (2003). Faces and objects in macaque cerebral cortex. *Nature Neuroscience*, 6(9), 989-995.
- Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, 311(5761), 670-674.
- Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008a). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences, USA*, 105(49), 19514-19519.
- Tsao, D. Y., Schweers, N., Moeller, S., & Freiwald, W. A. (2008b). Patches of face-selective cortex in the macaque frontal lobe. *Nature Neuroscience*, 11(8), 877-879.
- Turano, M. T., Lao, J., Richoz, A. R., De Lissa, P., Desgoscio, S. B. A., Viggiano, M. P., & Caldara, R. (2017). Fear boosts the early neural coding of faces. *Social, Cognitive, and Affective Neuroscience*, 1-13.
- Turati, C. (2004). Why faces are not special to newborns: An alternative account of the face preference. *Current Directions in Psychological Science*, 13(1), 5-8.
- Turati, C., Simion, F., Milani, I., & Umiltà, C. (2002). Newborns' preference for faces: What is crucial? *Developmental Psychology*, 38(6), 875-881.
- Uono, S., Sato, W., & Toichi, M. (2010). Brief report: Representational momentum for dynamic facial expressions in pervasive developmental disorder. *Journal of Autism and Developmental Disorders*, 40(3), 371-377.
- Valenza, E., Simion, F., Cassia, V. M., & Umiltà, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4), 892-903.
- Van Borsel, J., Janssens, J., & De Bodt, M. (2009). Breathiness as a feminine voice characteristic: A perceptual approach. *Journal of Voice*, 23(3), 291-294.
- Van den Hurk, J., Van Baelen, M., & Op de Beeck, H. P. (2017). Development of visual category selectivity in ventral visual cortex does not require visual experience. *Proceedings of the National Academy of Sciences, USA*, 114(22), E4501-E4510.

- Van Duuren, M., Kendell-Scott, L., & Stark, N. (2003). Early aesthetic choices: Infant preference for attractive premature infant faces. *International Journal of Behavioral Development*, 27, 212-219.
- Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin*, 122(3), 231-249.
- Vermeire, B. A., Hamilton, C. R., & Erdmann, A. L. (1998). Right-hemispheric superiority in split-brain monkeys for learning and remembering facial discriminations. *Behavioral Neuroscience*, 112(5), 1048-1061.
- 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.
- Vizioli, L., Foreman, K., Rousselet, G. A., & Caldara, R. (2010a). Inverting faces elicits sensitivity to race on the N170 component: A cross-cultural study. *Journal of Vision*, 10(1), 1-23.
- Vizioli, L., Rousselet, G. A., & Caldara, R. (2010b). Neural repetition suppression to identity is abolished by other-race faces. *Proceedings of the National Academy of Sciences, USA*, 107(46), 20081-20086.
- Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia*, 45(1), 174-194.
- Walker-Andrews, A. S. (1986). Intermodal perception of expressive behaviors: Relation of eye and voice? *Developmental Psychology*, 22(3), 373-377.
- Walker-Andrews, A. S., Bahrick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infants' bimodal perception of gender. *Ecological Psychology*, 3(2), 55-75.
- Wallraven, C., Breidt, M., Cunningham, D. W., & Bülthoff, H. H. (2008). Evaluating the perceptual realism of animated facial expressions. *ACM Transactions on Applied Perception (TAP)*, 4(4), 4.
- Walton, G. E., & Bower, T. (1993). Newborns form "prototypes" in less than 1 minute. *Psychological Science*, 4(3), 203-205.
- Wang, G., Tanaka, K., & Tanifuji, M. (1996). Optical imaging of functional organization in the monkey inferotemporal cortex. *Science*, 272(5268), 1665-1668.

- Wang, G., Tanifuji, M., & Tanaka, K. (1998). Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. *Neuroscience Research*, 32(1), 33-46.
- Warrington, E. K. (1984). *Recognition Memory Test*. Windsor, Berkshire: NFER.
- Wasserman, E. A., Lazareva, O., Gibson, B., Gosselin, F., Schnys, P., & Biederman, I. (2004). Geons and Bubbles: Object recognition by pigeons. *Journal of Vision*, 4(8), 353.
- Watson, J. B. (1930). *Behavior*. Chicago: Chicago University Press.
- Watson, R., Latinus, M., Charest, I., Crabbe, F., & Belin, P. (2014). People-selectivity, audiovisual integration and heteromodality in the superior temporal sulcus. *Cortex*, 50, 125-136.
- Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, 78(1), 105-119.
- Weiner, B., & Graham, S. (1984). An attributional approach to emotional development. In C. E. Izard, J. Kagan & R. B. Zajonc (Eds.), *Emotions, cognition, and behavior* (pp. 167-191). New York: Cambridge University Press.
- Weiner, K. S., & Grill-Spector, K. (2012). The improbable simplicity of the fusiform face area. *Trends in Cognitive Sciences*, 16(5), 251-254.
- Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, 103(1), 147-162.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49-63.
- Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology*, 46(3), 233-251.
- West, J. T., Horning, S. M., Klebe, K. J., Foster, S. M., Cornwell, R. E., Perrett, D., Burt, D. M., & Davis, H. P. (2012). Age effects on emotion recognition in facial displays: From 20 to 89 years of age. *Experimental Aging Research*, 38(2), 146-168.
- West, M. J. (1993). Regionally specific loss of neurons in the aging human hippocampus. *Neurobiology of Aging*, 14(4), 287-293.

- 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.
- Widen, S. C., & Russell, J. A. (2015). Do dynamic facial expressions convey emotions to children better than do static ones? *Journal of Cognition and Development*, 16(5), 802-811.
- Wiese, H., Schweinberger, S. R., & Hansen, K. (2008). The age of the beholder: ERP evidence of an own-age bias in face memory. *Neuropsychologia*, 46(12), 2973-2985.
- Wigan, A. L. (1844). *The duality of the mind: Proved by the structure, functions, and diseases of the brain, and by the phenomena of mental derangement, and shown to be essential to moral responsibility*. London: Longman, Brown, Green, and Longmans.
- Wilcock, R. A., Bull, R., & Vrij, A. (2007). Are old witnesses always poorer witnesses? Identification accuracy, context reinstatement, own-age bias. *Psychology, Crime & Law*, 13(3), 305-316.
- Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, 42(3), 671-684.
- Williams, L. M., Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., & Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology*, 31(3), 257-277.
- Winston, J., O'doherty, J., & Dolan, R. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *NeuroImage*, 20(1), 84-97.
- Winston, J. S., Henson, R. N., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of Neurophysiology*, 92(3), 1830-1839.
- Wong, B., Cronin-Golomb, A., & Neargarder, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology*, 19(6), 739-749.
- Wood, S.N. (2006). *Generalized additive models: An introduction with R*. Boca Raton: CRC Press.
- Wright, D. B., & Stroud, J. N. (2002). Age differences in lineup identification accuracy: People are better with their own age. *Law and Human Behavior*, 26(6), 641-654.
- Xiao, N. G., Perrotta, S., Quinn, P. C., Wang, Z., Sun, Y.-H. P., & Lee, K. (2014). On the facilitative effects of face motion on face recognition and its development. *Frontiers in Psychology*, 5, 1-16.

- Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2015). Eye tracking reveals a crucial role for facial motion in recognition of faces by infants. *Developmental Psychology*, 51(6), 744-757.
- Xu, X., & Biederman, I. (2010). Loci of the release from fMRI adaptation for changes in facial expression, identity, and viewpoint. *Journal of Vision*, 10(14), 36-36.
- Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *Journal of Neuroscience*, 29(26), 8525-8529.
- Yang, H., Susilo, T., & Duchaine, B. (2015). The anterior temporal face area contains invariant representations of identity that can persist despite the loss of right FFA and OFA. *Cerebral Cortex*, 26(3), 1096-1107.
- Yang, T. T., Menon, V., Eliez, S., Blasey, C., White, C. D., Reid, A. J., Gotlib, I. H., & Reiss, A. L. (2002). Amygdalar activation associated with positive and negative facial expressions. *Neuroreport*, 13(14), 1737-1741.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81(1), 141-145.
- Yoshikawa, S., & Sato, W. (2006). Enhanced perceptual, emotional, and motor processing in response to dynamic facial expressions of emotion. *Japanese Psychological Research*, 48(3), 213-222.
- Young, A., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, 16, 747-759.
- Young, A. W., McWeeny, K. H., Hay, D. C., & Ellis, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. *Psychological Research*, 48(2), 63-68.
- Young, A. W., Newcombe, F., de Haan, E. H., Small, M., & Hay, D. C. (1993). Face perception after brain injury. Selective impairments affecting identity and expression. *Brain*, 116(Pt 4), 941-959.
- Younger, B. A., & Fearing, D. D. (1999). Parsing items into separate categories: Developmental change in infant categorization. *Child Development*, 70(2), 291-303.
- Yovel, G., & Freiwald, W. A. (2013). Face recognition systems in monkey and human: Are they the same thing? *F1000Prime Reports*, 5, 1-8.
- Yovel, G., & O'Toole, A. J. (2016). Recognizing people in motion. *Trends in Cognitive Sciences*, 20(5), 383-395.

- Yu, H., Garrod, O. G. B., & Schyns, P. G. (2012). Perception-driven facial expression synthesis. *Computer & Graphics*, 36(3), 152-162.
- Zangl, R., & Mills, D. L. (2007). Increased brain activity to infant-directed speech in 6-and 13-month-old infants. *Infancy*, 11(1), 31-62.
- Zelinski, E. M., & Kennison, R. F. (2007). Not your parents' test scores: Cohort reduces psychometric aging effects. *Psychology and Aging*, 22(3), 546-557.
- Zhao, K., Zhao, J., Zhang, M., Cui, Q., & Fu, X. (2017). Neural responses to rapid facial expressions of fear and surprise. *Frontiers in Psychology*, 8, 1-8.
- Zhao, M.-F., Zimmer, H. D., Shen, X., Chen, W., & Fu, X. (2016). Exploring the cognitive processes causing the age-related categorization deficit in the recognition of facial expressions.

Anne-Raphaëlle RICHOZ

Department of Psychology, University of Fribourg
Rue P.A. Faucigny 2, 1700 Fribourg, Switzerland
+ 41 (0)26 300 76 83, + 41 (0)79 281 11 73
anne-raphaelle.richoz@unifr.ch
Nationality: Swiss

ACADEMIC POSITION

- | | |
|-----------------|---|
| 2017–now | University of Fribourg, Switzerland
Lecturer. Department of Psychology. Unit for social and visual neuroscience |
| 2015–now | University of Fribourg, Switzerland
Advisor of studies for the MSc students in Psychology – Cognitive Neuroscience. Department of Psychology. |
| 2015–now | University of Fribourg, Switzerland
Advisor of studies for the MSc students in Psychology – Development. Department of Psychology. |
| 2013–now | University of Fribourg, Switzerland
Teaching assistant for the following classes: Cognitive Psychology, Cognitive Neuroscience, Practical experiments in psychology. Supervision of seminar dissertations, Supervision of Bachelor and Master theses. Department of Psychology. |

EDUCATION

- | | |
|------------------|--|
| 2015 –now | PhD Student at the “Collège Doctoral” Université Grenoble-Alpes, under the supervision of Prof. Olivier Pascalis |
| 2014 –now | PhD Student in the “Programme Doctoral Romand en Psychologie” |
| 2013 –now | PhD Student in social and visual neuroscience, Department of Psychology, University of Fribourg under the supervision of Prof. Roberto Caldara |
| 2012–13 | English writing and communication courses, Northeastern University, Boston |
| 2010–12 | Master in Clinical Psychology with the distinction <i>insigni cum laude</i> , University of Fribourg |
| 2011–12 | Teaching tutor for the class E-Prime, University of Fribourg |
| 2007–10 | Bachelor in Psychology with the distinction <i>insigni cum laude</i> , University of Fribourg |
| 2009 | Superior Certificate in Piano at the Lausanne Conservatory of Music |
| 2006–07 | Music studies and German classes, Berlin |
| 2006 | High School Diploma (Maturité fédérale), Gymnasium August Piccard, special section for musicians and elite athletes |
| 2005 | Certificate of musical theory, Conservatory of Lausanne, Switzerland |
| 1994–2002 | Ecole Rudolf Steiner, Lausanne |

PUBLICATIONS

- Turano, M. T., Lao, J., Richoz, A. R., De Lissa, P., Desgosciu, S. B. A., Viggiano, M. P., & Caldara, R. (2017). Fear boosts the early neural coding of faces. *Social, Cognitive, and Affective Neuroscience*, 1-13.
- Fiset, D., Blais, C., Royer, J., Richoz, A.-R., Dugas, G., & Caldara, R. (2017). Mapping the impairment in decoding static facial expressions of emotion in prosopagnosia. *Social, Cognitive, and Affective Neuroscience*, 12(8), 1334–1341.
- Richoz, A.-R., Quinn, P. C., Hillairet de Boisferon, A., Berger, B., Loevenbruck, H., Lewkowicz, D. J., Lee, K., Dole, M., & Caldara, R., & Pascalis, O. (2017). Audio-Visual Perception of Gender by Infants Emerges Earlier for Adult-Directed Speech. *PLoS One*, 12(1), 1–15.
- Richoz, A.-R., Jack, R. E., Garrod, O. G. B., Schyns, P. G. & Caldara, R. (2015). Reconstructing Dynamic Mental Models of Facial Expressions in Prosopagnosia Reveals Distinct Representations for Identity and Expression. *Cortex*, 65, 50–64.

In preparation

- Richoz, A.-R., Lao, J., Pascalis, O., Caldara, R. (submitted). Tracking the recognition of static and dynamic facial expressions of emotion across life span.
- Richoz, A.-R., Elphick, C., & Caldara, R. (in preparation). The right occipital face area is not mandatory for facial symmetry perception.
- Luisier, A.-C., Petitpierre, G., Richoz, A.-R., Lao, J., Caldara, R., Clerc Béro, A., Bensafi, M. (under review). Visual exploration and hedonic processing in children with autism spectrum disorders and their relationship to food neophobia.

CONFERENCE PRESENTATIONS

Talks

- Richoz, A.-R., Lao, J., Pascalis, O., Caldara, R. (2017). Tracking the recognition of static and dynamic facial expressions of emotion across life span. Swiss Psychological Society (SSP), Lausanne, Switzerland.
- Richoz, A.-R., Lao, J., Pascalis, O., Caldara, R. (2017). Tracking the recognition of static and dynamic facial expressions of emotion across life span. Vision Sciences Society, St. Pete Beach, USA.
- Richoz, A.-R., Jack, R.E., Garrod, O.G.B., Schyns, P.G., & Caldara, R. (2014). Reconstructing emotions in motion in prosopagnosia supports separate representations for identity and emotion. European Conference for Visual Perception (ECPV), Belgrade, Serbia.
- Richoz, A.-R., Jack, R.E., Garrod, O.G.B., Schyns, P.G., & Caldara, R. (2013). Reconstructing emotions in motion abolishes facial expression categorization impairment in prosopagnosia. Congress of the Swiss Psychological Society (SSP), Basel, Switzerland.

Posters

- Richoz, A.-R., Quinn, P.C., Hillairet de Boisferon, A., Berger, B., Loevenbruck, H., Lewkowicz, D. J., Lee, K., Dole, M., & Caldara, R., & Pascalis, O. (2016). Audio-visual perception of gender by infants emerges earlier for adult-directed speech. Poster presented at the Psychonomic Society's 57th Annual Meeting, Boston, USA.
- Richoz, A.R., Elphick, C., & Caldara, R. (2015). The right occipital face area is not necessary for facial symmetry perception. Poster presented at the Swiss Society of Psychology (SSP), Geneva, Switzerland.
- Richoz, A.R., Jack, R.E., Garrod, O.G.B., Schyns, P.G., & Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. Poster presented at the 20th Alpine Brain Imaging Meeting (ABIM), Champéry, Switzerland.
- Richoz, A.R., Jack, R.E., Garrod, O.G.B., Schyns, P.G., & Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. Poster presented at the SSN Annual Meeting (Swiss Society of Neuroscience), Fribourg, Switzerland.

GRANT

2014: University of Fribourg, Switzerland. CHF 10'000. Scholarship from the CRUS (Swiss Universities) for a joint PhD supervision with the University of Grenoble-Alpes, France.

SUPERVISION

Bachelor theses

Flavie, Déglise; Marcel Zwissig; Pauline Rotzetter; Vanessa Ferrari; Qendresa Shkodra; Pauline Schaller; Linda Pigozzo; Martina Studer

Master theses (co-supervision with Prof. Caldara)

Leonora Alimetaj; Mélanie Fernandez; Martino Ceroni; Benedetta Leidi-Maimone; Pauline Schaller ; Anna Speckert

Internships

Rémy Lachelin, Yann Girardin, Mattia Moggio, Lauriane Beffa, Garance Lehmann, Martina Studer, Christel Aichele, Adrian Kolly, Serena Malagnino, Sylvie Huguenot, Delphine Waeber, Eva Huguenin-Elie

TEACHING

Cognitive Psychology class (2013–now)
Cognitive Neuroscience class (2014–2015)
Research Colloquium for Bachelor students (2016–now)
Practical Experience in Psychology class, 2nd year BA students (2017–2018)

SKILLS

Informatics

Software: Word, Power point, Excel, Endnote, SPSS

Matlab (user), R (user)

Languages

French, native speaker

Swiss-German, native speaker

German, advanced, spoken and written

English, advanced, spoken and written

Italian, advanced, spoken and written

HOBBIES AND INTERESTS

Music studies: 15 years of piano studies at the Conservatory of Music, Lausanne.

Sports: Running, walking, hiking.

I am fascinated by artists. This fascination includes different works of art, particularly from Italian and French cultures, such as music (from the classical or even baroque era to recent music), opera, poetry, paintings, cinema, and modern literature. I am also interested in traveling and discovering different world cultures, particularly Asian ones.

I also enjoy spending time in nature, hiking in the Swiss mountains, eating good and healthy food, playing piano, or watching interesting and outstanding (French) movies.

Fribourg, January 2018

Publication List

Experimental research papers presented in this thesis

Richoz, A.-R., Lao, J., Pascalis, O., Caldara, R. (submitted). Tracking the recognition of static and dynamic facial expressions of emotion across life span.

Richoz, A.-R., Quinn, P. C., Hillairet de Boisferon, A., Berger, B., Loevenbruck, H., Lewkowicz, D. J., Lee, K., Dole, M., & Caldara, R., & Pascalis, O. (2017). Audio-visual perception of gender by infants emerges earlier for adult-directed speech. *PLoS One*, 12(1), 1–15.

Richoz, A.-R., Jack, R. E., Garrod, O. G. B., Schyns, P. G., & Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. *Cortex*, 65, 50–64.

Other publications

Richoz, A.-R., Elphick, C., & Caldara, R. (in preparation). The right occipital face area is not mandatory for facial symmetry perception.

Luisier, A.-C., Petitpierre, G., Richoz, A.-R., Lao, J., Caldara, R., Clerc Béro, A., & Bensafi, M. (under review). Visual exploration and hedonic processing in children with autism spectrum disorders and their relationship to food neophobia.

Turano, M. T., Lao, J., Richoz, A. R., De Lissa, P., Desgosciu, S. B. A., Viggiano, M. P., & Caldara, R. (2017). Fear boosts the early neural coding of faces. *Social, Cognitive, and Affective Neuroscience*, 1-13.

Fiset, D., Blais, C., Royer, J., Richoz, A.-R., Dugas, G., & Caldara, R. (2017). Mapping the impairment in decoding static facial expressions of emotion in prosopagnosia. *Social, Cognitive, and Affective Neuroscience*, 12(8), 1334–1341.

Je déclare sur mon honneur que ma thèse est une œuvre personnelle, composée sans concours extérieur non autorisé, et qu'elle n'a pas été présentée devant une autre Faculté.

