

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

L'AUTOMATICITÉ DE L'EFFET CONTEXTUEL DANS LA PERCEPTION DES
EXPRESSIONS ÉMOTIONNELLES FACIALES ET CORPORELLES : APPROCHE
NEUROSCIENTIFIQUE

THÈSE PRÉSENTÉE
COMME EXIGENCE PARTIELLE DU
DOCTORAT CONTINUUM D'ÉTUDES EN PSYCHOLOGIE
(PROFIL RECHERCHE)

PAR
ANNE-SOPHIE PUFFET

SEPTEMBRE 2025

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire, de cette thèse ou de cet essai a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire, de sa thèse ou de son essai.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire, cette thèse ou cet essai. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire, de cette thèse et de son essai requiert son autorisation.

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

DOCTORAT CONTINUUM D'ÉTUDES EN PSYCHOLOGIE
(PROFIL RECHERCHE) (Ph. D.)

Direction de recherche :

Simon Rigoulot, Ph. D. directeur de recherche
Université du Québec à Trois-Rivières

Jury d'évaluation :

Simon Rigoulot, Ph. D. directeur de recherche
Université du Québec à Trois-Rivières

Benjamin Boller, Ph. D. président du jury
Université du Québec à Trois-Rivières

Isabelle Blanchette, Ph. D. évaluatrice interne
Université Laval

Jorge Armony, Ph. D. évaluateur externe
Université McGill

Thèse soutenue le [jj/mm/année]

Ce document est rédigé sous la forme d'article(s) scientifique(s), tel qu'il est stipulé dans les règlements des études de cycles supérieurs (Article 360) de l'Université du Québec à Trois-Rivières. Les articles ont été rédigés selon les normes de publication de revues reconnues et approuvées par le Comité de programmes de cycles supérieurs du département de psychologie. Le nom du directeur de recherche pourrait donc apparaître comme co-auteur de l'article soumis pour publication.

Sommaire

À travers nos interactions sociales quotidiennes, nous percevons continuellement les émotions des personnes qui nous entourent. Ces expressions émotionnelles jouent un rôle central en orientant nos comportements et en nous informant sur les ressentis émotionnels d'autrui. Traditionnellement, la recherche s'est longtemps focalisée sur le traitement des expressions faciales isolées en négligeant d'autres indices tels que la posture corporelle, la scène environnante, la voix ou encore les odeurs. Or, c'est la prise en compte simultanée de l'ensemble de ces signaux qui permet de saisir toute la richesse de la perception émotionnelle. Dans ce cadre, un champ de recherche a exploré l'influence du contexte (p. ex., la posture corporelle, la scène environnante, les sons, les odeurs, etc.) sur le traitement des expressions faciales émotionnelles, montrant une interaction entre ces différents canaux. L'interaction visage-corps apparaît particulièrement pertinente, car, dans la vie quotidienne, les expressions faciales sont rarement perçues isolément des expressions corporelles. Elle se manifeste par une meilleure reconnaissance émotionnelle lorsque les expressions faciales et corporelles sont congruentes (exprimant la même émotion), et à l'inverse, par une reconnaissance entravée lorsqu'elles sont incongruentes (exprimant des émotions différentes). Cependant, la recherche sur la perception simultanée d'indices faciaux et corporels présente plusieurs lacunes : Premièrement, (1) la majorité des études utilise des images distinctes pour les expressions faciales et corporelles, produisant des stimuli de faible qualité (en noir et blanc, mauvaises proportions, etc.) lorsqu'ils sont combinés, limitant ainsi la compréhension de la perception réelle des émotions. Deuxièmement, (2) l'intégration entre les expressions

faciales et corporelles reste mal comprise, notamment quant à son caractère automatique. Alors que certaines études suggèrent un traitement émotionnel efficace, rapide et inconscient, d'autres montrent que l'attention volontaire portée au visage ou au corps influence cet effet, remettant en cause l'idée d'une automaticité complète. Enfin, le décours temporel de l'activité cérébrale a été peu étudié, réduisant nos connaissances sur les mécanismes sous-jacents. Ces lacunes freinent notre compréhension de la perception réaliste des expressions émotionnelles et des processus qui la sous-tendent. L'objectif de cette thèse est d'examiner, sur les plans comportemental et neural, l'automaticité de l'intégration des expressions faciales et corporelles, dans des conditions plus proches de la perception réelle. Trois études ont été menées : la création et la validation d'une banque de stimuli congruents et incongruents (ECIFBSS; article 1), puis l'exploration de l'automaticité à travers l'efficacité (étude 2) et la rapidité (étude 3) du traitement émotionnel. Les résultats soutiennent l'hypothèse d'une automaticité de l'intégration, indépendamment de la rapidité du traitement émotionnel, mais montrent que, aux stades cérébraux précoces, celle-ci reste partielle, car modulée par la charge cognitive. Cette thèse apporte ainsi un double apport : une ressource unique de stimuli validés (ECIFBSS) et de nouvelles connaissances fondamentales sur les mécanismes d'intégration des signaux émotionnels, ouvrant la voie à des recherches futures, notamment dans d'autres modalités sensorielles ou auprès de populations cliniques.

Table des matières

Sommaire	iv
Liste des tableaux	xiv
Liste des figures	xvi
Remerciements	xx
Introduction générale	1
Expressions faciales émotionnelles isolées.....	4
Fondements théoriques majeurs des expressions faciales émotionnelles	5
Théorie des émotions discrètes	6
Théories dimensionnelles	7
Expressions faciales émotionnelles comme préparation à l'action	11
Expressions faciales émotionnelles comme une évaluation cognitive	12
Corrélat neuronal de la perception des expressions faciales isolées	15
Les trois stades du traitement des expressions faciales émotionnelles	16
Expressions corporelles émotionnelles isolées	26
Fondements théoriques des postures corporelles comme vecteur émotionnel ..	27
Corrélat neuronal de la perception des expressions corporelles isolées.....	32
Influence bidirectionnelle entre les expressions faciales et corporelles	37
Effet contextuel entre les expressions faciales et corporelles.....	38
Décours temporel de l'intégration des expressions faciales et corporelles.....	46
Automaticité dans le traitement émotionnel	51
Automaticité d'un processus cognitif	52

Automaticité d'un processus émotionnel : les expressions émotionnelles isolées.....	56
Recherches en faveur d'une automaticité du traitement émotionnel.....	56
Nuances sur l'automaticité du traitement émotionnel	59
Automaticité de l'effet contextuel	63
Objectif et hypothèses de recherche	65
Objectifs et hypothèses de la première étude.....	67
Objectifs et hypothèses de la seconde étude	69
Objectifs et hypothèses de la troisième étude	71
Chapitre 1. Création et validation d'une banque de stimuli d'expressions faciales et corporelles congruents et incongruents.....	73
Stimuli d'expressions faciales et corporelles émotionnelles.....	74
Méthodologie de création et de validation de stimuli émotionnels	75
Création de stimuli congruents et incongruents dans l'étude de l'effet contextuel ..	78
Chapitre 2. Article 1 – Validation of the Emotionally Congruent and Incongruent Face-Body Static Set (ECIFBSS)	81
Abstract	83
Introduction.....	84
Stimuli Material Creation	91
Method	91
Participants.....	91
Materials	92
Procedure	94
Validation of Stimuli: Part 1	95

Methods.....	95
Participants.....	95
Materials	96
Procedure	97
Measures of Interest.....	98
Statistical Analysis.....	99
Results.....	100
Descriptive Statistics.....	100
Face Stimuli	101
Reliability.....	101
“Raw Hit Rates” and “Unbiased Hit Rates”	102
Recognition of Facial Basic Emotions: Emotion and Gender Difference	105
Confusion Between Facial Basic Emotions.....	106
Body Stimuli	107
Reliability.....	107
“Raw Hit Rate” and “Unbiased Hit Rate”	108
Recognition of Basic Body Postures: Emotion and Gender Difference ...	111
Confusion Between Basic Body Postures.....	112
Discussion	113
Validation of Stimuli: Part 2	116
Methods.....	116
Participants.....	116

Materials	117
Procedure	117
Measures and Statistical Analyses	118
Results.....	119
Descriptive Statistics.....	119
Facial Stimuli	119
Reliability.....	119
“Raw Hit Rates” and “Unbiased Hit Rates”	120
Recognition of Facial Basic Emotions: Emotion and Gender Difference	123
Confusion Between Facial Basic Emotions.....	124
Valence, Arousal, Intensity, and Authenticity: Emotions and Gender Differences.....	125
Valence	127
Arousal.....	127
Intensity	127
Authenticity	128
Body Stimuli	128
Reliability.....	128
“Raw Hit Rates” and “Unbiased Hit Rates”	129
Recognition of Basic Body Postures: Emotion and Gender Difference...	132
Confusion Between Basic Body Postures.....	133
Intensity, Valence, Arousal, and Authenticity: Emotions and Gender Differences.....	135

Valence	137
Arousal.....	137
Intensity	137
Authenticity	137
Discussion	137
General Discussion	139
Advantages and Limits of the ECIFBSS	141
Conclusion	144
References.....	147
Transition vers le Chapitre 3.....	160
Chapitre 3. Article 2 – The Role of Cognitive Load in Automatic Integration of Emotional Information from Face and Body	161
Abstract	163
Introduction.....	164
Methods.....	171
Participants.....	171
Materials	171
Experimental Design.....	175
General Procedure.....	175
Statistics Analysis	179
Behavioral Analysis.....	179
ERP Recording, Preprocessing, and Analysis	181
Results.....	183

Memorization Task	183
Preliminary Analyses: Behavioral and EEG Data	183
Hypothesis 1: Contextual Effect	186
Hypothesis 2: Attentional Focus Effect	188
Hypothesis 3: Cognitive Load Effect.....	194
Discussion	195
The Contextual Effect Between Face and Bodily Cues.....	196
The Role of Attentional Focus on the Facial and Bodily Expressions Integration	197
The Role of Cognitive Load on the Facial and Bodily Expressions Integration	200
Limitations	201
Conclusion	202
References	204
Transition vers le Chapitre 4.....	214
Chapitre 4. Article 3 – How Speed of Processing Influences the Integration of Facial and Bodily Emotional Expressions: An Event-Related Potential Study	215
Abstract	217
Introduction.....	218
Methods.....	223
Participants.....	223
Materials	224
Experimental Design.....	227
General Procedure.....	228

Statistics Analysis	229
Behavioral Analysis	229
ERP Recording, Preprocessing, and Analysis	231
Results	233
Preliminary Analyses: Behavioral and EEG Data	233
Hypothesis 1: Contextual Effect	235
Hypothesis 2: Attentional Focus Effect	238
Hypothesis 3: Speed Effect	244
Discussion	246
The Contextual Effect	247
The Role of Attentional Focus on the Facial and Bodily Expressions Integration	250
The Role of Speed on the Facial and Bodily Expressions Integration	252
Limitation	253
Conclusion	254
References	255
Discussion générale	263
Création d'une banque de stimuli pour une approche plus écologiquement valide	266
Automaticité de l'intégration des expressions faciales et corporelles	269
Charge cognitive	270
Rapidité de traitement	271
Intégration des deux articles expérimentaux	273

Proposition d'un modèle intégratif	277
Limites et perspectives	285
Conclusion générale	293
Références générales	296
Appendice. Certificats d'éthique	315

Liste des tableaux

Liste des tableaux dans l'Article 1 :

Tableau

1	Summary Table of “Raw Hit Rate” and “Unbiased Hit Rates” for Facial Emotions	103
2	Confusion Matrix for the “Raw Hit Rate” of Facial Emotions.....	104
3	Summary Table of “Raw Hit Rate” and “Unbiased Hit Rates” for Body Postures	109
4	Confusion Matrix for the “Raw Hit Rate” of Body Postures	110
5	Summary Table of “Raw Hit Rate” and “Unbiased Hit Rates” for Facial Emotions	121
6	Confusion Matrix for the “Raw Hit Rate” of Facial Emotions.....	122
7	Means and SD for Arousal, Valence, Intensity, and Authenticity of Each Stimuli Emotions by Gender Actors	129
8	Summary Table of “Raw Hit Rate” and “Unbiased Hit Rates” for Body Posture	130
9	Confusion Matrix for the “Raw Hit Rate” of Body Postures	131
10	Means and SD for Arousal, Valence, Intensity, and Authenticity of Each Stimuli Emotions by Gender Actors	137

Liste des tableaux dans l'Article 2 :

Tableau

1	Mean and Standard Deviation of Valence, Arousal and Intensity for Facial and Bodily Emotions	173
---	--	-----

Liste des tableaux dans l'Article 3 :

Tableau

1	Mean and Standard Deviation of Valence, Arousal and Intensity for Facial and Bodily Emotions	226
---	--	-----

Liste des figures

Liste des figures dans la thèse :

Figure

1	Modèle dimensionnel de Russell (1980).....	8
2	Représentation de la théorie d'évaluation cognitive des émotions d'Arnold	13
3	Représentation du Component Process Model (CPM)	15
4	Étapes de l'acquisition et du pré-traitement du discours temporel de l'activité cérébrale par EEG	17
5	Schéma du modèle en trois phases du traitement des expressions faciales émotionnelles de Luo et al. (2010)	18
6	Graphique des PEs pour la joie, la neutralité et la peur sur l'électrode P8	22
7	Grand moyennage des PEs sur l'électrode FCz lorsque le focus attentionnel est dirigé vers l'expression faciale (gauche) et corporelle (droite).....	35
8	Exemple de stimuli congruent et incongruent utilisés dans la littérature	39
9	Image représentant des stimuli alignés (à gauche) et désalignés (à droite)	41
10	Schéma récapitulatif du traitement des expressions faciales ou corporelles	50
11	Stimuli Congruents et Incongruents Utilisés dans de Précédentes Etudes sur l'Effet Contextuel.....	80
12	Schéma représentant le continuum d'automaticité dans lequel s'inscrivent nos résultats comportementaux et électrophysiologiques.....	280
13	Représentation graphique du modèle proposé sur base du travail de thèse.....	284

Liste des figures dans l'Article 1 :

Figure

1	Examples of Congruent Stimuli.....	93
2	Examples of Incongruent Stimuli	93
3	Examples of Face Emotion Stimuli	97
4	Example of a Trial for the Face (left) and Body (right) Expressions	98
5	Boxplot of the Arcsine Transformation of the Hu Score by Facial Emotions	106
6	Cumulative of Accuracy (%) for Each Facial Emotion.....	107
7	Boxplot of the Arcsine Transformation of the Hu Score by Body Postures.....	112
8	Cumulative of Accuracy (%) for Each Body Postures	113
9	Boxplot of the Arcsine Transformation of the Hu Score by Facial Expression.....	124
10	Cumulative of Accuracy (%) for Each Facial Emotion.....	125
11	Scatterplot Representing the Mean Intensity as a Function of the Mean Arousal for Each Emotional Stimulus	126
12	Boxplot of the Arcsine Transformation of the Hu Score by Body Postures.....	133
13	Cumulative of Accuracy (%) for Each Body Postures	134
14	Scatterplot Representing the Mean Intensity as a Function of the Mean Arousal for Each Emotional Stimulus	135

Liste des figures dans l'Article 2 :

Figure

1	Schema of Block Structure	178
2	Plot ERP Waveforms of All Electrodes (A), Grand Averaged ERP for Poz and Oz Electrodes (B), and FMUT map of Focus and Cognitive Load Effects (C)	185
3	Histogram of Means of Arcsine Transformation (A) and Response Time (B) for the Congruence	187
4	Histogram of Means of Arcsine Transformation (A) and Response Time (B) for Interaction between Congruence and Focus Type	189
5	Scalp map, FMUT map and Grand Averaged ERP on the Fronto-Central, and Parieto-Occipital Sites.....	191
6	Scalp map, FMUT map and Grand Averaged ERP on the Fronto-Central, and Parieto-Occipital Sites.....	192
7	Scalp map, FMUT map and Grand Averaged ERP on Occipital Sites.....	193
8	Histogram of Congruence in Function of Focus Type for Each Cognitive Load Conditions for Means of Arcsine Transformation (A) and for Response Times (C), and Violin Plot of Accuracy Means in Congruence in Function of Cognitive Load and Focus for Means of Arcsine Transformation (B) and Response Times (D).....	195

Liste des figures dans l'Article 3 :

Figure

1	Shema of Experimental Procedure.....	229
2	Plot ERP Waveforms of All Electrodes (A), Scalp Mapping for Focus and Congruence Conditions (B), and FMUT map of Focus and Congruence Effects (C and D)	234
3	Histogram of Means of Arcsine Transformation (A) and Response Time (B) for the Congruence	236

Figure

4	FMUT map of Congruence Effect (A), Grand Averaged ERP on the Parietal Sites(B), Scalp map (C)	237
5	FMUT map of Congruence Effect (A), Grand Averaged ERP on the Centro-Parietal Sites(B), Scalp map (C).....	238
6	Histogram of Means of Arcsine Transformation (A) and Response Time (B) for the Congruence and Focus Interaction.....	239
7	FMUT map of Focus Effect (A), Grand Averaged ERP on the Fronto-Central and Parieto-Occipital sites(B), Scalp map (C)	241
8	FMUT map of Focus Effect (A), Grand Averaged ERP on the Fronto-Central and Occipital Sites(B), Scalp map (C)	243
9	Histogram of Means of Arcsine Transformation (A) and Response Time (B) for the Speed *Congruence*Focus Interaction.....	244
10	FMUT map of Congruence * Speed Effect (A), Grand Averaged ERP on the Frontal Sites(B), Scalp map (C).....	246

Remerciements

Je tiens à remercier particulièrement Simon Rigoulot, mon directeur de thèse, pour la confiance qu'il m'a accordée tout au long de ces années. Tes conseils avisés et ton expertise ont été essentiels dans mon développement en tant que chercheuse. Merci d'avoir soutenu mes idées, même lorsque celles-ci sortaient des sentiers battus. Ton encadrement bienveillant, ton écoute et ta disponibilité ont été un moteur précieux tout au long de mon parcours. Merci de m'avoir embarquée dans cette aventure, même si, paraît-il, je n'étais pas ton premier choix... Comme quoi, le meilleur n'est pas toujours celui qu'on attend 😊

Je souhaite également remercier Solène Kalénine et Yannick Wamain qui ont su éveiller en moi l'intérêt et la passion pour la recherche scientifique. Nos échanges, vos conseils et votre présence ont profondément marqué mon parcours. Réaliser mon mémoire de master à vos côtés fut une décision déterminante, tant pour les compétences que j'ai pu acquérir que pour la qualité de votre accompagnement. Je vous remercie aussi pour les nombreuses opportunités que vous m'avez offertes, notamment celle de contribuer à ma première publication scientifique.

Je remercie Guillaume Vallet et Benjamin Boller, membres de mon comité d'évaluation, pour la richesse de leurs retours, la justesse de leurs conseils et l'attention portée à mon travail, qui a permis d'en améliorer significativement la qualité. Je remercie

également Pierre Maurage et Fabien D'Hondt pour leur accueil dans leur laboratoire et leur confiance.

Je suis également reconnaissante envers mes collègues du laboratoire : Simon, Gasser, Mathias et Annabelle. Au-delà de la recherche, vous m'avez accueillie, fait découvrir le Québec et partagé votre culture. Votre amitié, votre écoute et votre humour ont été un vrai soutien au quotidien. Vous serez toujours les bienvenus en Europe, et je répondrai toujours présente à chaque : « demande à Anne-So »!

Je ne saurais clore ces remerciements sans évoquer Laetitia Tiri, ma meilleure amie, ma partenaire de vie, ma femme. Merci d'avoir tout quitté pour me suivre dans cette folle aventure de doctorat au Canada. Ton amour, ta confiance en moi, ta patience face aux doutes, ton humour, ta sensibilité et ton énergie m'ont portée à chaque étape. Tu as été ma force dans l'aboutissement de ce travail. Merci de me pousser à croire en moi, de m'écouter, et de me faire sortir de ma grotte quand cela est nécessaire. Merci d'écrire cette vie à mes côtés depuis maintenant sept ans. J'ai hâte de vivre ce nouveau chapitre de notre histoire avec la naissance de notre petite fille, Maxine.

Je remercie également ma famille : Christine, Gilbert et Monique. Même si vous n'avez pas toujours compris en quoi consistait exactement mon travail, vous m'avez toujours soutenue, encouragé et montré votre fierté. Cela m'a donné la force d'avancer.

Enfin, je tiens à me remercier moi-même. Une thèse représente un investissement important sur plusieurs années, une aventure à la fois intellectuelle et humaine, exigeante, formatrice et souvent bouleversante. J'ai persévéré malgré les difficultés, cru en mes capacités et appris qu'il fallait parfois ralentir pour mieux avancer. Je suis fière d'être allée au bout de ce projet et d'avoir mené un travail de recherche rigoureux et ambitieux.

Introduction générale

Imaginez une scène dans un café à Lille, en France : deux personnes sont assises à une table. Le visage d'une femme affiche des sourcils froncés et des lèvres serrées, tandis qu'elle parle rapidement en gesticulant. Ses mains sont tendues, ouvertes devant elle, et son corps légèrement incliné vers l'avant. Un homme assis en face d'elle évite son regard, les bras croisés sur la poitrine et la mâchoire crispée. Dans cette mise en situation, il semble évident que ces deux personnes ressentent et expriment de la colère. Leurs émotions sont rapidement identifiables grâce aux indices visuels congruents, c'est-à-dire que leurs expressions faciales (par exemple, les sourcils froncés, les lèvres serrées) et corporelles (par exemple, les bras croisés, les mains tendues) expriment une émotion similaire, la colère. Maintenant, imaginons une variante dans laquelle la femme exprime des émotions plus ambiguës synonymes d'une ambivalence dans son ressenti émotionnel. Elle sourit, ses yeux pétillent, et ses sourcils sont légèrement relevés, mais sa posture reste rigide, les bras croisés et les poings serrés, comme si elle voulait dissimuler une émotion qu'elle ne souhaite pas laisser transparaître. Dans ce cas, l'émotion qu'elle exprime devient plus difficile à identifier. Ses expressions faciales et corporelles sont incongruentes, c'est-à-dire exprimant des émotions différentes tels un visage joyeux et un corps de colère, entravant la reconnaissance émotionnelle (Kret, Stekelenburg et al., 2013; Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013). Ces deux mises en situation soulignent non seulement l'importance des expressions faciales et corporelles dans la reconnaissance des émotions, mais aussi l'influence réciproque de ces deux

canaux. En effet, lorsque ces canaux sont perçus simultanément, leurs indices émotionnels s'intègrent facilitant ou perturbant la reconnaissance émotionnelle (Lecker et al., 2020). Cette influence réciproque entre les expressions faciales et corporelles dans la reconnaissance émotionnelle, nommée habituellement dans la littérature « l'effet contextuel », suggère que la perception des émotions repose sur une intégration de plusieurs canaux non verbaux permettant ainsi une reconnaissance rapide et précise (Kret, Stekelenburg et al., 2013; Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013; Willis et al., 2011; Zhang et al., 2019). Cependant, bien que l'effet contextuel soit bien documenté dans la littérature scientifique, des questions subsistent. Une de ces questions concerne le caractère automatique ou volontaire de ce traitement. Certaines études montrent que cet effet survient de manière automatique, c'est-à-dire sans nécessiter de conscience (Karaaslan et al., 2020) ni mobiliser de ressources cognitives (Aviezer et al., 2011), et se réalisant rapidement (Meeren et al., 2005). D'autres, en revanche, indiquent que cet effet est modulé par des consignes spécifiques, par exemple en demandant aux participants de focaliser leur attention uniquement sur les expressions faciales ou corporelles, remettant ainsi en question l'automatisme de l'effet contextuel (Lecker et al., 2020) car s'il était automatique, il serait identique, quelles que soient les consignes données.

Le cadre théorique de cette thèse se divise en trois sections principales. La première et seconde section se concentrent sur la perception des expressions faciales et corporelles de manière isolée. Ces deux types d'expressions émotionnelles ont historiquement été

étudiés séparément, ce qui a permis de mieux comprendre leur contribution spécifique à la reconnaissance émotionnelle. Bien que les expressions faciales aient été au centre de nombreuses recherches, il est important de souligner que les expressions corporelles jouent également un rôle crucial dans la reconnaissance émotionnelle et mobilisent des traitements similaires aux expressions faciales. La troisième section explore les interactions bidirectionnelles entre les expressions faciales et corporelles lorsqu'elles sont perçues simultanément. Initialement, les recherches se sont intéressées aux expressions corporelles en tant que contexte aux expressions faciales, montrant que des indices corporels congruents avec des expressions faciales facilitent la reconnaissance des émotions faciales, contrairement à des expressions incongruentes (Kret, Stekelenburg et al., 2013; Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013; Willis et al., 2011; Zhang et al., 2019). Plus récemment, les études ont montré que les expressions faciales agissent également comme contexte émotionnel pour la perception des expressions corporelles, soulignant ainsi que les indices faciaux peuvent moduler la reconnaissance des expressions corporelles (Abo Foul et al., 2018; Lecker et al., 2020; Reschke & Walle, 2021). Enfin la quatrième section se concentre sur les mécanismes sous-jacents à l'intégration des expressions faciales et corporelles, en particulier sur leur caractère automatique.

Expressions faciales émotionnelles isolées

Historiquement, la recherche sur la perception des émotions s'est focalisée sur les expressions faciales. Cette prédominance s'explique notamment par le rôle central

qu'elles jouent dans notre communication quotidienne lors de conversations, de présentations orales ou même en visionnant des films.

Dans cette section, nous examinerons les principales théories ayant contribué à une meilleure compréhension de la nature et du rôle des expressions faciales et nous analyserons les différentes activations cérébrales impliquées dans le traitement des expressions faciales émotionnelles.

Fondements théoriques majeurs des expressions faciales émotionnelles

Les expressions faciales, qui regroupent des activations musculaires telles que l'activation des zygomatiques et corrugateurs, le retroussement du nez ou le froncement des sourcils, jouent un rôle essentiel dans notre vie quotidienne en remplissant plusieurs fonctions. Elles permettent non seulement de communiquer notre état émotionnel à un interlocuteur, mais aussi de signaler l'élaboration d'un processus cognitif et de réguler nos interactions sociales (Sander & Scherer, 2019). Par exemple, un sourire, produit de la contraction des zygomatiques, est généralement associé à de la joie, mais il peut également encourager à poursuivre une conversation ou indiquer que le sujet discuté est plaisant. De manière similaire, un froncement de sourcils, lié à la contraction des muscles corrugateurs, peut refléter de l'incompréhension tout en signalant un état de concentration ou de réflexion. Cela fait des expressions faciales des outils de communication multifonctionnels, fondamentaux pour nos interactions sociales (Sander & Scherer, 2019).

L'importance des expressions faciales dans la vie quotidienne a suscité un intérêt considérable parmi les chercheurs, qui ont exploré leur fonction et leur influence sur le comportement.

Théorie des émotions discrètes

Charles Darwin fut l'un des premiers à amener l'idée que nos états internes s'extériorisaient à travers nos expressions faciales émotionnelles. Il a notamment mis en lumière au travers de son ouvrage « *The expression of emotions in man and animals* » (Darwin, 1872) que ces expressions faciales avaient des fonctions adaptatives. Dans ce livre, il examine la relation entre les expressions faciales émotionnelles et les comportements humains, suggérant que les émotions sont associées à des modifications comportementales spécifiques, telles que l'évitement ou l'approche (Darwin & Prodger, 1998). Ces observations ont inspiré de nombreuses théories, parmi lesquelles celle des émotions discrètes développée par Paul Ekman (Ekman, 1992, 1999; Ekman & Friesen, 1976). Selon Ekman, les émotions se limitent à un ensemble restreint d'émotions dites « de base » ou « primaires » comprenant la joie, la tristesse, la colère, la peur, le dégoût et la surprise. Ces émotions de base sont caractérisées par des réponses physiologiques distinctes, telles qu'une augmentation du rythme cardiaque ou de la sudation. Elles s'accompagnent également de réactions préprogrammées et automatiques, souvent déclenchées par des événements, comme une menace ou une perte. De plus, Ekman propose que ces émotions se manifestent par des expressions faciales universelles, observables à travers différentes cultures et présentes chez les primates, suggérant un

fondement évolutif commun. Enfin, ces émotions se distinguent également par leur rapidité, leur courte durée, et leur caractère spontané (Ekman, 1992, 1999).

Bien que la théorie des émotions discrètes ait profondément influencé la recherche sur les expressions faciales émotionnelles, en particulier sur leurs liens avec les réponses physiologiques (James, 1994; LeDoux, 1990; Sander & Scherer, 2019), elle présente une limite conceptuelle majeure. En effet, en postulant un nombre restreint d'émotions universelles (telles que la joie, la colère ou la peur), cette approche ne permet pas de rendre compte de la richesse et de la complexité de l'expérience émotionnelle humaine. Elle tend ainsi à exclure des émotions dites sociales ou secondaires comme la honte, la jalousie ou l'émerveillement, qui jouent pourtant un rôle central dans nos interactions sociales. Cette limite n'est pas retrouvée dans d'autres modèles, tels que les modèles dimensionnels, qui proposent une représentation plus nuancée des émotions plutôt qu'une catégorisation fixe.

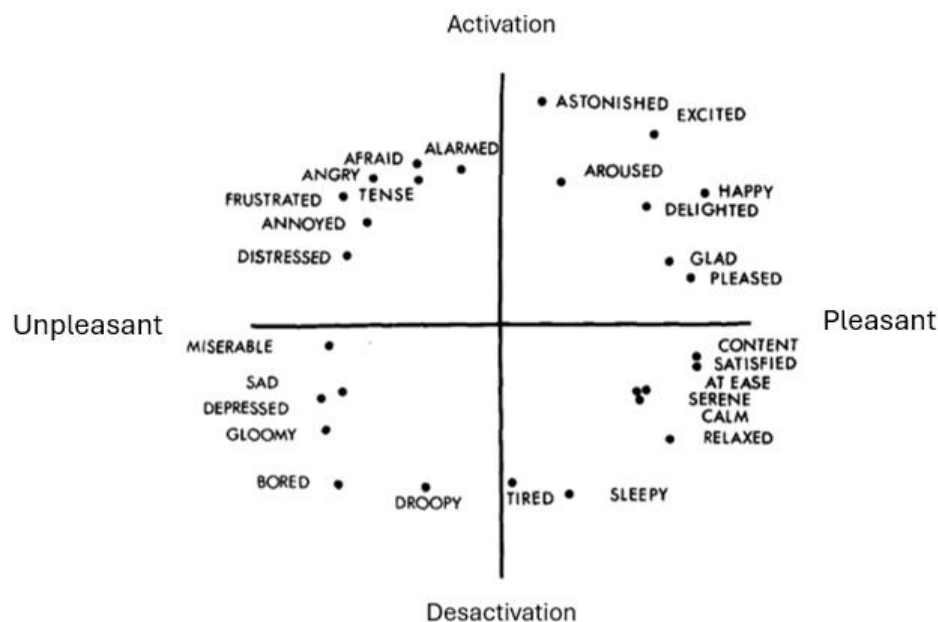
Théories dimensionnelles

Russell a développé une approche alternative, la théorie dimensionnelle des émotions, postulant que les expressions émotionnelles peuvent être représentées dans un espace bidimensionnel défini par la valence (positif vs. négatif) et le niveau d'activation (haute vs. Élevée; Russell, 1980; voir Figure 1). Selon cette théorie, l'évaluation d'une expression faciale émotionnelle chez autrui se déroule en deux étapes. La première, automatique et universelle, consiste à juger les niveaux de valence et d'activation de l'émotion perçue. La seconde étape, plus complexe et influencée par des facteurs culturels,

correspond à la catégorisation émotionnelle. Celle-ci repose non seulement sur l'évaluation initiale de la valence et de l'activation, mais également sur des indices contextuels issus de l'environnement de l'interaction (Russell, 1980, 1994; Russell & Fernández-Dols, 1997). Ainsi, Russell propose une conceptualisation des émotions à travers un nombre non fini d'émotions, suggérant qu'elles s'organisent selon un espace bidirectionnel et que leur reconnaissance repose sur un processus en deux temps : une première phase d'évaluation automatique et universelle, suivi d'une catégorisation plus élaborée et modulée par le contexte et les spécificités culturelles (Sander & Scherer, 2019).

Figure 1

Modèle dimensionnel de Russell (1980)



Note. Modèle dimensionnel de Russell sur lequel l'axe horizontal représente la valence et l'axe vertical l'activation. Cette image est reprise de l'article de Russell (1980).

Le modèle dimensionnel de Russell est largement utilisé dans la littérature scientifique, mais il demeure controversé. Il est notamment critiqué pour son caractère simpliste, ne prenant pas en compte l'ensemble des caractéristiques constitutives d'une expression faciale émotionnelle. A la suite du modèle de Russell, d'autres modèles dimensionnels ont émergé, notamment celui de Gao et ses collègues (2010). Ce modèle positionne les sept émotions de base (joie, dégoût, tristesse, peur, surprise, colère et neutre) selon quatre dimensions : la valence (positif vs. négatif), l'activation (haut vs. bas), la potentialité (faiblesse vs. puissance) et l'intensité (forte vs. faible) (Gao et al., 2010). Il enrichit ainsi le modèle de Russell en y intégrant les dimensions de la potentialité et de l'intensité. L'axe de la potentialité reflète l'évaluation du contrôle exercé dans une interaction sociale. Des valeurs élevées sur cette dimension sont associées à des sentiments de dominance et une impulsion à agir, tandis que des valeurs faibles traduisent des sentiments de soumission et une inhibition à agir. Par ailleurs, Gao et ses collègues mettent en évidence des variations développementales dans ces dimensions. En effet, les enfants de 7 ans ne semblent pas structurer les émotions selon les mêmes dimensions que les adultes, ni regrouper les expressions faciales de manière similaire. Notamment, l'absence de la dimension de potentialité chez les enfants suggère qu'elle émerge plus tard au cours du développement. De plus, certains regroupements observés chez l'adulte, comme l'association entre la surprise et la peur, ne sont pas présents chez l'enfant, ce qui suggère des différences dans les capacités cognitives impliquées dans la perception et la catégorisation des émotions (Gao et al., 2010). Ce modèle, plus complexe que celui de Russell, intègre des dimensions émotionnelles supplémentaires approfondissant la

compréhension des mécanismes perceptifs et cognitifs sous-jacents à l'interprétation des expressions faciales dans les interactions quotidiennes.

Bien que les modèles dimensionnels, et en particulier celui de Russell (1980), soient encore aujourd'hui largement utilisés comme cadre de référence pour étudier les expressions faciales émotionnelles, ils présentent certaines limites. A l'instar du modèle d'Ekman qui postule l'universalité des émotions discrètes, le modèle dimensionnel de Russell suppose que les dimensions émotionnelles de valence et d'activation sont universelles, c'est-à-dire communes à tous les individus, indépendamment de leur culture. Ce postulat a toutefois été remis en question par plusieurs travaux soulignant le rôle déterminant de facteurs culturels dans la reconnaissance émotionnelle. Par exemple, dans l'étude de Gendron et ses collègues (2014), des participants issus de la culture américaine et de la culture Himba (Namibie) devaient trier des expressions faciales selon l'émotion exprimée. Les résultats ont montré que les participants Himba ne catégorisaient pas les émotions de la même manière que les participants américains. Alors que ces derniers utilisaient principalement des étiquettes émotionnelles discrètes (comme « joie » ou « colère »), les Himbas tendaient à se baser sur des comportements observables (comme « rire » ou « regarder quelque chose »), suggérant une préférence pour les descripteurs d'actions plutôt que pour ceux relatifs aux états mentaux. Ces résultats remettent en cause l'universalité des émotions postulée par Ekman ou supposée par Russell, et appuient l'idée d'une influence culturelle marquée dans la perception émotionnelle (Gendron et al., 2014). Par ailleurs, si les modèles dimensionnels ont l'avantage de proposer une représentation

simple et synthétique des émotions à travers des dimensions, ils négligent un aspect fondamental du fonctionnement émotionnel, c'est-à-dire leur fonction adaptative. En effet, ces modèles peinent à expliquer comment une émotion spécifique prépare à une réponse comportementale particulière. Par exemple, bien que la peur et la colère puissent partager une valence négative et un niveau d'activation élevé, elles entraînent des réponses comportementales différentes telles que l'évitement ou la confrontation. C'est dans cette optique que des modèles centrés sur la fonction des émotions, comme celui de Frijda (1989), apportent un éclairage complémentaire, en soulignant le rôle des émotions dans la préparation à l'action.

Expressions faciales émotionnelles comme préparation à l'action

Des chercheurs se sont intéressés aux émotions en prenant un angle conceptuel différent. Notamment, Frijda a conceptualisé les émotions faciales comme des mécanismes de préparation à l'action (Frijda, 1987, 1988; Frijda et al., 1989). Selon cette théorie, les expressions faciales ne reflètent pas uniquement des émotions spécifiques, mais constituent des indicateurs de la manière dont un individu interagit avec son environnement et se prépare à agir avec lui. La préparation à l'action implique une orientation comportementale, qui peut être une tendance à l'approche ou à l'évitement, en fonction des objectifs et motivations de l'individu. Ainsi, les expressions faciales renseignent l'observateur sur les intentions d'action d'une personne, lui permettant d'ajuster son propre comportement en conséquence. Frijda identifie par exemple le concept d'attente environnementale, selon lequel l'expression faciale d'un individu dirige

l'attention de l'observateur vers un élément précis de l'environnement. Dans ce cadre, lorsqu'une personne exprime une expression faciale de peur, cela signale la présence d'un potentiel danger dans l'environnement orientant ainsi l'attention des autres personnes vers l'objet menaçant (Frijda, 1987).

L'approche de Frijda conçoit les expressions faciales émotionnelles sous l'angle d'un processus d'interaction avec l'environnement. Dans ce contexte, l'expression faciale est un outil d'interaction sociale fournissant des informations sur la manière dont l'individu tend à interagir avec son environnement immédiat.

Expressions faciales émotionnelles comme une évaluation cognitive

D'autres chercheurs ont, quant à eux, mis l'accent sur le rôle des processus cognitifs dans la génération des émotions. Ces théories postulent que les émotions faciales résultent d'une évaluation cognitive complexe de l'environnement (Lazarus, 2000; Scherer, 2009). Autrement dit, l'émotion ressentie n'est pas une réaction immédiate et automatique aux stimuli extérieurs, mais découle d'une analyse subjective de la situation par l'individu (Sander & Scherer, 2019). L'une des premières chercheuses à avoir formalisé cette perspective est Arnold (1960) proposant un modèle en quatre étapes du traitement émotionnel (voir Figure 2). D'abord, l'individu perçoit la situation dans laquelle il se trouve. Ensuite, il évalue cette situation comme étant bénéfique ou nuisible à son bien-être. Sur la base de cette évaluation, une émotion spécifique est générée, suivie d'une réaction comportementale cohérente avec l'émotion ressentie (Arnold, 1960). Ce modèle met en

évidence le rôle fondamental de l'évaluation cognitive dans la production des émotions et souligne le lien étroit entre l'expérience émotionnelle, l'expression faciale et l'action.

Figure 2

Représentation de la théorie d'évaluation cognitive des émotions d'Arnold



Dans une approche similaire, Lazarus (2000) a établi un modèle distinguant deux types d'évaluations successives : l'évaluation primaire déterminant si une situation est pertinente en fonction des objectifs et du bien-être d'un individu et l'évaluation secondaire permettant d'apprécier les ressources disponibles pour faire face aux conséquences de la situation. Ce processus d'évaluation successif conduit à une catégorisation émotionnelle, expliquant ainsi la diversité des émotions ressenties en fonction du contexte et des ressources perçues par l'individu (Lazarus, 2000).

Finalement, Scherer (2009) a proposé un modèle dynamique et multifactoriel, connu sous le nom de *Component Process Model* (CPM). Ce modèle postule que l'évaluation d'un événement ne repose pas uniquement sur une évaluation dichotomique (positif/négatif), mais sur une série d'évaluations cognitives successives qui interagissent entre elles et influencent les réponses émotionnelles. Lorsqu'un individu est confronté à un événement, il effectue d'abord une évaluation de la nouveauté de la situation,

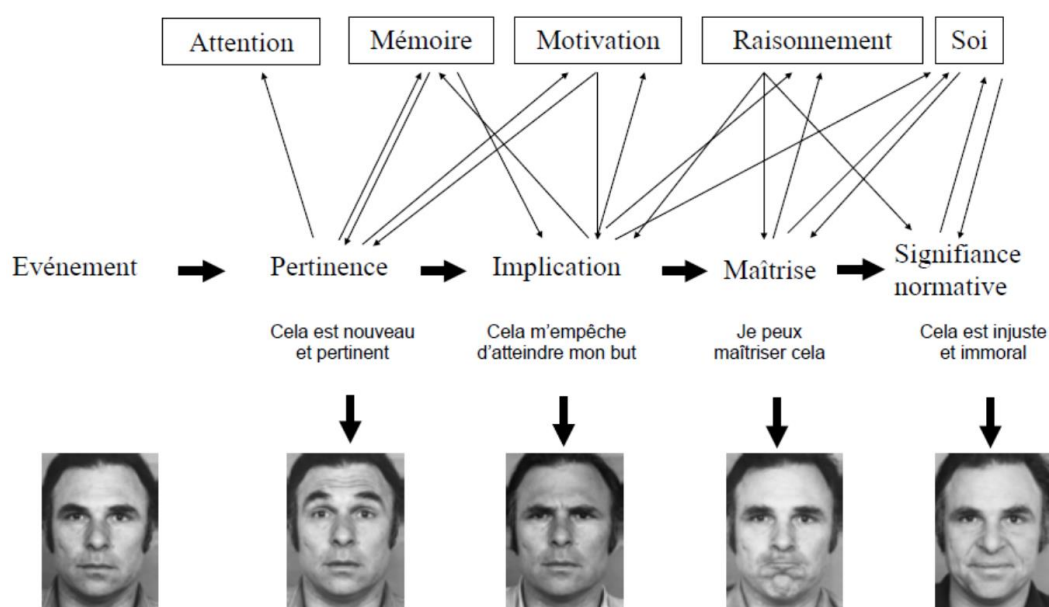
déterminant ainsi si celle-ci est familière ou inattendue. Il évalue ensuite la congruence aux buts, en jugeant si l'évènement est en adéquation avec ses objectifs et ses besoins. Cette évaluation est suivie par une analyse de l'agentivité, qui permet d'identifier la source de la situation, qu'il s'agisse de l'individu lui-même, d'autrui ou d'un facteur externe. Par la suite, l'individu évalue la compatibilité morale de l'évènement, en déterminant s'il est conforme à ses valeurs personnelles. Enfin, il estime sa capacité d'adaptation, c'est-à-dire s'il dispose des ressources nécessaires pour gérer efficacement la situation (Scherer, 2009). Le CPM constitue donc une avancée considérable à travers la conceptualisation de plusieurs types d'évaluations successive et les interactions entre ces évaluations influençant directement les réponses émotionnelles (voir Figure 3). Par ailleurs, le modèle postule que ces réponses ne sont pas seulement des réactions passives, mais qu'elles modifient en retour l'évaluation cognitive de la situation, générant ainsi un processus d'adaptation dynamique (Scherer, 2009).

L'ensemble des théories et modèles présentés met en évidence la diversité des approches dans la conceptualisation des émotions. Alors que certaines théories amènent à penser les émotions comme des réponses à des changements physiologiques (James, 1894, 1994) ou encore comme des tendances à l'action (Frijda et al., 1989), d'autres les représentent comme un ensemble de processus cognitifs complexes (Scherer, 2009) intervenant dans leur expression. Comprendre la définition des émotions et la manière dont elle s'intègre dans le fonctionnement cognitif et comportemental des individus requiert donc une vision multifactorielle. C'est pourquoi, d'autres travaux ont adopté une

perspective neuroscientifique, cherchant à établir un lien entre l'activité cérébrale et la perception émotionnelle.

Figure 3

Représentation du Component Process Model (CPM)



Note. Image reprise du livre « traité de psychologie des émotions » de David Sander et Klaus R. Scherer (2009).

Corrélatés neuronaux de la perception des expressions faciales isolées

Afin d'étudier le déroulement temporel dans le traitement émotionnel au niveau de l'activité cérébrale, l'électro-encéphalographie de surface (EEG) est une technique de neuroimagerie privilégiée. Cette méthode non-invasive enregistre l'activité électrique du cerveau à l'aide d'électrodes placées sur le scalp de l'individu. Grâce à un ensemble de techniques d'analyse du signal, les données EEG retracent précisément le déroulement temporel des processus cognitifs qui suivent la présentation d'un stimulus sensoriel, en se

basant sur l'analyse des potentiels évoqués (PE). Un potentiel évoqué correspond à une réponse électrique du cerveau, qui se manifeste sous forme de pics à différentes millisecondes après la présentation d'un stimulus. Chaque composante d'un PE est caractérisée par une amplitude (positive ou négative) et une latence (intervalle de temps entre la présentation d'un stimulus et la pointe (*peak*) d'une composante), et reflète un processus cognitif spécifique (Luck, 2014) (voir Figure 4).

Les trois stades du traitement des expressions faciales émotionnelles

Sur la base de données EEG, il a été proposé que le traitement des expressions faciales émotionnelles suit une série de trois phases distinctes (voir Figure 5). La première correspond à un traitement automatique et rapide, spécifiquement orienté vers les expressions faciales négatives. La seconde phase introduit une différenciation entre les expressions faciales neutres et émotionnelles. Enfin, la dernière phase permet une distinction fine entre les différentes émotions spécifiques (Luo et al., 2010).

Figure 4

Étapes de l'acquisition et du pré-traitement du décours temporel de l'activité cérébrale par EEG

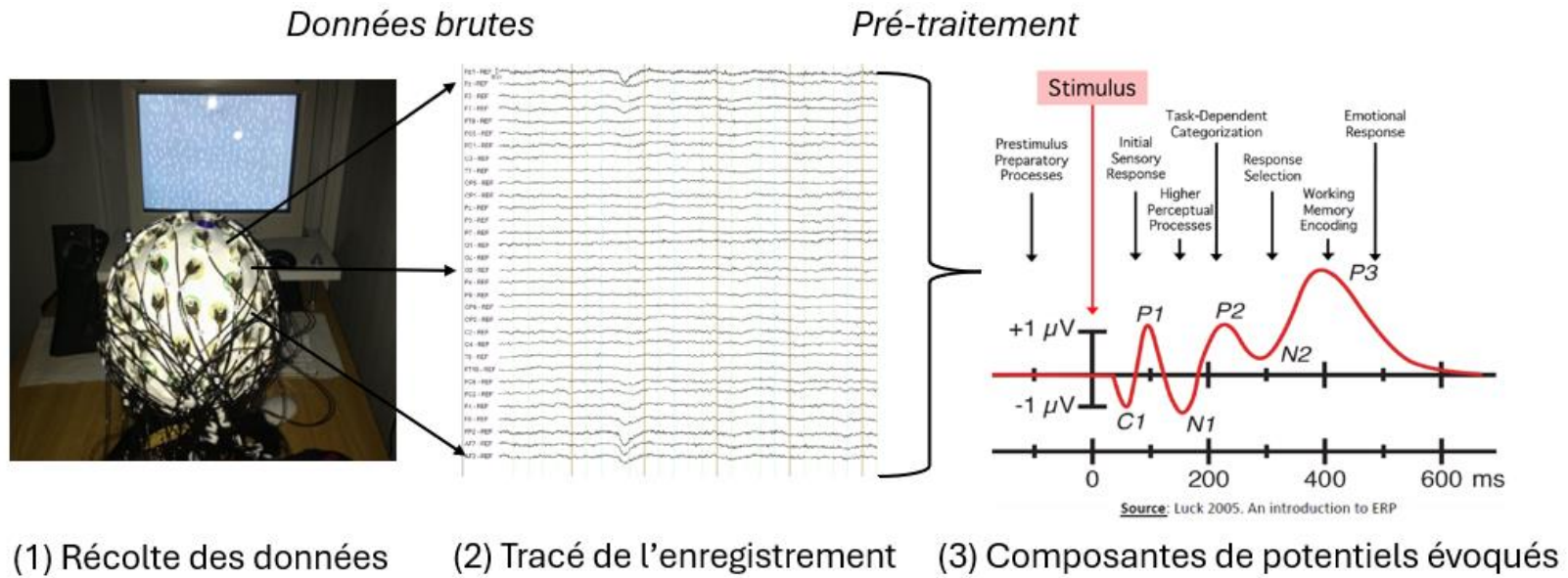
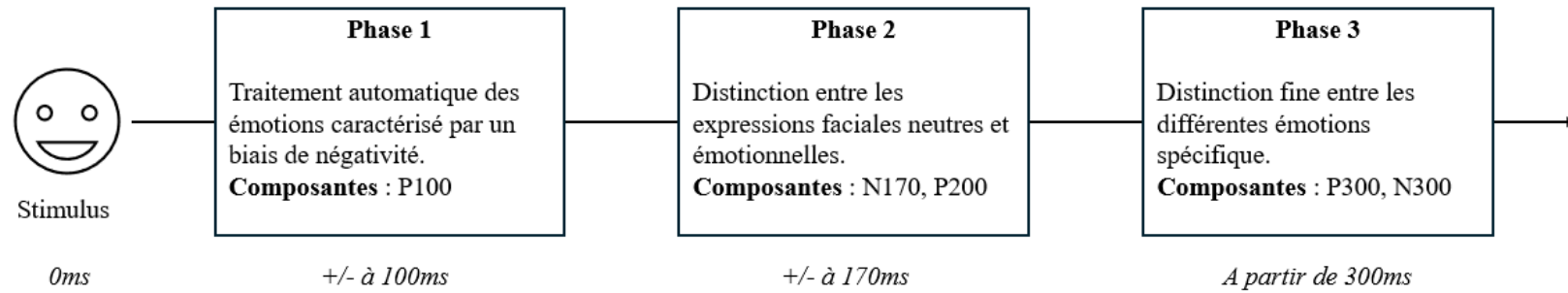


Figure 5

Schéma du modèle en trois phases du traitement des expressions faciales émotionnelles de Luo et al. (2010)



La *première phase* précoce est caractérisée par la composante P100. Il s'agit d'une composante positive localisée principalement dans les zones pariéto-occipitales, apparaissant entre 60 et 120ms après la présentation d'un stimulus, avec un pic autour de 100ms (Batty & Taylor, 2003; Luo et al., 2010; Utama et al., 2009). L'implication de la P100 dans le traitement émotionnel est cependant actuellement controversée dans la littérature. En effet, certaines études ont montré que l'amplitude de la P100 était modulée par la perception d'expressions faciales émotionnelles négatives, telles que la peur ou la colère, par rapport aux expressions positives comme la joie, suggérant un traitement automatique et précoce des expressions négatives dès les premières étapes (Batty & Taylor, 2003; Luo et al., 2010). Cette observation suggère l'existence d'un biais de négativité précoce dans le traitement des expressions émotionnelles faciales (Luo et al., 2010). Ce biais pourrait être expliqué par des impératifs adaptatifs, la détection rapide des signaux de menace étant essentielle à la survie en facilitant une réponse automatique aux stimuli potentiellement dangereux (Santos et al., 2008). D'autres études ne trouvent aucun effet des expressions faciales émotionnelles sur la P100 (Calvo et al., 2013; Durston & Itier, 2021, 2025). Ces résultats divergents dans la littérature sont notamment expliqués par le fait que cette composante précoce est particulièrement sensible aux caractéristiques de bas niveau des stimuli, tels que la couleur, la luminosité ou le contraste (Puce et al., 2013; Yang et al., 2020) ainsi qu'au paradigme expérimental utilisé (Schindler & Bublatzky, 2020). Afin de comprendre l'implication de la P100 dans le traitement précoce des expressions faciales émotionnelles, Durston et Itier (2025) ont réalisé une étude de reconnaissance émotionnelle sur 80 personnes en contrôlant la luminance et l'intensité des

pixels des expressions faciales, et en utilisant un grand nombre de stimuli provenant de quatre banques de stimuli différents (NimStim, Chicago Face Database, Radboud Face Database, et FACES databases) (Ebner et al., 2010; Langner et al., 2010; Ma et al., 2015; Tottenham et al., 2009) afin de prendre en compte les potentielles différences d'intensité émotionnelle entre les banques de stimuli. Les résultats de cette étude ne suggèrent aucune influence des émotions ni des niveaux de valence et d'activation des expressions faciales perçues sur la P100 remettant en considération l'implication de cette composante précoce dans le traitement automatique des expressions faciales émotionnelles (Durston & Itier, 2025).

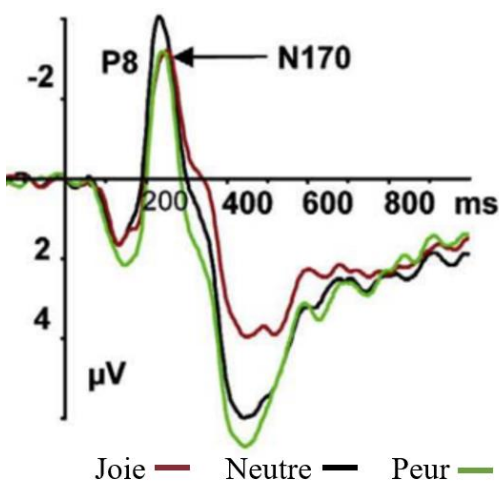
La *seconde phase de traitement* est principalement caractérisée par l'activité de la composante N170. Celle-ci a une polarisation négative et est située dans les zones occipito-temporal autour de 120 et 220ms avec un pic à 170ms après la présentation du stimulus. Le traitement des expressions faciales est sous-tendu par l'activation de plusieurs régions cérébrales, notamment les aires *fusiformes face* (FFA) et *occipital face* (OFA) située respectivement dans le gyrus fusiforme et le cortex occipital (Gauthier et al., 2000; Kanwisher et al., 1997). L'aire FFA, située dans le gyrus fusiforme latéral moyen, joue un rôle crucial dans la génération de la composante N170, qui est sensible aux expressions faciales (Deffke et al., 2007) et particulièrement impliquée dans l'analyse de la configuration faciale (Schindler et al., 2017; Sollfrank et al., 2021). En effet, la perception des expressions faciales semble reposer sur un traitement configurationnel globale, contrairement à la reconnaissance des objets non-faciaux, qui s'appuie davantage

sur l'identification de leurs parties isolées. Ce phénomène, bien connu dans la littérature, est principalement mis en évidence par l'effet d'inversion des visages suggérant que lorsque les visages sont présentés à l'envers, la composante N170 présente une amplitude plus élevée et un délai plus important que lorsque les visages sont présentés à l'endroit. En revanche, ce même effet n'est pas observé pour les objets, suggérant que la N170 est spécifiquement impliquée dans le traitement configurationnel des visages (Watanabe et al., 2003). Plus récemment, cette implication de la N170 a été explorée par Schindler et ses collègues (2017), qui ont examiné sa sensibilité à différents niveaux d'abstraction des visages. Pour ce faire, ils ont présenté des stimuli avec six degrés de stylisation, allant du plus abstrait au plus réaliste, permettant ainsi d'évaluer dans quelle mesure cette composante reflète le traitement configurationnel du visage. Les résultats ont révélé une modulation de la N170 avec des amplitudes plus élevées pour les visages les plus abstraits et les plus réalistes, comparativement aux niveaux intermédiaires. Ces résultats suggèrent que la N170 est sensible aux aspects structurels des expressions faciales et pourraient refléter un traitement différencié en fonction du degré de stylisation des stimuli faciaux (Schindler et al., 2017). Par ailleurs, la N170 est également plus fortement activée en réponse aux visages humains qu'à ceux d'autres espèces, comme les primates, indiquant une spécificité de cette composante au traitement des visages humains uniquement (Gajewski & Stoerig, 2011). Finalement, la N170 est sensible aux expressions faciales émotionnelles ainsi qu'à la valence et l'activation de ces émotions faciales. Notamment, l'étude de Durston et Itier (2025) mentionnée plus haut a montré une différenciation d'amplitude de la composante N170 en fonction des niveaux d'activation et de valence perçues sur des expressions faciales

émotionnelles de peur, joie et colère contrairement à d'autres composantes comme la P200 (Durston & Itier, 2025), renforçant l'idée que la N170 joue un rôle dans la détection de la structure des expressions faciales. Par ailleurs, l'amplitude de la N170 est plus importante lors de perception d'expressions faciales émotionnelles plutôt que neutre suggérant qu'une différenciation entre du contenu émotionnel et neutre, mais pas encore entre les émotions à ce stade (Luo et al., 2010) (voir Figure 6).

Figure 6

Graphique des PEs pour la joie, la neutralité et la peur sur l'électrode P8



Note. Cette image provient de l'article de Luo et al. (2010).

A contrario, une méta-analyse réalisée par Hinojosa et ses collègues (2015) a mis en évidence que l'amplitude de la N170 était significativement différente lors de la perception d'expressions faciales émotionnelles, notamment entre la colère, la peur et la joie (Hinojosa et al., 2015). Cela a également été observé plus récemment dans d'autres études notamment utilisant des stimuli d'expressions faciales joyeuses et peureuses

dynamiques, suggérant que l'utilisation d'expressions faciales plus proche d'une perception réaliste impacte le traitement émotionnel (Jacques & Caharel, 2022). Des divergences persistent donc concernant le rôle de la composante N170 dans la reconnaissance et, plus spécifiquement, dans la différenciation des émotions faciales. Ces variations pourraient être attribuées à des différences dans le degré de réalisme des stimuli émotionnels utilisés. Malgré ces contradictions dans la littérature, il semble que la N170 joue un rôle clé dans le traitement intermédiaire des visages, en intégrant à la fois des aspects structuraux et émotionnels des expressions faciales. Cependant, d'autres recherches ont quant à elle remis en cause la place centrale de la N170 dans le traitement des expressions faciales. Notamment l'étude de Tanaka et Curran (2001) a montré que la N170 était modulée par le niveau d'expertise des participants. Les auteurs partent du postulat que nous sommes experts en reconnaissance d'expressions faciales émotionnelles et que la N170 n'est pas modulée par les expressions faciales elles-mêmes, mais notre expertise envers elles. Afin de tester cette hypothèse, lors de leur étude, les chercheurs ont comparé deux groupes de participants, un groupe expert en chien et un autre expert en oiseaux. La tâche des participants était de catégoriser des noms de chiens ou d'oiseaux à l'image correspondante. Durant cette tâche, l'activité cérébrale des participants était enregistrée à l'aide d'un EEG. Les résultats de l'étude ont montré qu'effectivement l'amplitude de la N170 était plus importante pour les experts en oiseau lorsqu'ils percevaient des images d'oiseau plutôt que des chiens et inversement pour les experts en chien (Tanaka & Curran, 2001). Cette étude est l'une des premières à remettre en considération la spécificité de la N170 dans le traitement des expressions faciales.

D'autres études ont également remis en cause la spécificité de la N170 aux expressions faciales en montrant que certains effets tels que la sensibilité à l'inversion et aux modifications configurationnelles de la N170 pouvaient également provenir d'un effet d'expertise (Bukach et al., 2006; Gauthier & Curby, 2005; Gauthier & Tarr, 2002).

La *dernière phase du traitement* des émotions est principalement caractérisée par les composantes N300 et P300, apparaissant autour de 300 ms après la présentation du stimulus, reflètent une analyse plus approfondie du contenu émotionnel des expressions faciales (Campanella et al., 2002). La composante N300 est située dans la zone fronto-temporale et la P300, quant à elle est divisé en deux sous-composantes : la p3a localisée dans les zones frontales et la p3b localisée dans les zones pariétales (Polich, 2007). Bien que la différenciation entre les émotions puisse déjà émerger à des stades plus précoces (Hinojosa et al., 2015; Jacques & Caharel, 2022), plusieurs études suggèrent que le cerveau opère une distinction plus marquée entre les différentes émotions autour de 300 ms après la présentation du stimulus (Luo et al., 2010). Par exemple, la perception d'une expression faciale de colère est associée à une amplitude plus importante de la N300 comparativement aux expressions de joie ou neutres (Schutter et al., 2004). A ce stade, l'analyse des expressions faciales émotionnelles devient plus élaborée, impliquant des processus d'évaluation cognitive. Par ailleurs, il a été montré que l'orientation du regard joue un rôle clé dans ce traitement cognitif, indiquant une intégration des indices attentionnels dans l'évaluation émotionnelle à cette étape tardive du traitement (McCrackin & Itier, 2021). Ces résultats soulignent l'interaction entre les mécanismes

attentionnels et émotionnels, mettant en évidence la manière dont le cerveau ajuste dynamiquement son traitement des émotions en fonction des indices contextuels disponibles.

En conclusion, au fil des décennies, de nombreuses théories ont souligné l'importance des informations émotionnelles véhiculées par les expressions faciales et leur rôle central dans la régulation des interactions sociales. Ces signaux non verbaux influencent non seulement la perception des états émotionnels d'autrui, mais également les comportements qui en découlent (Frijda, 1987; Gao et al., 2010; Scherer, 2009). D'un point de vue neurocognitif, le traitement des expressions faciales émotionnelles s'opère à travers plusieurs étapes successives, allant d'un traitement précoce, automatique et rudimentaire à une analyse plus tardive et élaborée, chacune étant associée à des composantes spécifiques (Batty & Taylor, 2003; Hinojosa et al., 2015; Jacques & Caharel, 2022; Luo et al., 2010; McCrackin & Itier, 2021; Utama et al., 2009).

Malgré l'ensemble de la littérature sur la perception des expressions faciales émotionnelles, il semblerait qu'entre 2000 et 2020, la méthodologie employée dans ce domaine de recherche a peu évolué, restant majoritairement centrée sur l'utilisation de banques de stimuli statiques et standardisés, dans lesquelles des paramètres tels que la luminosité, le contraste et l'orientation du visage sont rigoureusement contrôlés (Dawel et al., 2022). Ces stimuli normalisés sont essentiels pour garantir la reproductibilité des résultats, mais ils négligent la spontanéité, le dynamisme et la mise en contexte des

expressions faciales émotionnelles (Dawel et al., 2022), pouvant pourtant avoir un impact potentiel sur le traitement de ces expressions émotionnelles (Jacques & Caharel, 2022). Par ailleurs, dans la vie quotidienne, la reconnaissance des émotions ne repose pas uniquement sur le visage, mais implique une intégration multisensorielle de plusieurs canaux sensoriels. En particulier, les expressions corporelles jouent un rôle clé dans l'identification et l'interprétation des émotions. Postures, gestes et mouvements corporels contribuent de manière significative à l'élaboration du signal émotionnel, en renforçant ou en modulant les informations véhiculées par le visage (de Gelder et al., 2006; Kret et al., 2011; Kret, Roelofs et al., 2013).

Expressions corporelles émotionnelles isolées

Lors de la plupart de nos interactions sociales, nous percevons les postures corporelles de notre interlocuteur. Celles-ci nous fournissent des informations sociales pertinentes pour comprendre les intentions d'autrui, mais également son ressenti émotionnel. La posture, les gestes ou encore les contractions musculaires véhiculés par le corps sont donc essentiels pour interagir efficacement avec les autres et adopter des comportements adaptés aux situations sociales. Par exemple, lorsqu'une personne retrouve un ami après une longue absence, elle peut spontanément ouvrir les bras, se pencher légèrement vers l'autre et adopter une posture ouverte. Ces gestes et cette posture transmettent de la proximité et communiquent ainsi des émotions de joie. Malgré la place déterminante des postures corporelles dans la perception émotionnelle, cet indice n'a pas toujours été au centre des recherches dans le domaine.

Cette section abordera les théories principales qui ont exploré la place des postures corporelles dans la perception émotionnelle ainsi que les mécanismes neuronaux qui sous-tendent la perception des indices corporels.

Fondements théoriques des postures corporelles comme vecteur émotionnel

Contrairement aux expressions faciales, l'importance des expressions corporelles dans la perception des émotions a longtemps été sous-estimée. Pourtant, dès l'époque de Darwin, les expressions corporelles étaient reconnues pour leur caractère universel et adaptatif, au même titre que les expressions faciales (Darwin, 1872). Dans ses travaux, Darwin amène la notion que les expressions corporelles soient, elles aussi, des réponses émotionnelles universelles et biologiquement ancrées permettant de faire face à des événements compromettant le bien-être, la survie ou la valeur reproductibles d'un individu (Darwin, 1872). Plus récemment, des théoriciens tels que Frijda (1988) et Tomkins (1995) ont proposé que les émotions sont intrinsèquement liées à l'action. Selon ces chercheurs, les expressions corporelles traduisent directement les préparations à l'action motivées par des émotions spécifiques (Frijda, 1988; Tomkins, 1995). Par exemple, une posture défensive pourrait refléter un état émotionnel lié à la peur, tandis qu'un geste de rejet pourrait indiquer du dégoût. Cependant, malgré ces contributions, les expressions corporelles ont longtemps été considérées comme secondaires par rapport aux expressions faciales, perçues simplement comme des éléments amplifiant l'émotion exprimée par le visage. Par exemple, une posture fermée (bras croisés, épaules rentrées) était interprétée comme renforçant une expression faciale de tristesse ou de peur, tandis qu'une posture

ouverte et expansive accentuait des émotions faciales telles que la joie ou la fierté (Ekman & Friesen, 1967).

Cette vision a été remise en question par de récents travaux démontrant que les postures corporelles véhiculent des informations spécifiques et distinctes sur les états émotionnels notamment à travers des mouvements et des gestes, indépendamment des expressions faciales (Atkinson et al., 2004; Coulson, 2004; Dael et al., 2012; de Gelder, 2006, 2009; de Gelder et al., 2015). Par ailleurs, les performances de reconnaissance des expressions faciales et corporelles semblent similaires (de Gelder et al., 2015). Notamment, lors de la création de la banque de stimuli d'expressions corporelles, the BEAST, les taux de reconnaissance des expressions corporelles atteignaient 85 % (de Gelder, 2009; de Gelder et al., 2015; de Gelder & van den Stock, 2011), suggérant une haute reconnaissance des émotions corporelles similairement à des banques de stimuli d'expressions faciales (Goeleven et al., 2008; Tottenham et al., 2009). Bien que les expressions faciales et corporelles partagent certaines similarités, elles véhiculent des informations distinctes et complémentaires dans la perception des émotions. Les expressions faciales jouent notamment un rôle central dans l'identification des individus (Calder & Young, 2005) alors que les expressions corporelles fournissent plutôt des indices sur les actions et les intentions (de Gelder, 2006; de Gelder et al., 2015). En outre, les expressions corporelles, contrairement aux expressions faciales, facilitent la reconnaissance des émotions exprimées lors de l'observation d'une interaction sociale émotionnelle entre plusieurs individus (Abramson et al., 2021). Ainsi, les expressions

corporelles participent activement à la reconnaissance des émotions d'autrui de la même manière que les expressions faciales, mais se différenciant par la nature des informations transmises.

Progressivement, la conception d'une émotion s'est transformée pour inclure la posture corporelle comme étant un véritable vecteur émotionnel. Notamment, certains modèles auparavant centrés exclusivement sur les expressions faciales ont inclus des caractéristiques posturales afin d'expliquer comment sont traitées les émotions perçues. C'est le cas du modèle dimensionnel de Fontaine et ses collègues (2007) qui se base sur six composantes émotionnelles (l'évaluation des événements, les sensations corporelles, les expressions motrices du visage, les expressions corporelles et vocales, les tendances à l'action, les expériences subjectives et la régulation des émotions) afin d'identifier quatre dimensions sur lesquelles se positionnent 24 émotions : l'activation (haute *vs.* élevée), la valence, la potentialité (puissance *vs.* faiblesse) et l'imprévisibilité (prévisible *vs.* imprévisible). La potentialité est également une dimension qui a été incluse dans le modèle de Gao et ses collègues (2010) concernant les expressions faciales émotionnelles, excepté que Fontaine et ses collègues (2007) associe cette dominance ou soumission interpersonnelle à des modifications du système nerveux parasympathique. Par exemple, les émotions associées à la dominance peuvent être liées à des yeux grands ouverts au niveau facial, une augmentation du rythme cardiaque au niveau corporel, et un besoin d'agir au niveau motivationnel. A l'inverse, la soumission pourrait se manifester par des yeux fermés, une sensation de faiblesse dans les bras ou jambes et une motivation réduite

à entreprendre des actions (Fontaine et al., 2007). Ce modèle positionne donc les émotions en fonction de ces quatre dimensions. Par exemple, le dégoût est caractérisé par une valence négative, une faible activation, une forte imprévisibilité et une sensation de puissance, tandis que la peur est associée à une valence négative, une forte activation, une forte imprévisibilité et un sentiment de faiblesse (Fontaine et al., 2007). Le modèle de Fontaine et ses collègues reprend donc l'hypothèse du modèle dimensionnel de Russell concernant les expressions faciales émotionnelles, en y ajoutant non seulement une composante corporelle, mais également vocale et motivationnelle afin d'avoir une vision plus complète de ce que représente une émotion.

Les postures corporelles, les gestes, ainsi que les sensations internes telles que les palpitations cardiaques ou les contractions musculaires, constituent des composantes essentielles de la perception émotionnelle. C'est sur cette base que s'appuie la théorie des émotions incarnées, qui propose une conception selon laquelle les émotions ne sont pas uniquement traitées de manière cognitive, mais également à travers des activations corporelles. Dans cette perspective, les sensations corporelles jouent un rôle fondamental, en fournissant des informations cruciales à la reconnaissance des émotions (Niedenthal, 2007; Winkielman et al., 2008). Par exemple, lors d'un échange avec une personne exprimant de la colère, nous pouvons adopter spontanément des mimiques faciales et corporelles, des contractions musculaires, une posture tendue, une accélération du rythme cardiaque, qui nous permettent d'identifier l'émotion ressentie par autrui. Cette approche théorique a été largement étudiée à travers le phénomène de rétroaction faciale, selon

lequel la reconnaissance des expressions faciales émotionnelles d'autrui repose sur notre capacité à imiter rapidement et inconsciemment ces expressions (Borgomaneri et al., 2020; Neal & Chartrand, 2011; Oberman et al., 2007). Toutefois, il a également été établi que ce mimétisme ne se limite pas à la perception d'expressions faciales. En effet, des réactions mimétiques similaires peuvent être observées lors de la perception d'émotions transmises par d'autres indices sensoriels, tels que les vocalisations (Hawk et al., 2012) ou les postures corporelles (Magnée et al., 2007). Ces résultats suggèrent que le mimétisme émotionnel découle d'un processus dynamique de simulation sensori-motrice, et ne constitue pas une simple reproduction des expressions faciales perçues (Moody et al., 2017). De manière similaire, certaines études montrent que la perception d'expressions faciales émotionnelles entraîne des réponses motrices faciales, mais également corporelles. Ainsi Moody et ses collègues (2017) ont observé qu'en réponse à des visages exprimant la colère, les participants activaient non seulement les muscles du visage associés à cette émotion, mais aussi les muscles des bras, avec une activité compatible à des mouvements de fermeture des poings, typiques de l'expression corporelle de la colère. De même, en réponse à des expressions faciales de peur, une activation des muscles faciaux et des bras a été enregistrée, traduisant une tendance à lever les mains, comportement souvent associé à la peur (Moody et al., 2017). Enfin, l'importance des composantes corporelles dans la reconnaissance émotionnelle est renforcée par les travaux de Lenzoni et ses collègues (2020), qui ont montré que les patients atteints de dystrophie myotonique, une pathologie affectant la contraction des muscles faciaux et corporels, présente des difficultés à reconnaître les expressions émotionnelles faciales et corporelles.

Ces résultats soulignent le rôle central du corps dans le traitement émotionnel et indiquent que le mimétisme émotionnel dépasse largement le cadre des seules expressions faciales.

Corrélat neuronal de la perception des expressions corporelles isolées

Contrairement aux expressions faciales, peu d'études ont exploré les corrélats neuronaux impliqués dans le traitement des expressions corporelles et particulièrement son déroulement temporel. De manière similaire aux émotions négatives exprimées par le visage, la composante P100 présente une amplitude plus importante lors de perception d'expressions corporelles de peur plutôt que neutres, ce qui suggère une propension précoce de l'activité cérébrale à traiter en priorité les signaux négatifs plutôt que neutres (Zhang et al., 2015). Cette activité, ne pouvant pas être expliquée par une sensibilité de la composante P100 aux caractéristiques de bas niveau, suggère un traitement rapide et automatique des expressions corporelles de peur (van Heijnsbergen et al., 2007).

Le traitement des expressions corporelles se poursuit par l'activation de la N170. Alors que la composante N170 a fortement été associée à la perception des expressions faciales (Almeida et al., 2016; Batty & Taylor, 2003; Campanella et al., 2002; Deffke et al., 2007; Gajewski & Stoerig, 2011; Hinojosa et al., 2015; Jacques & Caharel, 2022; Watanabe et al., 2003), il semblerait qu'elle soit également sensible à la perception des expressions corporelles (Stekelenburg & de Gelder, 2004; Thierry et al., 2006). En effet, l'effet d'inversion des visages, précédemment évoqué (Watanabe et al., 2003), est également observé dans la perception des expressions corporelles, c'est-à-dire que la

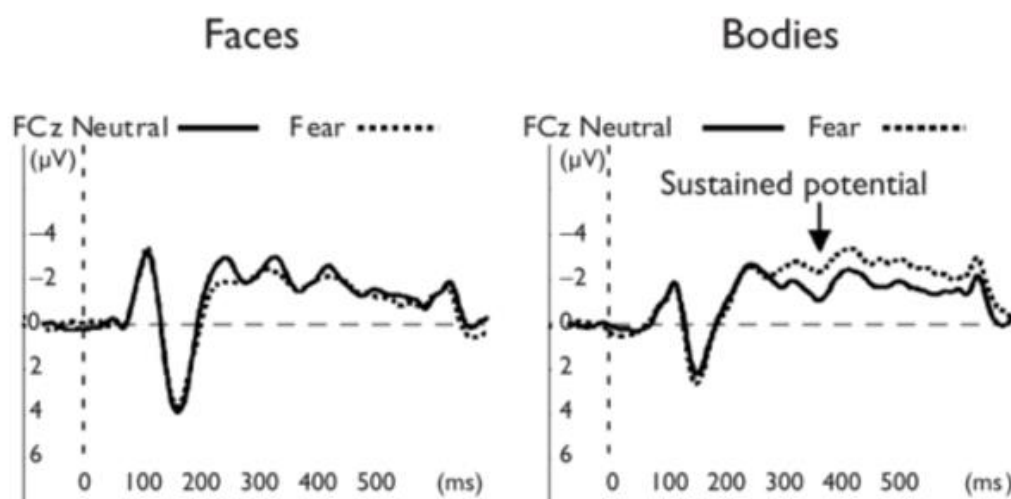
composante N170 présente une amplitude plus élevée et tardive lorsque les expressions corporelles sont présentées à l'envers plutôt qu'à l'endroit (Stekelenburg & de Gelder, 2004). Cela suggère une similarité dans l'encodage précoce des expressions faciales et corporelles, indiquant que le N170 n'est pas exclusivement dédié au traitement des expressions faciales, mais qu'elle joue également un rôle dans la perception des expressions corporelles. Plus récemment, une étude a montré que l'amplitude de la N170 était plus importante lors de perception d'expressions corporelles plutôt que d'objets, renforçant l'hypothèse du rôle important de la N170 dans le traitement de la configuration des expressions corporelles (Hietanen & Nummenmaa, 2011). Ces similarités dans l'implication de la N170 dans le traitement configurationnel des expressions faciales et corporelles peut s'expliquer par un processus se déroulant au niveau du cortex fusiforme dont l'activité de la N170 semble provenir. Ce dernier fait état d'un chevauchement entre le traitement des expressions faciales et corporelles, suggérant une base neuronale commune pour l'analyse des indices émotionnels véhiculés par le visage et le corps (de Gelder, 2006; de Gelder et al., 2015; Hadjikhani & de Gelder, 2003). Bien qu'il semble y avoir un traitement similaire entre les expressions faciales et corporelles particulièrement au niveau structurel, en ce qui concerne les indices émotionnels véhiculés par le visage et le corps, des différences apparaissent. La N170 ne semble pas intervenir dans le traitement émotionnel des expressions corporelles, contrairement à son rôle dans celui des expressions faciales émotionnelles (Hinojosa et al., 2015). En revanche, la composante VPP, localisée dans la région frontale et apparaissant dans la même fenêtre temporelle que la N170, présente une amplitude plus élevée face à des expressions corporelles de peur

qu'à des expressions neutres (Stekelenburg & de Gelder, 2004; van Heijnsbergen et al., 2007).

Plus tardivement autour de 300ms après la présentation du stimulus, la perception des expressions corporelles de peur induit un potentiel fronto-central soutenu, reflétant une activité cérébrale prolongée. Cet effet n'est pas observé lors de la perception des expressions faciales, ce qui pourrait refléter une différence de traitement entre ces deux canaux émotionnels, avec notamment une attention plus soutenue portée aux expressions corporelles qu'aux expressions faciales. Cette ambiguïté nécessiterait un traitement cognitif plus approfondi afin de reconnaître les émotions exprimées par le corps (Stekelenburg & de Gelder, 2004) (voir Figure 7).

Figure 7

Grand moyennage des PEs sur l'électrode FCz lorsque le focus attentionnel est dirigé vers l'expression faciale (gauche) et corporelle (droite)



Note. Le grand moyennage des PE montrent un potentiel évoqué soutenu à partir de 300ms reflétant une attention soutenue sur les expressions corporelles émotionnelles. Cette image est reprise de l'article de Stekelenburg et de Gelder (2004).

Pour conclure, bien que les expressions corporelles émotionnelles aient longtemps été négligées dans la recherche sur la perception des émotions, elles suscitent un intérêt croissant au cours des dernières années, même si moindre que celui accordé aux expressions faciales. Malgré cet intérêt grandissant, peu d'études ont encore mis en place des modèles théoriques récents sur la perception des expressions corporelles émotionnelles et les corrélats neuronaux impliqués dans le traitement de ces expressions. Pourtant, les émotions véhiculées par le corps constituent un vecteur essentiel de communication affective, fournissant des indices cruciaux sur les actions et les mouvements d'autrui, facilitant ainsi l'inférence de ses intentions (de Gelder, 2006; de Gelder et al., 2015). Sur le plan des corrélats cérébraux, plusieurs études indiquent que le

traitement des expressions corporelles émotionnelles partage des mécanismes précoces avec celui des expressions faciales, comme en témoigne l'activation des composantes P100 (Zhang et al., 2015) et N170 (Hietanen & Nummenmaa, 2011). Toutefois, au-delà de cette similarité, le traitement de l'information émotionnelle issue des expressions corporelles implique des activations cérébrales distinctes, notamment par l'activité de la composante VPP ainsi que d'un potentiel fronto-central soutenu tardivement. Ce dernier marque une allocation attentionnelle prolongée aux expressions corporelles émotionnelles, suggérant qu'elles nécessitent un traitement plus approfondi pour être pleinement reconnu (Stekelenburg & de Gelder, 2004; van Heijnsbergen et al., 2007).

Bien que ces recherches portant sur les expressions faciales ou corporelles aient permis de mieux comprendre comment ces indices émotionnels sont perçus et traités, elles négligent une dimension essentielle de notre expérience quotidienne. En effet, dans la vie de tous les jours, les émotions sont le plus souvent perçues à travers plusieurs canaux sensoriels en parallèle. Les expressions faciales et corporelles sont généralement observées simultanément, véhiculant des informations émotionnelles de manière conjointe. Ainsi, les études qui considèrent ces indices de façon séparée présentent une limite importante, car elles ne reflètent pas les conditions réelles de perception des émotions, et restreignent par conséquent notre compréhension des processus émotionnels impliqués dans une perception conjointe des indices faciaux et corporels.

Influence bidirectionnelle entre les expressions faciales et corporelles

Afin de mieux caractériser la perception des émotions au plus près d'une perception réelle, il apparaît essentiel d'examiner le traitement émotionnel en contexte écologique. Dans la vie quotidienne, nous percevons des expressions faciales accompagnées d'indices émotionnels contextuels tels que la posture corporelle, la scène environnante ou les sons. En effet, les émotions sont par nature multisensorielles, c'est-à-dire qu'elles sont perçues à travers plusieurs modalités sensorielles simultanément comme la vision (expressions faciales et corporelles), l'audition (voix, sons) ou l'olfaction (odeurs). La reconnaissance émotionnelle résulte donc de l'intégration simultanée d'indices provenant de l'ensemble de nos sens. Cette intégration facilite la reconnaissance lorsque les indices véhiculent la même émotion et l'entrave lorsqu'ils sont discordants (Abo Foul et al., 2018; Balconi & Carrera, 2011; Collignon et al., 2008; Falagiarda & Collignon, 2019; Pourtois et al., 2002; Seubert et al., 2010; Song et al., 2022; Van den Stock & de Gelder, 2014; Van den Stock et al., 2007). Par exemple, dans un film d'horreur, une musique sinistre facilite la reconnaissance et l'intensité de la peur perçue, alors qu'une musique joyeuse réduit cette reconnaissance. De même, si des acteurs sourient et semblent joyeux, l'émotion devient difficile à interpréter, poussant l'observateur à chercher des indices dans d'autres modalités sensorielles pour mieux identifier l'émotion (Collignon et al., 2008).

Ainsi, la reconnaissance des émotions d'autrui repose sur une intégration multisensorielle des indices émotionnels. Cette intégration se manifeste également au sein d'une même modalité sensorielle, par exemple entre expressions faciales et corporelles

(Lecker et al., 2020; Meeren et al., 2005; Willis et al., 2011). Explorer l'influence de ces deux types d'indices est particulièrement pertinent, car dans la vie quotidienne, nous percevons fréquemment le visage et le corps d'autrui simultanément.

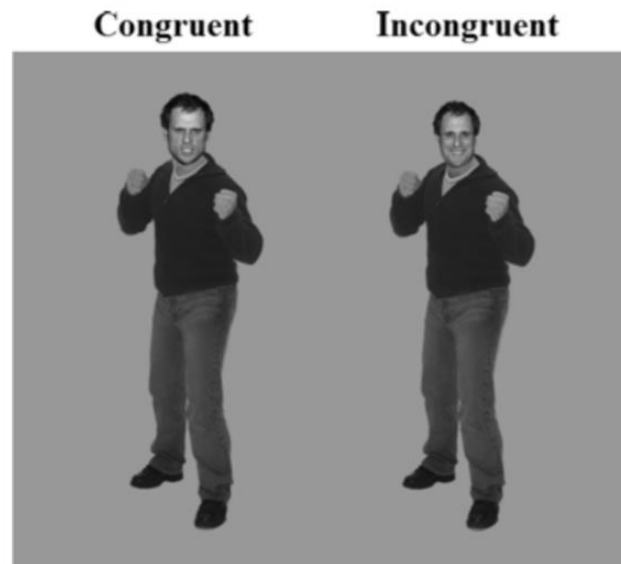
Dans cette section, nous nous focaliserons sur une seule modalité sensorielle, c'est-à-dire la vision, à travers les recherches actuelles sur les influences bidirectionnelles entre les expressions faciales et corporelles dans la reconnaissance émotionnelle, ainsi que sur le déroulement temporel de leur intégration. .

Effet contextuel entre les expressions faciales et corporelles

Afin d'étudier la perception simultanée des expressions faciales et corporelles ainsi que d'examiner les interactions entre ces canaux émotionnels, les recherches ont développé un protocole expérimental spécifique utilisant des stimuli émotionnels congruents et incongruents (Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013). Les participants sont ainsi invités à réaliser des tâches de reconnaissance émotionnelle sur des images de personnes exprimant des émotions similaires (congruence) ou différences (incongruence) avec leur visage et leur corps (voir Figure 8).

Figure 8

Exemple de stimuli congruent et incongruent utilisés dans la littérature



Note. L'image à gauche représente une personne exprimant de la colère avec son visage et son corps (congruence) et l'image de droite représente une personne exprimant de la joie avec son visage et de la colère avec son corps (incongruence). L'image provient de l'article de Kret et de Gelder (2013).

Une critique émise concernant ce type de protocole est qu'il est peu probable, dans la vie quotidienne, d'observer une personne exprimant simultanément deux émotions clairement distinctes entre le visage et le corps. Cependant, l'objectif de ce paradigme n'est pas de reproduire une situation naturelle au sens strict, mais de mettre en place un conflit émotionnel contrôlé qui permette d'étudier les mécanismes d'intégration. Ce type d'ambiguïté est en réalité très fréquent dans la vie réelle. En effet, les expressions émotionnelles sont souvent peu claires ou encore faiblement intenses, générant parfois des signaux non verbaux plus difficiles à reconnaître. Par exemple, nous pouvons sourire par politesse tout en adoptant une posture tendue, ou encore montrer une expression neutre

alors que le corps trahit la peur. Ces situations sont importantes, car elles poussent l'observateur à chercher des informations supplémentaires dans le contexte pour établir une reconnaissance émotionnelle efficace. Le paradigme congruent/incongruent permet donc de modéliser ce besoin d'intégration, tout en conservant un haut niveau de contrôle expérimental.

Les recherches utilisant ce protocole expérimental montrent que la perception simultanée des expressions faciales et corporelles influence la reconnaissance émotionnelle (Kret, Stekelenburg et al., 2013; Lecker et al., 2020). En effet, cette reconnaissance est d'autant plus rapide et précise lorsque les expressions faciales et corporelles véhiculent une émotion identique, c'est-à-dire dans des situations de congruence émotionnelle. A l'inverse, lorsque les expressions faciales et corporelles transmettent des émotions différentes (incongruence) ou qu'elles sont perçues isolément du contexte, la reconnaissance émotionnelle devient plus difficile et moins précise (Abo Foul et al., 2018; Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013; Willis et al., 2011). Cet effet de facilitation contextuelle, nommé dans la littérature l'effet contextuel, indique que les informations provenant des expressions faciales et corporelles sont intégrées afin d'obtenir une perception précise des émotions (Aviezer, Hassin, Ryan et al., 2008; Aviezer et al., 2012a; Kret, Roelofs et al., 2013; Kret, Stekelenburg et al., 2013; Meeren et al., 2005). Cette intégration se déroule dans une perspective d'unité individuelle (*single person unit*) (Aviezer et al., 2012a), c'est-à-dire que contrairement aux conceptions traditionnelles postulant un traitement séparé des expressions faciales et corporelles avant une pondération finale (Meeren et al., 2005; Wallbott, 1998), la perception émotionnelle

semble adopter une approche holistique et intégrée, dans laquelle les deux canaux sont traités ensemble comme une seule unité (Aviezer et al., 2012a; de Gelder, 2009). En effet, lors de trois études, Aviezer et ses collègues (2012b) ont demandé à des participants de réaliser une tâche de reconnaissance d'expressions faciales lorsque les visages étaient alignés ou décalés par rapport aux corps congruents ou incongruents (voir Figure 9).

Figure 9

Image représentant des stimuli alignés (à gauche) et désalignés (à droite)



Note. L'image provient de l'article de Aviezer et ses collègues (2012b).

Les résultats montrent que le désalignement de l'expression faciale avec le corps diminue l'influence de l'expression corporelle sur la reconnaissance de l'émotion faciale, réduisant ainsi l'effet facilitateur des expressions corporelles congruentes et l'effet perturbateur des expressions corporelles incongruentes. Ces observations confirment la notion de perception globale et holistique des expressions faciales et corporelles (Aviezer et al., 2012a). En résumé, cela met en lumière l'importance de la contribution simultanée des expressions faciales et corporelles dans la perception des émotions et souligne l'intégration globale des informations émotionnelles.

Nous avons montré dans une revue systématique de la littérature que l'effet contextuel entre les expressions faciales et corporelles repose sur cinq éléments fondamentaux (Puffet et al., 2025). Premièrement, *cet effet n'est pas homogène* entre les différentes paires d'expressions faciales et corporelles, certaines combinaisons d'expressions entraînant un taux d'erreurs plus élevé dans la catégorisation des émotions faciales que d'autres (Karaaslan et al., 2020; Mondloch et al., 2013; Reschke & Walle, 2021). Par exemple, une étude menée par Lecker et ses collègues (2020) a révélé que les expressions faciales de dégoût étaient mieux reconnues lorsqu'elles étaient associées à des expressions corporelles de peur avec un taux de reconnaissance de 66 % que lorsqu'elles étaient combinées avec des expressions corporelles de colère avec un taux de reconnaissance de 11 %, ce qui suggère que l'effet contextuel varie en fonction des paires d'émotions provenant du visage et du corps.

Deuxièmement, l'influence des signaux émotionnels dans les situations d'incongruence est *bidirectionnelle* (Lecker et al., 2020; Reschke & Walle, 2021; Willis et al., 2011; Zhang et al., 2019). La majorité des recherches sur l'effet contextuel se sont concentrées sur l'impact des expressions corporelles sur la reconnaissance des expressions faciales (Abo Foul et al., 2018; Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013; Willis et al., 2011), toutefois les expressions faciales influencent également la perception des émotions véhiculées par le corps, mettant ainsi en évidence une interaction réciproque entre ces deux canaux (Lecker et al., 2020). Par exemple, la reconnaissance

d'expressions corporelles de dégoût est rendue plus difficile lorsque celles-ci sont accompagnées d'expressions faciales de colère plutôt que de dégoût (Lecker et al., 2020). Cela montre que les expressions faciales fournissent un contexte émotionnel influençant la reconnaissance des émotions corporelles, et réciproquement les expressions corporelles modulent également la reconnaissance des émotions faciales (Abo Foul et al., 2018; Kret, Stekelenburg et al., 2013; Lecker et al., 2020).

Troisièmement, cette bidirectionnalité dans l'effet contextuel présente une *asymétrie* (Karaaslan et al., 2020; Mondloch et al., 2013). Cette asymétrie se manifeste par le fait que certaines combinaisons spécifiques d'émotions faciales et corporelles influencent différemment la reconnaissance des expressions faciales. Par exemple, lorsque le visage exprime de la peur et que le corps exprime de la tristesse, la reconnaissance de la peur est altérée avec un taux de reconnaissance de 73 %, alors qu'à l'inverse lorsque le visage exprime de la tristesse et le corps de la peur, la reconnaissance de la tristesse est moins affectée avec un taux de reconnaissance de 91 % (Mondloch et al., 2013). Ce type d'observation indique que l'influence contextuelle du corps sur la reconnaissance de l'émotion faciale varie en fonction de la nature des émotions impliquées.

Quatrièmement, l'effet contextuel est *modulé par le focus attentionnel* (Abo Foul et al., 2018; Kret, Stekelenburg et al., 2013; Lecker et al., 2020), c'est-à-dire que la force de l'influence exercée par chaque canal, visage ou corps, sur la reconnaissance émotionnelle dépend de l'orientation de l'attention sur l'un de ces deux canaux. En effet, les expressions

corporelles exercent une influence plus marquée sur la reconnaissance des expressions faciales qu'inversement, cette prédominance pouvant s'expliquer par plusieurs facteurs. D'une part, les expressions corporelles lorsqu'elles sont présentées isolément sont généralement moins bien reconnues que les expressions faciales (Kret, Stekelenburg et al., 2013; Lecker et al., 2020). D'autre part, dans les interactions sociales quotidiennes, les individus ont tendance à focaliser leur attention sur les visages et à traiter les signaux corporels comme un contexte émotionnel secondaire. Cette focalisation sur le visage, combinée à une attention moindre portée aux signaux corporels, pourrait expliquer pourquoi les expressions faciales sont plus sensibles à l'influence des expressions corporelles (Lecker et al., 2020). Une autre explication possible est que les expressions corporelles émotionnelles sont moins ambiguës que les émotions faciales engendrant donc moins de besoins d'informations secondaires pour reconnaître les expressions corporelles (Karaaslan et al., 2020; Kret, Stekelenburg et al., 2013; Lecker et al., 2020; Meeren et al., 2005).

Enfin, l'effet contextuel entre les expressions faciales et corporelles semble être modulé par des *différences interindividuelles*. En effet, les individus ne présentent pas une sensibilité égale au contexte émotionnel lors de la perception d'émotions. Certaines personnes ont tendance à s'appuyer davantage sur le contexte, notamment les expressions corporelles, tandis que d'autres privilégient une focalisation sur le visage (Ensenberg-Diamant et al., 2025). Ces différences dans les stratégies d'observation entraînent des variations dans l'influence exercée par les expressions corporelles sur la reconnaissance des émotions faciales. Notamment, les individus sensibles au contexte adoptent un traitement plus holistique des stimuli émotionnels, ce qui les rend plus réceptifs à l'effet contextuel. Par ailleurs, ces différences interindividuelles persistent même lorsque l'attention est explicitement dirigée vers les expressions faciales ou corporelles, suggérant que la tendance à privilégier le visage ou le corps ne semble pas modulable par une manipulation attentionnelle (Ensenberg-Diamant et al., 2025).

En résumé, les recherches comportementales sur la perception simultanée des indices faciaux et corporels soulignent l'importance de ces deux canaux dans la reconnaissance des émotions (Karaaslan et al., 2020; Kret, Stekelenburg et al., 2013; Lecker et al., 2020; Meeren et al., 2005). Ces travaux montrent également que les indices faciaux et corporels s'influencent de manière bidirectionnelle et asymétrique (Lecker et al., 2020) et que des différences interindividuelles peuvent moduler l'effet contextuel (Ensenberg-Diamant et al., 2025). Toutefois, si ces résultats éclairent les mécanismes perceptifs de l'intégration des indices faciaux et corporels, ils ne permettent pas d'en saisir les dynamiques cérébrales

précises. C'est pourquoi plusieurs études neuroscientifiques ont exploré la perception simultanée des expressions faciales et corporelles en utilisant des mesures EEG afin de mieux comprendre la temporalité de l'activité cérébrale dans l'intégration de ces indices.

Décours temporel de l'intégration des expressions faciales et corporelles

Afin de caractériser le décours temporel de l'intégration des expressions faciales et corporelles, plusieurs études ont utilisé l'analyse des composantes de potentiel évoqué (PE).

Trois étapes de traitement ont été identifiées caractérisant un traitement précoce autour de 100ms (p. ex., la composante P100), intermédiaire autour de 200ms (p. ex., la composante P200) et tardif autour de 300ms (p. ex., la composante P300 ou LPP) après la présentation des stimuli faciaux et corporels émotionnels (Chen et al., 2022; Gu et al., 2013; Hu et al., 2020; Li, 2021; Teufel et al., 2019).

La première étape du traitement des expressions faciales et corporelles reflète un traitement automatique et rapide des expressions corporelles menaçantes, mettant en lumière un biais de négativité précoce dans le traitement émotionnel. Ce biais est notamment mis en évidence par une augmentation de l'amplitude de la composante P100 lors de la perception des expressions corporelles exprimant de la peur, en comparaison à des expressions corporelles de joie, et cela, indépendamment que l'attention soit portée sur les expressions faciales ou corporelles (Chen et al., 2022; Gu et al., 2013; Wang et

al., 2017). Cette observation suggère un traitement rapide des signaux émotionnels négatifs provenant des expressions corporelles en présence des expressions faciales. Par ailleurs, la détection de l'incongruence entre les expressions faciales et corporelles semble également se produire à ce stade précoce du traitement, comme en témoigne une augmentation de l'amplitude de la composante P100 lors de la perception d'expressions émotionnelles incongruentes par rapport aux expressions congruentes (Meeren et al., 2005; Wang et al., 2017), suggérant une détection rapide et automatique des conflits émotionnels. Concernant la composante N170, celle-ci ne semble pas être modulée par la congruence ou l'incongruence entre les expressions faciales et corporelles (Hu et al., 2020; Meeren et al., 2005). Cependant, comme montrée précédemment, cette composante est impliquée dans le traitement configurationnel des expressions faciales et corporelles lorsqu'elles sont présentées isolément (Hu et al., 2020; Stekelenburg & de Gelder, 2004; Watanabe et al., 2003). Il est donc plausible que la N170 joue également un rôle dans le traitement configurationnel de ces expressions perçues conjointement, bien qu'aucune étude ne se soit encore spécifiquement penchée sur cette question.

La *seconde étape du traitement* est caractérisée par la détection de l'incongruence entre les expressions faciales et corporelles (Chen et al., 2022; Gu et al., 2013; Li, 2021). Cette détection se manifeste notamment par une augmentation de l'amplitude des composantes N200 et P200, situées respectivement dans les régions fronto-centrales et occipito-temporales, lors de la perception d'expressions faciales et corporelles incongruentes par rapport aux expressions congruentes (Chen et al., 2022; Gu et al., 2013;

Hu et al., 2020; Li, 2021). Toutefois, comme précisées précédemment, certaines études ont observé cette détection plus précocement dès 100ms, via une modulation de la P100 (Meeren et al., 2005; Wang et al., 2017). Ces divergences pourraient être attribuables à la nature des émotions utilisées, par exemple Meeren et ses collègues (2005) ont restreint leurs stimuli aux expressions de colère et de dégoût tandis que Li (2021) a inclus plus d'émotions comprenant la colère, le dégoût, la peur, la tristesse et la surprise. De plus, des variables dans les protocoles expérimentaux, telles que la durée de présentation des stimuli et les banques de stimuli utilisées, pourraient également expliquer ces disparités dans les résultats. Malgré ces différences, l'ensemble des résultats convergent vers l'hypothèse d'un traitement rapide et automatique de l'intégration des expressions faciales et corporelles (Chen et al., 2022; Gu et al., 2013; Hu et al., 2020; Li, 2021; Meeren et al., 2005; Wang et al., 2017). Contrairement aux composantes précoces comme la P100 et N170, la P200 semble également intervenir dans l'intégration des expressions faciales et corporelles via sa sensibilité aux configurations plausibles et impossibles des expressions faciales et corporelles (Vakli et al., 2017). Une étude a notamment montré que la P200 présente une amplitude plus élevée lorsqu'un visage est accompagné de mains inversées à 180 degrés par rapport à une configuration corporelle plausible (Vakli et al., 2017). Ces résultats suggèrent que la P200 est impliquée non seulement dans la détection des conflits émotionnels, mais également dans l'analyse de la cohérence structurelle des expressions corporelles par rapport aux expressions faciales (Hu et al., 2020; Vakli et al., 2017).

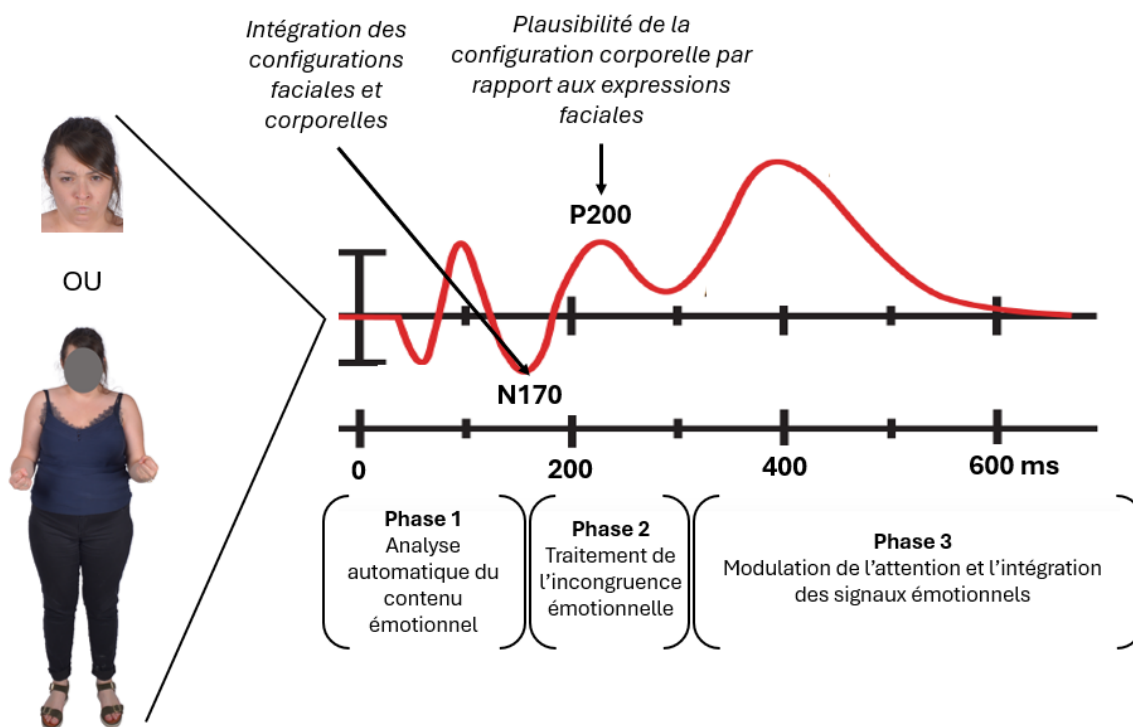
Enfin, la *dernière étape du traitement* correspond à une évaluation attentionnelle et à une analyse plus approfondie des émotions perçues. Lorsque l'attention est dirigée vers les expressions faciales, l'amplitude de la P300 est plus élevée en cas de congruence par rapport à l'incongruence émotionnelle (Gu et al., 2013; Li, 2021; Wang et al., 2017). En revanche, lorsque l'attention est focalisée sur les expressions corporelles, l'amplitude de la P300 est plus marquée pour les expressions de peur que pour celles de joie, sans effet significatif de la congruence émotionnelle (Gu et al., 2013). Ces résultats indiquent que l'attention module, à un stade tardif du traitement, la manière dont les informations émotionnelles sont intégrées lorsque les expressions faciales et corporelles sont perçues conjointement (Gu et al., 2013; Lecker et al., 2020).

Ainsi, le traitement des expressions faciales et corporelles semble s'organiser en trois étapes. La première, autour de 100ms, correspond à une analyse rapide et automatique des signaux émotionnels, en particulier ceux provenant des expressions corporelles négatives (Chen et al., 2022; Gu et al., 2013; Wang et al., 2017). La N170, quant à elle, serait impliquée dans l'intégration des configurations faciales et corporelles, notamment en raison de sa sensibilité à l'inversion de ces expressions (Hu et al., 2020; Stekelenburg & de Gelder, 2004; Watanabe et al., 2003). La seconde étape est caractérisée par le traitement de l'incongruence entre les expressions faciales et corporelles qui intervient précocement, dès 100 ou 200ms après la présentation du stimulus (Chen et al., 2022; Gu et al., 2013; Hu et al., 2020; Li, 2021; Meeren et al., 2005; Wang et al., 2017). La P200 apparaît notamment sensible à la plausibilité de la configuration corporelle (particulièrement les

mains) par rapport aux expressions faciales (Vakli et al., 2017). Enfin, vers 300ms, la dernière phase débute avec une modulation de l'attention et l'intégration des signaux émotionnels issus des expressions faciales et corporelles (voir Figure 10) (Chen et al., 2022; Gu et al., 2013; Hu et al., 2020; Wang et al., 2017).

Figure 10

Schéma récapitulatif du traitement des expressions faciales ou corporelles



Bien qu'un grand nombre de recherches ont été effectuées ces dernières années sur l'intégration des expressions faciales et corporelles, certains aspects restent encore à explorer. Notamment, il est essentiel de déterminer si cette intégration se déroule de manière automatique ou si elle nécessite une intervention volontaire des individus. Cette

question est centrale, car elle permettrait de mieux comprendre les mécanismes sous-jacents au traitement des émotions et leur influence par exemple sur les interactions sociales. Si cette intégration est automatique, cela suggérerait qu'elle repose sur des processus perceptifs rapides et efficaces et inconscients, facilitant une réponse émotionnelle immédiate et adaptée aux signaux de l'environnement. En revanche, si elle nécessite une intervention volontaire, cela impliquerait l'implication de processus cognitifs plus contrôlés. Ainsi, la question de l'automatisme de cette intégration constitue un enjeu majeur restant à approfondir dans la littérature afin d'améliorer notre compréhension des processus émotionnels.

Automatisme dans le traitement émotionnel

Déterminer le caractère automatique d'un processus cognitif est une question récurrente en psychologie. Il a notamment été investigué dans divers domaines tels que la mémoire (Abbenhuis et al., 1990) ou encore les émotions (Tracy & Robins, 2008). Historiquement, les processus cognitifs étaient considérés comme entièrement automatiques ou, au contraire, entièrement volontaires. Toutefois, cette vision dichotomique a progressivement évolué afin d'admettre aujourd'hui qu'un processus cognitif peut être automatique sans pour autant correspondre à l'ensemble des critères d'automatisme (Tzelgov, 1997). Dans ce cadre, trois critères indépendants ont été proposés pour définir l'automatisme d'un processus cognitif : la conscience, l'efficacité et la rapidité (Moors & De Houwer, 2006).

Dans cette section, nous explorerons d'abord le concept d'automatisme d'un processus cognitif à travers les trois critères qui le définissent. Ensuite, nous examinerons la littérature actuelle sur le processus automatique de l'effet contextuel entre les expressions faciales et corporelles selon les trois critères d'automatisme.

Automatisme d'un processus cognitif

Tout comme certains comportements quotidiens tels que faire du vélo, manger, respirer, conduire ou nager, des processus cognitifs comme la perception des émotions, peuvent aussi se dérouler de manière automatique. Savoir si un processus cognitif est automatique ou volontaire est crucial pour comprendre comment l'information est traitée par le cerveau et pour déterminer le niveau de ressources cognitives, d'attention et de conscience requis pour accomplir une tâche donnée. Il existe plusieurs propositions (voir plus loin) mais dans le cadre de cette thèse, nous considérerons qu'un processus cognitif est considéré comme automatique lorsqu'il se caractérise par trois critères principaux : sa rapidité, son caractère conscient et son efficacité (Moors & De Houwer, 2006). Le critère de rapidité, tel que présenté par Moors et De Houwer (2006), reste assez vague, les auteurs précisent uniquement qu'il correspond à un processus rapide. L'efficacité désigne, quant à lui, le fait qu'il nécessite peu ou pas de ressources cognitives ou attentionnelles, permettant ainsi son déroulement en parallèle à d'autres processus. Enfin, l'automatisme comprend également un aspect inconscient, signifiant que le processus peut se réaliser sans que l'individu en ait nécessairement conscience (Moors & De Houwer, 2006). Ces trois critères indépendants permettent de distinguer les processus automatiques, qui sont

rapides, efficaces et inconscients, des processus contrôlés, qui, eux, sont plus lents, demandent une plus grande mobilisation des ressources et sont conscients.

Afin d'évaluer les trois critères d'automatisme, différentes méthodologies ont été proposées dans la littérature. Le critère d'*efficacité* est généralement investigué à l'aide de paradigmes de double tâche, dans lequel les participants sont amenés à réaliser simultanément une tâche principale et une tâche secondaire mobilisant fortement les ressources cognitives (p. ex., une tâche de mémorisation de chiffres). L'idée sous-jacente est que si la performance dans la tâche principale demeure stable malgré l'ajout de la tâche secondaire, cela suggère que le processus cognitif est efficace, c'est-à-dire qu'il nécessite peu de ressources cognitives pour fonctionner (Moors & De Houwer, 2006). Ce paradigme a notamment été appliqué dans des études sur la reconnaissance d'expressions émotionnelles, où les participants devaient identifier des émotions faciales tout en mémorisant des séquences de chiffres, tâche secondaire reconnue pour sa forte charge cognitive (Aviezer et al., 2011; Cao et al., 2022; Lima et al., 2019; Moors & De Houwer, 2006; Tracy & Robins, 2008). Le maintien de la performance en reconnaissance émotionnelle malgré cette surcharge cognitive permet alors de déduire que ce processus se déroule de manière automatique au sens de l'efficacité.

Le second critère, celui de la *rapidité*, repose sur l'idée qu'un processus automatique s'initie rapidement après la présentation d'un stimulus. Toutefois, il convient de distinguer la durée du traitement cognitif de celle de la présentation du stimulus. En effet, un

processus peut se poursuivre même après la disparition d'un stimulus (Moors & De Houwer, 2006). En pratique, la rapidité est évaluée à travers la mesure du temps minimal requis pour traiter un stimulus donné, comme une expression émotionnelle faciale par exemple (Tracy & Robins, 2008). Il reste néanmoins difficile d'établir un seuil temporel précis permettant de qualifier un traitement de rapide, dans la mesure où cette durée dépend du type de tâche, de la nature des stimuli et du contexte expérimental (Moors & De Houwer, 2006).

Finalement, le troisième critère d'automatisme est l'*inconscience*. Un processus est considéré comme automatique s'il peut se déclencher et se dérouler sans accès conscient, que ce soit au moment de la perception du stimulus, lors des étapes de traitement, ou encore au niveau de la réponse comportementale subséquente (Moors & De Houwer, 2006). Pour explorer ce critère, plusieurs manipulations expérimentales sont possibles. Par exemple, certaines études ont recours à des durées très brèves de présentation des stimuli, combinées à des techniques de masquage visuel, afin d'empêcher leur traitement conscient (Smith, 2012). D'autres recherches manipulent la position spatiale du stimulus dans le champ visuel (p. ex., en le présentant en périphérie) afin de limiter son accès à la conscience (Mumenthaler & Sander, 2015). Ces différentes méthodes permettent ainsi d'évaluer dans quelle mesure un traitement cognitif donné peut se dérouler de manière inconsciente.

En conclusion, l'évaluation de l'automatisme d'un processus cognitif repose sur plusieurs critères indépendants. Selon Moors et De Houwer (2006), trois critères permettent de qualifier un processus comme automatique. Il s'agit d'abord de son efficacité, c'est-à-dire la faible mobilisation des ressources cognitives et attentionnelles. Ensuite, sa rapidité, c'est-à-dire le court laps de temps nécessaire à son traitement. Enfin, son caractère non conscient, c'est-à-dire le fait qu'il puisse se produire sans que l'individu en ait conscience (Moors & De Houwer, 2006).

D'autres modèles ont proposé divers critères pour définir l'automatisme d'un processus cognitif. Bargh, notamment, identifie quatre caractéristiques : le traitement inconscient, l'absence de contrôle volontaire (le processus ne peut être arrêté une fois initié), l'absence d'effort, et le caractère non intentionnel (Bargh, 2014; Bargh & Tota, 1988; Bargh et al., 1996). Contrairement à Moors et De Houwer (2006), cette approche met l'accent sur le fait qu'un processus peut se déclencher sans intention préalable, et semble irrésistible. De manière générale, ces modèles classiques reposent sur une approche multidimensionnelle et flexible de l'automatisme durant lequel un processus peut être plus ou moins automatique en fonction de la présence ou de l'importance relative de différents critères. Cette approche a été également soutenue par d'autres auteurs, comme Tzelgov (1997), qui soutiennent qu'il n'est pas nécessaire que tous les critères soient simultanément présents. Selon cette perspective, l'automatisme serait mieux comprise comme un continuum, dans lequel un processus cognitif peut être considéré comme plus ou moins automatique selon qu'il remplit une combinaison particulière de ces critères.

Cette conception plus nuancée permet de rendre compte de la diversité des processus automatiques observés (Tzelgov, 1997).

Dans le cadre du traitement des informations affectives, ces critères ont également été explorés pour identifier dans quelle mesure la perception d'indices émotionnels se produit de manière automatique ou volontaire.

Automaticité d'un processus émotionnel : les expressions émotionnelles isolées

La question de l'automaticité du traitement des informations émotionnelles a une grande importance tant en recherche fondamentale et clinique que dans la compréhension de notre comportement au quotidien. En effet, un traitement automatique des émotions suggère que celles-ci influencent notre comportement et nos prises de décisions de manière rapide, inconsciente et sans nécessité de ressource. Savoir que les émotions sont traitées automatiquement a également une utilité clinique notamment dans la prise en charge de certaines pathologies telles que les troubles d'usage d'alcool ou encore le psycho traumatisme, dont on sait qu'ils présentent une altération du traitement émotionnel. Cependant, il n'est pas encore clair dans la littérature si les informations émotionnelles sont traitées automatiquement ou nécessite un traitement volontaire.

Recherches en faveur d'une automaticité du traitement émotionnel

L'automaticité du traitement des informations émotionnelles suscite un grand intérêt scientifique depuis plusieurs années, bien que la question demeure largement débattue

dans la littérature. Dès les premières recherches, de nombreuses études utilisant divers protocoles expérimentaux ont mis en évidence un *traitement plus rapide* et plus précis des expressions faciales d'émotions négatives, en comparaison avec les émotions positives. Ce traitement préférentiel concerne plus spécifiquement les émotions à caractère menaçant, telles que la colère ou la peur (Anderson et al., 2013; Öhman et al., 2001; Vuilleumier, 2002). Ce biais de négativité s'exprime tant au niveau comportemental, par des temps de réponse plus courts, qu'au niveau cérébral, par une activation accrue à un stade précoce, lors de la reconnaissance d'expressions de colère par rapport à des expressions de joie (Anderson et al., 2013; Carretie et al., 2004; Öhman et al., 2001; Palermo & Rhodes, 2007; Pegna et al., 2005; Tracy & Robins, 2008; Vuilleumier, 2002). Ce traitement prioritaire des signaux émotionnels menaçants est généralement interprété à la lumière de considérations évolutionnistes et adaptatives. Une réaction rapide face à un danger potentiel favorise la mise en place de comportements adaptés, augmentant ainsi les chances de survie. A l'inverse, un traitement cognitif trop lent ou trop coûteux en ressources pourrait compromettre la détection et l'évitement rapides de ces menaces. Dans cette perspective, l'attention sélective automatique dirigée vers les signaux émotionnels menaçants constitue un mécanisme fonctionnel fondamental (Palermo & Rhodes, 2007).

Au-delà de leur traitement rapide, les signaux émotionnels semblent également être traités en l'absence de *conscience* (Mumenthaler & Sander, 2015; Palermo & Rhodes, 2007; Vuilleumier, 2002). Plusieurs études ont ainsi montré que des expressions faciales présentées très brièvement (environ 33ms) peuvent être perçues et influencer les

comportements et les prises de décision, même sans accès conscient (Mumenthaler & Sander, 2015; Palermo & Rhodes, 2007). Sur le plan cérébral, l'amygdale, une structure clé dans le traitement des émotions, s'active lors de la présentation d'expressions faciales émotionnelles, y compris lorsque ces dernières ne sont pas perçues consciemment (Pegna et al., 2005). Ces résultats suggèrent que certaines dimensions du traitement émotionnel peuvent se dérouler de manière inconsciente, renforçant l'idée d'un traitement automatique.

Par ailleurs, l'automatisme du traitement émotionnel a également été étudiée via des paradigmes de double tâche, afin de déterminer la nécessité de *ressources cognitives* dans le traitement. Notamment, Shaw et ses collègues (2011) ont montré que, malgré la réalisation simultanée d'une tâche de discrimination auditive et une seconde tâche de reconnaissance du genre, les participants portaient préférentiellement leur attention vers les visages des émotions, en particulier la colère. Ces résultats suggèrent que, même lors d'une tâche non émotionnelle (discrimination auditive et reconnaissance du genre), l'attention est attirée préférentiellement par du contenu émotionnel plutôt que neutre (Shaw et al., 2011).

En conclusion, il semblerait que les indices émotionnels sont traités rapidement (Anderson et al., 2013; Öhman et al., 2001; Vuilleumier, 2002), inconsciemment (Mumenthaler & Sander, 2015; Palermo & Rhodes, 2007; Vuilleumier, 2002) et ne nécessitant pas de ressource cognitive (Shaw et al., 2011), suggérant ainsi un traitement

automatique de ces informations émotionnelles. Cependant, d'autres travaux ont remis en question ces résultats.

Nuances sur l'automatisme du traitement émotionnel

Comme précédemment explicité, les premières études soutenant l'hypothèse d'un traitement automatique des expressions émotionnelles reposaient notamment sur la démonstration d'un biais de négativité, se traduisant par une orientation *rapide* préférentielle de l'attention vers des signaux menaçants comme la peur ou la colère plutôt que vers des stimuli neutres (Anderson et al., 2013; Carretie et al., 2004; Öhman et al., 2001; Palermo & Rhodes, 2007; Pegna et al., 2005; Tracy & Robins, 2008; Vuilleumier, 2002). Toutefois, des études plus récentes n'ont pas systématiquement mis en évidence ce biais attentionnel envers les stimuli émotionnels (Khalid et al., 2017; Lien et al., 2013; Puls & Rothermund, 2018; Victeur et al., 2020; Wirth & Wentura, 2020). Par exemple, Lien et ses collègues (2013) ont exposé brièvement (125ms) des visages exprimant la peur ou une émotion neutre avant une tâche de recherche visuelle. Si les visages apeurés capturaient automatiquement l'attention, les participants auraient dû répondre plus rapidement lorsque la cible apparaissait à l'endroit précédemment occupé par ces visages. Or, aucune différence de temps de réponse n'a été observée, suggérant une absence d'avantage attentionnel pour les expressions émotionnelles (Lien et al., 2013). Une explication possible de ces résultats contradictoires réside dans le rôle de la fréquence d'apparition des distracteurs émotionnels (Maxwell et al., 2023). En effet, Grimshaw et ses collègues (2018) ont montré que le biais attentionnel envers des scènes émotionnelles

était plus marqué lorsque ces distracteurs étaient présentés dans seulement 25 % des essais, par rapport à une fréquence plus élevée de 75 %. De manière complémentaire, Maxwell et ses collègues (2023) ont mené trois expériences portant sur des expressions faciales et ont observé que la fréquence d'apparition des distracteurs modulait l'orientation de l'attention vers ces derniers. Ces résultats soulignent l'influence du contexte expérimental sur la saillance des stimuli émotionnels, et remettent ainsi en question l'idée d'un traitement émotionnel entièrement automatique. En effet, une exposition fréquente aux stimuli émotionnels, comme c'est souvent le cas dans les paradigmes expérimentaux (au moins 50 % voire 100 % des essais pourraient générer un effet d'habituation ou d'anticipation, diminuant leur pouvoir de captation attentionnelle (Maxwell et al., 2023).

Concernant le critère de *conscience*, plusieurs études ont montré que le traitement des stimuli émotionnels *implique des ressources attentionnelles* (Lichtenstein-Vidne et al., 2017; Schindler & Bublatzky, 2020; Schindler et al., 2022). Il apparaît notamment que certaines structures cérébrales impliquées dans la perception des émotions, telles que l'amygdale, requièrent une allocation attentionnelle suffisante pour assurer une analyse efficace des expressions faciales émotionnelles (Pessoa et al., 2002). Une étude menée à l'aide d'un paradigme de double tâche illustre ce point (Erthal et al., 2005). Les auteurs ont observé que la présentation d'images émotionnelles désagréables interférait avec la tâche visuo-spatiale principale d'orientation de barres. Les participants devaient indiquer l'orientation de ces barres tout en ignorant les images émotionnelles (désagréables ou

neutres). Les résultats ont mis en évidence un allongement des temps de réponse en présence de stimuli émotionnels désagréables par rapport aux stimuli neutres indiquant une interférence due à la charge émotionnelle des images (Erthal et al., 2005). De plus, cette interférence était modulée par la difficulté de la tâche ainsi que par la consommation d'alcool des participants, suggérant que la *charge cognitive globale* joue également un rôle déterminant dans l'impact des distracteurs affectifs sur la performance (Erthal et al., 2005). Ces résultats renforcent l'idée selon laquelle le traitement émotionnel n'est pas systématiquement automatique, mais dépend de divers facteurs contextuels tels que l'ambiguïté du stimulus (Maxwell et al., 2021) et le degré de difficulté de la tâche (Erthal et al., 2005; Pessoa, Padmala et al., 2005).

Enfin, concernant le critère d'*efficacité*, les résultats de Shaw et ses collègues (2011) contrastent avec ceux de Tomasik et ses collègues (2009), qui, en utilisant un protocole de double tâche similaire, n'ont pas observé de préférence attentionnelle pour les visages émotionnels dans une tâche de catégorisation émotionnelle (joie vs colère) (Tomasik et al., 2009). Cette contradiction est d'autant plus intrigante que les deux études utilisent les mêmes stimuli (Tottenham et al., 2009) et relativement le même paradigme expérimental. Une hypothèse avancée pour expliquer cette divergence concerne la nature des stimuli. Tomasik et ses collègues ont utilisé des stimuli moyennés (morphing de plusieurs visages), ce qui aurait pu introduire une forme d'ambiguïté perceptive. Sur la base de cette idée, Maxwell et ses collègues (2021) ont exploré l'impact de l'ambiguïté des expressions faciales sur le traitement automatique, en comparant des visages morphés

à des visages non altérés dans un paradigme de double tâche. Leurs résultats montrent que l'attention est automatiquement dirigée vers les expressions émotionnelles, sauf lorsqu'elles sont morphées (Maxwell et al., 2021). Selon les auteurs, cette absence d'automatisme dans les conditions de stimuli ambigus pourrait être due à une diminution de la saillance de certaines caractéristiques faciales clés, comme la bouche, dans les visages morphés. Cette altération rendrait l'expression plus difficile à interpréter, nécessitant alors une mobilisation accrue des ressources attentionnelles (Maxwell et al., 2021). Ces résultats soulignent l'importance de la clarté des signaux émotionnels pour permettre un traitement automatique, suggérant que ce dernier n'est pas systématique, mais dépend du degré d'ambiguïté des expressions faciales.

La question de l'automatisme dans la perception des expressions émotionnelles isolées fait l'objet d'un intérêt croissant dans la littérature scientifique. De nombreuses études se sont penchées sur la rapidité, l'efficacité ou la conscience impliquée dans ces processus (Amting et al., 2010; Baier et al., 2022; Kiss & Eimer, 2008; Lee et al., 2017; Lima et al., 2019; Smith, 2012; Tracy & Robins, 2008). Toutefois, les résultats demeurent hétérogènes : certaines recherches soutiennent l'existence d'un traitement automatique des informations émotionnelles, tandis que d'autres la remettent en question. Cependant, ces études se sont concentrées presque exclusivement sur des expressions émotionnelles faciales présentées de manière isolée, négligeant ainsi le rôle potentiel du contexte, notamment les expressions corporelles, dans l'émergence et la nature de cette automatisme.

Automaticité de l'effet contextuel

Comme évoqué précédemment, la perception conjointe des expressions faciales et corporelles influence la reconnaissance émotionnelle, suggérant une interaction entre ces deux indices (App et al., 2012; Kret, Stekelenburg et al., 2013; Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013). Cependant, la nature des mécanismes impliqués dans cet effet contextuel demeure débattue, notamment pour savoir s'il s'agit d'un processus cognitif automatique ou volontaire.

Peu d'études ont exploré la question du traitement automatique de l'intégration des expressions émotionnelles faciales et corporelles. Parmi elles, Aviezer et ses collègues (2011) ont mené une étude comportementale portant sur l'influence des expressions corporelles sur la reconnaissance des expressions faciales, tout en manipulant la *charge cognitive*. Les participants devaient identifier l'émotion d'un visage présenté dans un contexte corporel congruent ou incongruent, tout en réalisant simultanément une tâche de mémoire comprenant soit une seule lettre (faible charge cognitive), soit une série de cinq chiffres et une lettre (forte charge cognitive). Les effets contextuels se sont maintenus indépendamment de la charge cognitive, ce qui suggère un traitement efficace de l'effet contextuel, potentiellement automatique (Aviezer et al., 2011). D'autres travaux ont porté sur le critère de la *conscience*. Karaaslan et ses collègues (2020) ont démontré que cet effet s'observe même lorsque les stimuli sont présentés de façon subliminale (33ms), c'est-à-dire trop brièvement pour permettre une perception consciente. En effet, les résultats ont montré que la reconnaissance des émotions faciales restait plus précise dans

les conditions de congruences, appuyant l'idée d'un traitement inconscient (Karaaslan et al., 2020). Finalement, Meeren et ses collègues (2005) ont exploré la *rapidité* de ce traitement. Dans cette étude, les participants devaient identifier des expressions faciales de peur ou de colère, en contexte corporel congruent ou incongruent, sous une contrainte temporelle de 200ms et avec comme consigne de répondre le plus rapidement possible. Les auteurs ont mesuré à la fois les réponses comportementales et l'activité cérébrale via EEG. Les résultats ont montré une reconnaissance plus rapide et plus précise dans les conditions congruentes, ainsi qu'une détection du conflit dès 115ms après la présentation du stimulus (Meeren et al., 2005). Globalement, ces trois études semblent avancer l'hypothèse que l'intégration des expressions faciales et corporelles se déroule automatiquement via les critères d'efficacité, de conscience et de rapidité (Moors & De Houwer, 2006).

Néanmoins, certaines données remettent en question cette hypothèse. En effet, plusieurs travaux montrent que les expressions corporelles influencent davantage la reconnaissance des expressions faciales que l'inverse, révélant une interaction bidirectionnelle, mais asymétrique (Abo Foul et al., 2018; Lecker et al., 2020). Cette asymétrie pourrait indiquer que l'effet contextuel est sensible au focus attentionnel, et donc ne serait pas totalement indépendant des ressources attentionnelles allouées (Puffet et al., 2025). En outre, plusieurs limites présentes dans les études renforcent la méconnaissance des processus automatiques dans l'effet contextuel. Premièrement (1), il y a finalement assez peu d'études ont exploré le caractère automatique des effets contextuels, ce qui limite nos connaissances à ce sujet. Deuxièmement (2), certaines

conclusions sur l'absence d'effet contextuel reposent uniquement sur des analyses fréquentistes (pour un exemple, voir Aviezier et al., 2011), qui ne permettent pas de confirmer l'inexistence d'un effet. Une solution consiste dans l'utilisation d'analyses bayésiennes qui offrent un cadre statistique plus approprié pour tester cette hypothèse, en permettant d'estimer directement la probabilité d'absence ou de présence d'un effet (Lee & Wagenmakers, 2013; van den Bergh et al., 2020, 2023; van Doorn et al., 2021). Ce type d'analyse apporte ainsi un outil plus pertinent pour évaluer la robustesse des effets contextuels dans la perception émotionnelle. Troisièmement (3), une seule étude (Meeren et al., 2005) a examiné l'activité cérébrale liée à l'effet contextuel en manipulant le critère de rapidité. Aucune n'a encore exploré les critères d'efficacité ou de conscience au niveau neuronal, limitant ainsi notre compréhension des bases cérébrales de ce processus. Les données actuelles se restreignent à l'observation d'une détection rapide du conflit émotionnel et d'une différenciation précoce entre signaux faciaux et corporels (Chen et al., 2022; Gu et al., 2013; Li, 2021), sans pour autant montrer de manière claire l'automatisme du traitement au niveau cérébral. Enfin (4), la qualité des stimuli utilisés constitue une limite générant des biais dans les résultats obtenus. La plupart des études sur l'effet contextuel emploient des images émotionnelles de faible qualité, ne permettant pas d'étudier cet effet dans des conditions proches de la perception en situation réelle.

Objectif et hypothèses de recherche

Bien que la recherche sur l'effet contextuel entre les expressions faciales et corporelles se soit largement développée ces dernières années, plusieurs lacunes

persistent. (1) En particulier, la question de l'automatisme de ce traitement, à savoir s'il opère indépendamment de la conscience, des ressources cognitives disponibles et des limites de temps de traitement, reste encore peu explorée. Pourtant, déterminer dans quelle mesure ce processus est automatique est essentiel pour affiner les modèles actuels de la perception émotionnelle et mieux comprendre les mécanismes impliqués dans le traitement des indices émotionnels provenant de plusieurs canaux. (2) De plus, les études existantes se sont principalement concentrées sur l'influence des expressions corporelles sur la reconnaissance des émotions faciales. Or, la littérature récente montre que cette influence est bidirectionnelle et asymétrique, c'est-à-dire que les expressions faciales influencent également la reconnaissance des émotions corporelles, mais de manière différente (Lecker et al., 2020). Cette asymétrie suggère que l'effet contextuel n'est pas entièrement automatique, puisqu'il peut être modulé par l'attention. (3) Par ailleurs, à notre connaissance, une seule étude a manipulé un critère d'automatisme, la rapidité, ce qui ne permet pas de comprendre les mécanismes sous-tendant les interactions entre les indices faciaux et corporels. (4) Enfin, la majorité des recherches utilisent des stimuli de qualité limitée, ce qui réduit la généralisation des résultats à la perception réelle des émotions.

L'objectif principal de cette thèse de doctorat est donc d'examiner le caractère automatique de l'intégration des expressions faciales et corporelles perçues simultanément, tant au niveau comportemental que neural, en utilisant des stimuli émotionnels se rapprochant d'une perception réelle des émotions. L'hypothèse générale de la thèse, sur la base de la littérature actuelle, est que cette intégration se produit de

façon automatique. Pour tester cette hypothèse, trois études ont été menées sous forme d'articles scientifiques, chacun apportant de nouvelles connaissances ou des outils méthodologiques essentiels à la compréhension de la perception simultanée des expressions faciales et corporelles. Le premier article visait à développer et valider des stimuli d'expressions faciales et corporelles en situation de congruence et incongruence. Le second a exploré l'automatisme de l'intégration des expressions faciales et corporelles au niveau comportemental et neural en manipulant la charge cognitive, tandis que le troisième article a examiné cette automatisme en s'intéressant à la rapidité du traitement. Le premier article a été publié auprès du journal *Behavior Research Methods* (Puffet & Rigoulot, 2025a) et le second auprès du journal *Scientific Report* (Puffet & Rigoulot, 2025b). L'Appendice contient le certificat d'éthique émis par l'Université du Québec à Trois-Rivières.

Objectifs et hypothèses de la première étude

Les objectifs de ce premier article sont (1) de créer une nouvelle banque de stimuli statiques d'expressions émotionnelles de base, nommé *Emotionally Congruent and Incongruent Face-Body Static Set* (ECIFBSS), incluant sept expressions faciales et corporelles (joie, colère, surprise, peur, tristesse, dégoût et neutre) congruentes et incongruentes, et (2) de valider cet ensemble de stimuli. Ceux-ci ont été construits pour dépasser les limites des banques existantes en incluant des expressions faciales et corporelles de haute qualité, en couleur, avec une variabilité en termes d'âge, de genre et d'ethnies.

La validation des stimuli se divise en deux parties. La première consiste en une tâche de reconnaissance d'expressions faciales et corporelles isolées. La seconde inclut une tâche de reconnaissance identique à la précédente ainsi qu'une évaluation de l'intensité, de la valence, de l'activation et de l'authenticité des stimuli faciaux et corporels. Nous avons formulé sept hypothèses : (1) les émotions faciales seront mieux reconnues que les expressions corporelles; (2) certaines émotions, comme la joie pour les expressions faciales, auront des taux de reconnaissance plus élevés; (3) certaines émotions seront systématiquement confondues entre elles. Par exemple, lors de la reconnaissance des émotions faciales, une perception simultanée de la colère par le visage et du dégoût par le corps entrainera une moins bonne reconnaissance de la colère en raison des confusions systématiques entre ces émotions (Aviezer, Hassin, Ryan et al., 2008); (4) la reconnaissance du dégoût dans les postures corporelles sera moins précise; (5) nous nous attendons à ce que l'expression faciale de joie soit perçue comme l'émotion la plus authentique (Dawel et al., 2017); (6) nous prévoyons une corrélation entre l'intensité et l'activation pour les expressions faciales et corporelles (O'Reilly et al., 2016); enfin, (7) concernant la valence, nous anticipons que les expressions de joie et de surprise soient évaluées positivement, tandis que la colère, la peur et la tristesse seront évaluées négativement (Russell, 1980). En termes d'activation, les expressions de tristesse et de neutralité devraient être perçues comme faibles, tandis que celles de joie, surprise, peur, dégoût et colère devraient être élevées (Russell, 1980).

Objectifs et hypothèses de la seconde étude

Cette deuxième étude a pour objectif d'examiner l'automatisme de l'intégration des expressions faciales et corporelles, en particulier son déroulement temporel, en se basant sur le critère d'efficacité (Moors & De Houwer, 2006). Pour cela, plusieurs facteurs ont été manipulés : le type de focus attentionnel (sur le visage ou le corps), la valence des expressions faciales et corporelles (colère, tristesse, joie), la charge cognitive (élevée ou faible) et la congruence entre les expressions faciales et corporelles (congruente ou incongruente). Sur la base des études antérieures (Aviezer et al., 2011; Karaaslan et al., 2020; Meeren et al., 2005) portant sur l'automatisme de l'effet contextuel entre les indices faciaux et corporels, nous posons comme hypothèse principale que l'intégration des expressions faciales et corporelles constitue un processus automatique. Par conséquent, nous nous attendons à ne pas observer de différence dans l'effet contextuel en fonction de la charge cognitive. Trois hypothèses ont été testées en utilisant des mesures comportementales et EEG :

- *Hypothèse 1* : nous nous attendons à un effet contextuel classique entre les indices faciaux et corporels, c'est-à-dire une reconnaissance facilitée des émotions en situation de congruence et une reconnaissance plus difficile en situation d'incongruence émotionnelle. Au niveau comportemental, nous devrions observer un plus haut taux de reconnaissance et un temps de réponse plus rapide lorsque les émotions faciales et corporelles sont congruentes plutôt qu'incongruentes (Gu et al., 2013; Lecker et al., 2020). Au niveau neural, nous nous attendons à ce que les expressions incongruentes engendrent une plus large

- amplitude pour des composantes PE précoces (P100, N100, P200, N200) reflétant une détection rapide du conflit (Gu et al., 2013; Meeren et al., 2005).
- *Hypothèse 2* : nous faisons l'hypothèse que le focus attentionnel va moduler l'effet contextuel entre les indices faciaux et corporels (Lecker et al., 2020). Nous nous attendons à observer une influence des expressions faciales sur la reconnaissance des expressions corporelles (dans la condition focus sur le corps) moins importantes que l'influence des expressions corporelles sur la reconnaissance des émotions faciales. En termes de composantes PE, nous nous attendons, lorsque l'attention est dirigée vers l'expression faciale, que les expressions congruentes soient associées à une plus grande amplitude de la composante P300, suggérant l'extraction facilitée et la mobilisation d'une attention soutenue envers les informations congruentes (Gu et al., 2013; Li, 2021).
 - *Hypothèse 3* : nous faisons l'hypothèse de l'absence d'influence de la charge cognitive sur les autres facteurs, suggérant une automaticité de ces processus. Cela devrait conduire à un haut niveau de confiance dans l'hypothèse alternative, c'est-à-dire, aucune différence dans l'effet contextuel en fonction de la charge cognitive, que ce soit sur les mesures comportementales (taux de bonne réponse et temps de réponse) ou sur les mesures cérébrales (amplitudes des potentiels évoqués : P100, N100, P200, N200).

Objectifs et hypothèses de la troisième étude

Cette troisième étude a pour objectif d'examiner l'automatisme de l'intégration des expressions faciales et corporelles, en particulier son déroulement temporel, en se basant sur le critère de rapidité (Moors & De Houwer, 2006). Pour cela, plusieurs facteurs ont été manipulés : le type de focus attentionnel (sur le visage ou le corps), la valence des expressions faciales et corporelles (colère, tristesse, joie), la rapidité (rapide ou non) et la congruence entre les expressions faciales et corporelles (congruente ou incongruente). L'hypothèse principale est similaire à celle de la seconde étude, c'est-à-dire que l'intégration des expressions faciales et corporelles est un processus automatique. Par conséquent, nous nous attendons à ne pas observer de différences dans l'effet contextuel en fonction de la rapidité. Trois hypothèses ont été testées en utilisant des mesures comportementales et EEG :

- *Hypothèse 1* : nous nous attendons à observer un effet contextuel entre les indices faciaux et corporels, caractérisé par une amélioration de la précision de reconnaissance et un temps de réponse plus rapide lors de congruence plutôt que d'incongruence (Lecker et al., 2020). Au niveau neural, nous nous attendons à ce que les expressions émotionnelles incongruentes engendrent une plus grande amplitude des composantes précoces (P100, N100, P200, N200) reflétant une détection rapide des conflits (Gu et al., 2013).
- *Hypothèse 2* : nous prédisons que l'instruction de diriger l'attention sur les indices faciaux ou corporels va influencer l'effet contextuel (Lecker et al., 2020). Spécifiquement, l'influence des expressions faciales sur la reconnaissance des

émotions corporelles (quand le focus est sur le corps) devrait être plus faible que l'effet des indices corporels sur la reconnaissance des émotions faciales (quand le focus est sur le visage). Concernant les composantes PE, nous nous attendons à ce que le focus attentionnel sur les expressions faciales engendre une amplitude de la composante P300 plus importante lorsque les paires sont congruentes plutôt qu'incongruentes, indiquant un traitement émotionnel facilité et une allocation plus efficace de l'attention soutenue lors des situations de congruence émotionnelle (Gu et al., 2013; Li, 2021).

- *Hypothèse 3* : nous nous attendons à une intégration automatique des indices faciaux et corporels. Cela devrait se refléter par un fort niveau de confiance dans l'hypothèse alternative, c'est-à-dire l'absence de différences entre les conditions rapide et lente tant au niveau comportemental (taux de réponse et temps de réponse) que pour les mesures EEG (amplitudes des PE d'intérêt, P100, N100, P200, N200).

Chapitre 1

Création et validation d'une banque de stimuli d'expressions faciales et corporelles
congruents et incongruents

Les recherches sur l'effet contextuel entre les expressions faciales et corporelles ont majoritairement utilisé des stimuli d'émotions faciales et corporelles provenant de différentes banques de stimuli afin de générer des combinaisons congruentes et incongruentes (Abo Foul et al., 2018; Kret & de Gelder, 2013; Lecker et al., 2020; Reschke & Walle, 2021). Cette méthodologie engendre des questionnements sur la validité des stimuli utilisés pour évaluer avec précision l'influence des expressions faciales et corporelles au plus près d'une perception émotionnelle réelle.

Stimuli d'expressions faciales et corporelles émotionnelles

Divers ensembles d'expressions faciales et corporelles émotionnelles sont répertoriés dans la littérature. En ce qui concerne les expressions faciales, il existe un grand nombre d'ensembles d'expressions faciales émotionnelles (voir pour une liste plus exhaustive : www.face-rec.org/databases/ ou www.libguides.princeton.edu/facedatabase), incluant à la fois des expressions émotionnelles non basiques (Fernandes-Magalhaes et al., 2023) et des expressions émotionnelles basiques (Goeleven et al., 2008; Ma et al., 2015; Tottenham et al., 2009). Ces banques de stimuli varient en termes de qualité, cohérence, quantité de stimuli, ainsi que de diversité (par exemple, le genre, l'âge et l'ethnicité). De plus, certaines d'entre elles sont anciennes et ne peuvent plus concurrencer la qualité visuelle des images plus récentes. Contrairement au grand nombre d'ensembles d'expressions faciales, il n'en existe à notre connaissance que cinq pour les expressions corporelles

(Atkinson et al., 2004; de Gelder & van den Stock, 2011; Lopez et al., 2017; Schindler et al., 2008; Thoma et al., 2013). Ces outils sont essentiels et largement utilisés dans la recherche sur l'effet contextuel. Certains de ces ensembles d'expressions corporelles incluent sept émotions basiques (colère, dégoût, peur, joie, tristesse, surprise et neutralité) (Schindler et al., 2008; Thoma et al., 2013), tandis que d'autres en contiennent un plus petit nombre (4 à 5 basiques) (Atkinson et al., 2004; de Gelder & van den Stock, 2011; Lopez et al., 2017). Par ailleurs, un seul de ces ensembles propose des stimuli en couleur (Schindler et al., 2008). Il y a également une grande variabilité dans le nombre de femmes et d'hommes représentés (The BEAST : 31 femmes et 15 hommes; Atkinson et ses collègues (2004) : cinq femmes et cinq hommes; The Tilburg University Stimulus Set : 34 femmes et 16 hommes; the BESST : 41 femmes et 41 hommes). Enfin, tous ces ensembles de stimuli contiennent des acteurs de moins de 40 ans (Atkinson et al., 2004; Thoma et al., 2013) ou dont l'âge n'est pas précisé (de Gelder & van den Stock, 2011; Lopez et al., 2017; Schindler et al., 2008). Ces défauts nous ont amenés à considérer construire notre propre banque de stimuli, en nous inspirant des méthodologies utilisées précédemment.

Méthodologie de création et de validation de stimuli émotionnels

La méthodologie employée pour créer les stimuli d'expressions faciales et corporelles varie considérablement entre les études. Généralement, les chercheurs utilisent du matériel photographique de haute qualité, un éclairage contrôlé et un fond neutre pour standardiser les images (de Gelder & van den Stock, 2011; Lopez et al., 2017; Thoma et al., 2013). Toutefois, les instructions fournies aux participants pour produire des émotions

spécifiques diffèrent d'une étude à l'autre. Concernant les expressions faciales, certains chercheurs demandent explicitement aux participants de réaliser des expressions faciales spécifiques (p. ex., faire un visage souriant pour la joie) (Tottenham et al., 2009), ou encore d'exprimer chaque émotion de manière intense et claire (Goeleven et al., 2008). D'autres recherches, visant des expressions plus standardisées (van der Schalk et al., 2011), s'appuient sur les unités d'action faciales définies par Ekman, reliant des mouvements musculaires spécifiques aux émotions (Ekman, 1992, 1999; Ekman & Friesen, 1976). Dans ces cas, les participants reçoivent par exemple un manuel de formation à l'avance, suivi d'une session pratique pour reproduire avec précision les expressions à partir des unités d'action (van der Schalk et al., 2011). Concernant les stimuli d'expressions corporelles, une approche différente est plus habituellement adoptée, en utilisant des scénarios inspirés de situations quotidiennes générant spécifiquement chaque émotion (de Gelder & van den Stock, 2011; Thoma et al., 2013). Par exemple, pour susciter la colère, les expérimentateurs lisaient un scénario tel que « imaginez que vous avez une dispute avec quelqu'un » et les participants exprimaient ensuite de la colère à travers leur posture corporelle (Thoma et al., 2013).

Les diverses méthodologies employées pour créer des stimuli émotionnels ont permis de constituer plusieurs ensembles de stimuli ayant chacune des caractéristiques spécifiques. Cependant, certaines interrogations persistent, notamment sur la création d'expressions émotionnelles standardisées et prototypiques (van der Schalk et al., 2011) ou d'émotions de forte intensité (Goeleven et al., 2008). En réalité, les émotions se

manifestent souvent de manière variable dans la vie quotidienne. Par exemple, il existe plusieurs façons d'exprimer la peur ou la tristesse (van den Stock et al., 2007). Par conséquent, il apparaît essentiel que les ensembles de stimuli intègrent différentes variations d'une même émotion afin de mieux représenter la diversité de l'expression émotionnelle. De plus, se limiter à des émotions de forte intensité néglige un aspect fondamental de la perception émotionnelle, car nous sommes fréquemment confrontés à des émotions d'intensité variable et parfois ambiguë, ce qui peut compliquer la reconnaissance émotionnelle (Morningstar et al., 2021). L'utilisation de scénarios inspirés de la vie quotidienne, adoptée par certains ensembles de stimuli, semble offrir une approche plus écologique et réaliste. Cette méthode reflète davantage la variabilité naturelle dans l'expression et l'intensité des émotions (de Gelder & van den Stock, 2011; Thoma et al., 2013).

La validation des stimuli émotionnels repose sur diverses mesures dans la littérature. Premièrement, la reconnaissance des expressions faciales ou corporelles est souvent testée via un choix forcé entre plusieurs émotions de base, incluant parfois une option neutre ou « autre » (de Gelder & van den Stock, 2011; Goeleven et al., 2008; Lopez et al., 2017; Schindler et al., 2008; Thoma et al., 2013; Tottenham et al., 2009; van der Schalk et al., 2011). D'autres évaluations mesurent des dimensions spécifiques telles que l'activation (forte ou faible), la valence (positive ou négative), l'authenticité (naturel ou non naturel) et l'intensité (pas intense ou fortement intense) (Goeleven et al., 2008; Lopez et al., 2017; Thoma et al., 2013). La valence et l'activation sont essentielles pour vérifier que les

émotions sont perçues selon les deux dimensions du modèle dimensionnel, un modèle dominant dans le domaine de recherche (Russell, 1980). Par exemple, une émotion de joie devrait être évaluée avec une valence positive et une activation élevée. Les mesures d'intensité et d'authenticité offrent, quant à elles, la possibilité de manipuler ces aspects, en ajustant ou en garantissant une uniformité de ces deux éléments entre les stimuli.

Ainsi la variété des ensembles de stimuli faciaux et corporels disponibles constitue un atout précieux pour la recherche, en offrant une large gamme d'options concernant les émotions représentées, le nombre de stimuli, et les caractéristiques des acteurs, ainsi que des méthodes variées de création et de validation. Toutefois, cette diversité peut introduire des biais lorsque différents ensembles sont combinés pour générer des stimuli émotionnels congruents et incongruents, en raison par exemple de divergences méthodologiques lors de la création ou de la validation de ces expressions.

Création de stimuli congruents et incongruents dans l'étude de l'effet contextuel

Lorsque les chercheurs étudient spécifiquement l'effet contextuel entre les expressions faciales et corporelles, ils combinent souvent des expressions faciales et corporelles issues d'ensembles de stimuli différents (Abo Foul et al., 2018; App et al., 2012; Karaaslan et al., 2020; Kret & de Gelder, 2013; Kret, Stekelenburg et al., 2013; Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013; Willis et al., 2011; Zhang et al., 2019). Pour ce faire, ils sélectionnent des expressions faciales d'un ensemble et des expressions corporelles d'un autre, puis les fusionnent à l'aide de logiciels de traitement

d'images comme Photoshop, afin de créer des stimuli congruents et incongruents. Cependant, cette méthode aboutit souvent à la création de stimuli de faible qualité (en noir et blanc), avec des disproportions entre les visages et les corps, ainsi que des variations de luminosité, compromettant ainsi le réalisme des stimuli (voir Figure 11). Ces limitations peuvent significativement altérer les observations, car la qualité des images influence la reconnaissance des émotions (Wallbott, 1992). En conséquence, l'utilisation de stimuli de haute qualité est essentielle pour étudier les émotions fidèlement à la perception réelle. C'est pourquoi, nous avons donc (article 1) créé et validé un ensemble de stimuli intégrant à la fois des expressions faciales et corporelles dans des situations de congruences et incongruences constituant un outil précieux et complémentaire aux ensembles existants. Cela surmonte les limites actuelles et améliorerait la qualité des observations dans la recherche sur l'effet contextuel.

Figure 11

Stimuli Congruents et Incongruents Utilisés dans de Précédentes Etudes sur l'Effet Contextuel



Visage : Dégout
Corps : Dégout



Visage : Peur
Corps : Peur



Visage : Joie
Corps : Dégout

Note. L'image à gauche provient de Karaaslan et ses collègues (2020), l'image centrale de Lecker et ses collègues (2020) et celle de droite de Li (2021).

Chapitre 2

Article 1 – Validation of the Emotionally Congruent and Incongruent Face-Body Static Set (ECIFBSS)

Validation of the Emotionally Congruent and Incongruent Face-Body Static Set (ECIFBSS)¹

Anne-Sophie Puffet^{1,2} & Simon Rigoulot^{1,2}

¹ Department of Psychology, University of Quebec at Trois-Rivières.

² Research group CogNAC (Cognition, Neurosciences, Affect and Behaviour).

Author Note

Correspondence concerning to this article should be addressed to Anne-Sophie Puffet, anne-sophie.puffet@uqtr.ca and Simon Rigoulot, simon.rigoulot@uqtr.ca.

This work was supported by “Natural Sciences and Engineering Research Council of Canada” (NSERC) Discovery Grant (Grant numbers: RGPIN-2020-06706).

The authors declare that they have no conflict of interest, financial, or otherwise.

Author’s contribution:

Anne-Sophie Puffet: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing-original draft, Writing-reviewing, and editing. **Simon Rigoulot:** Conceptualization, Funding acquisition, Resources, Supervision, Writing-reviewing and editing, Investigation, Methodology, Project administration, Validation.

¹ Les normes de publication pour la revue *Behavior Research Methods* sont disponibles à <https://link.springer.com/journal/13428/submission-guidelines>

Abstract

Frequently, we perceive emotional information through multiple channels (e.g., face, voice, posture). These cues interact, facilitating emotional perception when congruent (similar across channels) compared to incongruent (different). Most previous studies on this congruency effect used stimuli from different sets, compromising their quality. In this context, we created and validated a new static stimulus set (ECIFBSS) featuring 1952 facial and body expressions of basic emotions in congruent and incongruent situations. We photographed 40 actors expressing facial emotions and body postures (anger, disgust, happiness, neutral, fear, surprise, and sadness) in both congruent and incongruent situations. The validation was conducted in two parts. In the first part, 76 participants performed a recognition task on facial and bodily expressions separately. In the second part, 40 participants performed the same recognition task, along with an evaluation of four features: intensity, authenticity, arousal, and valence. All emotions (face and body) were well recognized. Consistent with the literature, facial emotions were recognized better than body postures. Happiness was the most recognized facial emotion, while fear was the least. Among body expressions, anger had the highest recognition, while disgust was the least accurately recognized. Finally, facial and bodily expressions were considered moderately authentic, and the evaluation of intensity, valence and arousal aligned with the dimensional model. The ECIFBSS offers static stimuli for studying facial and body expressions of basic emotions, providing a new tool to explore the integration of emotional information from various channels and their reciprocal influence.

Keywords: emotion, facial emotions, body postures, congruency effect, stimulus set.

Introduction

Our daily social interactions are infused with emotions. Emotions play a fundamental role in our relationships and well-being, and help to adopt appropriate social behavior such as approach or avoidance (Marsh et al., 2005). Historically, research on emotions has primarily focused on visual modality, using especially facial emotions (Ekman, 1992, 1999; Russell, 1980, 1994). This led to a significant prevalence of these studies, such as that in 2009, it was calculated that 95% of studies on emotion perception were using emotional facial expressions (de Gelder, 2009).

In a broader sense, scientific literature has emphasized the paramount significance of facial expressions in the perception of emotions. Nonetheless, the reliability of facial emotional cues is not always absolute. Recent research has drawn attention to the adverse effects of concealing the face, a situation that has become prevalent during the pandemic, on the accurate assessment of expressed emotions (Tsantani et al., 2022). In this case, other sensory modalities are often used to disambiguate and recognize an emotion. Studies have revealed that when the facial emotions are ambiguous, such as expressions combining anger and disgust (Karaaslan et al., 2020), or when they are masked, as exemplified by the use of a gray mask to reduce the perception of fear and disgust (Collignon et al., 2008), participants tend to resort to other channels of emotional information, such as vocal cues (Collignon et al., 2008) and body posture (Karaaslan et al., 2020), for the categorization of facial emotions.

In fact, in everyday life, emotions are frequently communicated through a multitude of sensory channels, encompassing visual, auditory, olfactory, and more. The perception of emotions entails the integration of these multiple channels, combining information from facial expressions and body posture (Karaaslan et al., 2020; Lecker et al., 2020), background scene (Reschke & Walle, 2021), auditory (Talwar et al., 2023) and olfactory (Novak et al., 2015)

stimuli. These simultaneously perceived multimodal emotional channels influence each other (Liu et al., 2012; Novak et al., 2015; Pourtois et al., 2002; Seubert et al., 2010; van den Stock et al., 2007; Zimmer et al., 2022).

Several studies have focused on the influence of these channels on the perception of emotional facial expressions (Abo Foul et al., 2018; Balconi & Carrera, 2011; Collignon et al., 2008; Falagiarda & Collignon, 2019; Pourtois et al., 2002; Seubert et al., 2010; Song et al., 2022; van den Stock & de Gelder, 2014; van den Stock et al., 2007). These works have shown that in multimodal congruent situations, i.e., when emotions are perceived consistently through both facial expressions and another sensory modality, they are recognized more quickly and accurately than when this facial emotion is perceived alone (unimodal; (Kreifelts et al., 2007; Paulmann & Pell, 2011). However, when different emotional information is provided by sensory channels, making the emotional content incongruent between the face and other sensory inputs (incongruent situation), identifying facial emotions becomes more challenging (Collignon et al., 2008; Focker et al., 2011). This highlights how several sensory modalities are involved in emotion recognition and how conflicting emotional cues across sensory channels affect the recognition of facial emotions and the time required for this recognition (Gerdes et al., 2014). It appears that this influence occurs independently of attentional focus, whether on facial or vocal emotions (Collignon et al., 2008; Vroomen et al., 2001) and participants' tasks (Vroomen et al., 2001). Taken together, these studies have demonstrated that modalities of multimodal emotions either enhance emotion recognition during congruence or introduce bias during incongruence.

The interaction between multiple emotional channels also manifests within the context of emotions originating from various channels within the same sensory modality. For example, interactions between facial emotions and body postures, reflecting the congruency effect (Lecker et al., 2020; Mondloch et al., 2013), or contextual effect (Aviezer et al., 2017), or, in a more

specific context, the bodily context effect (Karaaslan et al., 2020), play a significant role in the study of emotion perception. Indeed, the recognition of facial emotions is notably improved when congruent with body posture, leading to faster recognition. On the other hand, incongruent body postures or emotional contexts introduce bias in the recognition of facial emotions (Aviezer et al., 2011; Meeren et al., 2005; Mondloch et al., 2013; Reschke & Walle, 2021). Moreover, incongruent body postures appear to influence accuracy (Lecker et al., 2020) in facial emotion recognition, findings corroborated by works at the brain level showing early detection (100ms) of incongruent emotional information from both facial expressions and body postures (Meeren et al., 2005).

Several observations in the literature have clarified the interaction between facial expression and body posture in the perception of emotions. First, this interference between facial expressions and body postures appears to be bidirectional (Albohn et al., 2022; Lecker et al., 2020). As previously described, the recognition of facial emotions is biased by incongruent body postures, but conversely, the recognition of body postures is also biased by incongruent facial emotions (Albohn et al., 2022; Lecker et al., 2020). Second, these bidirectional influences seem to be asymmetrical (Karaaslan et al., 2020; Mondloch et al., 2013). For instance, an emotion expressed through body posture, such as sadness, strongly affects the recognition of facial emotion, like anger. On the contrary, when the opposite configuration is perceived, with anger expressed through the body posture and sadness through facial expression, the influence of the body posture on the recognition of the facial emotion is less pronounced (Mondloch et al., 2013). Lastly, not all pairs of incongruent emotions (facial emotion and body posture) necessarily lead to substantial errors in recognition of facial expressions or body postures (Karaaslan et al., 2020; Mondloch et al., 2013; Reschke & Walle, 2021). Certain pairs of body and facial expressions generate minimal bias in the recognition of the target emotion (either facial or bodily), while

others result in a higher degree of errors in emotion recognition. For example, Aviezer et al. (2008) demonstrated that the simultaneous perception of the facial emotion of anger along with a body posture expressing disgust leads to the misrecognition of anger due to confusion with the body posture of disgust (Aviezer, Hassin, Ryan et al., 2008). Conversely, simultaneous perception of the facial emotion of happiness with a body posture expressing anger generates fewer recognition errors of facial emotion of happiness (Willis et al., 2011).

Two theories have been proposed to explain this bodily context effect, characterized by decreased accuracy (i.e., more emotion recognition errors) and increased response time in cases of incongruence between facial emotions and body postures (Mondloch et al., 2013): the emotional seed model (Aviezer, Hassin, Bentin et al., 2008; Susskind et al., 2007) and the dimensional model (Russell, 1980). These theories offer distinct explanations and predict various influences between emotions from facial and bodily expressions. The dimensional model (Russell, 1980) represents emotions in the form of a circumplex categorizing facial emotions along two axes: arousal (strong and weak) and valence (positive and negative). Emotions situated close to each other on the circumplex share position on both axes (e.g., anger and disgust both possess negative valence and strong arousal), whereas emotions positioned far apart share few or no positions on both axes (e.g., happiness, and sadness have different valence and arousal). According to the dimensional model, the congruency effect is most pronounced when facial and body postures are the closest on both axes (e.g., anger and disgust). Conversely, when facial emotions and body postures are further away, the congruency effect is weaker (e.g., happiness and sadness; (Aviezer, Hassin, Bentin et al., 2008; Aviezer, Hassin, Ryan et al., 2008; Mondloch et al., 2013).

The emotional seed model is based on the sharing of physical characteristics among emotions (Aviezer, Hassin, Bentin et al., 2008; Susskind et al., 2007). Facial emotions have

specific physical characteristics, and some of these characteristics are shared by several emotions (Aviezer, Hassin, Bentin et al., 2008). For instance, anger and disgust share many physical characteristics (e.g., furrowing of the brows, muscle tension, dilated nostrils) while anger and fear share some, and anger and surprise share few (Aviezer, Hassin, Bentin et al., 2008). The congruency effect between facial emotion and bodily postures can be explained in terms of the ambiguity levels generated by the number of physical characteristics shared by emotions from face and body posture. When emotions strongly resemble each other (e.g., anger and disgust), this generates a high level of ambiguity and a maximum congruency effect. Conversely, when emotions share few physical characteristics (e.g., anger and surprise), this results in lower ambiguity and a weaker congruency effect (Aviezer, Hassin, Bentin et al., 2008; Aviezer, Hassin, Ryan et al., 2008). It seems then essential to expand the current knowledge of the intricate interactions between emotions from multiple channels such as the face and the body posture. To achieve this, researchers often combine a facial emotion stimulus set with a bodily posture stimulus set to create stimuli that include both congruent and incongruent groupings of facial emotions and body postures (Abo Foul et al., 2018; Kret & de Gelder, 2013; Lecker et al., 2020; Reschke & Walle, 2021). Various sets of facial emotions and body postures are available in the literature for these purposes.

Indeed, a very large number of sets of emotional facial expressions exists (see an exhaustive list in www.face-rec.org/databases/), including *non-basic* (Fernandes-Magalhaes et al., 2023; Mende-Siedlecki et al., 2020) and *basic* emotional expressions databases (Goeleven et al., 2008; Ma et al., 2015; Tottenham et al., 2009). These stimulus sets possess unique characteristics, including quality, consistency, quantity of stimuli, and variability (e.g., gender, age, ethnicity). The set of facial and bodily expression stimuli created and validated in this article includes specifically emotional basic expressions. To the best of our knowledge, and in contrast

to the large number of facial stimuli databases, there appear to be only five stimulus sets of body posture images (Atkinson et al., 2004; de Gelder & van den Stock, 2011; Lopez et al., 2017; Schindler et al., 2008; Thoma et al., 2013). These sets of bodily expressions are essential tools and have been widely used in studying congruency effects (Karaaslan et al., 2020; Kret & de Gelder, 2013; Kret et al., 2013; Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013; Willis et al., 2011; Zhang et al., 2019).

Some sets of stimuli include seven basic emotions (anger, disgust, fear, happiness, sadness, surprise, and neutrality; (Schindler et al., 2008; Thoma et al., 2013), while others include a smaller number (4 to 5) of basic emotions (Atkinson et al., 2004; de Gelder & van den Stock, 2011; Lopez et al., 2017). Only one database is in color (Schindler et al., 2008). There is also a lot of variability in the number of women and men included in the databases (The BEAST: 31 women and 15 men; Atkinson et al. (2004): five women and five men; the Tilburg University Stimulus Set: 34 women and 16 men; the BESST: 41 women and 41 men). Finally, all databases contain actors whose ages are limited under 40 years (Atkinson et al., 2004; Thoma et al., 2013) or are unspecified (de Gelder & van den Stock, 2011; Lopez et al., 2017; Schindler et al., 2008).

In our opinion, the creation of a database taking these different characteristics into consideration and containing facial and bodily expressions in congruent and incongruent situations would be a very important and complementary tool to the existing databases. Firstly, to the best of our knowledge, the database in this article is one of the only ones containing facial and bodily expression stimuli from the same actors in both congruent and incongruent situations. Secondly, providing colored pictures of actors for whose gender is balanced and age range is wider (especially for body expressions) aims for more ecological stimuli. Finally, the inclusion of features of intensity, valence, arousal, and authenticity for facial and body emotion expressions

and our investigations of interaction effects between genders of actors and raters would bring unique relevant information for future users of our database.

Consequently, when authors (Karaaslan et al., 2020; Lecker et al., 2020; Li, 2021; Reschke & Walle, 2021) combine face expression and body posture from various stimulus sets to study congruency effect, these limitations impact the ecological validity of these stimuli. This manipulation results in low-quality stimuli, with facial sizes not always proportional to the body and variations in brightness and image quality, thus lacking realism. These limitations could significantly affect the quality of observations, as poor image quality can impact emotion recognition (Wallbott, 1992). Therefore, the use of high-quality stimuli is essential to study emotion perception with the highest realism.

Therefore, our objectives were (1) to create a novel stimulus set of basic static emotional expressions, the “Emotionally Congruent and Incongruent Face-Body Static Set” (the ECIFBSS), encompassing seven facial and bodily expressions of basic emotions (happiness, anger, surprise, fear, sadness, disgust, and neutrality) in both congruent and incongruent situations, and (2) to validate this stimulus set. This set of emotional stimuli was built to overcome the limitations of existing stimulus sets and includes emotional facial and body posture stimuli in congruent and incongruent situations. The stimuli are of high quality, full-color, and incorporate variability in age, gender, and basic emotions (surprise, happiness, disgust, sadness, fear, anger, neutrality).

Stimulus validation is divided into two parts. The first part solely comprises a facial and body expression recognition task. The second part includes a recognition task and an evaluation of the intensity, valence, arousal, and authenticity of the facial and bodily stimuli. We formulated seven hypotheses (1) facial emotion will be better recognized than body postures. (2) certain emotions, such as happiness, will have higher recognition rates compared to others. (3) certain emotions will exhibit consistent confusion with each other. For example, during facial emotion

recognition, simultaneous expressions of emotions through the face (e.g., anger) and body posture (e.g., disgust) often lead to frequent categorization errors due to systematic confusion between these emotions (i.e., the facial expression of anger will be frequently confused with the disgust; (Aviezer, Hassin, Ryan et al., 2008). (4) the recognition of disgust in body posture would be less accurate. (5) Following, we expect the facial emotion of happiness to be rated as the most authentic emotion (Dawel et al., 2017), (6) we also expect that intensity and arousal will be correlated for both facial and bodily expressions (O'Reilly et al., 2016). Finally (7) concerning valence, we expect the facial and bodily expressions of happiness and surprise to be evaluated as having positive valence, while anger, fear, and sadness are expected to be evaluated as having negative valence (Russell, 1980). Regarding arousal, we expect the facial and bodily expressions of sadness and neutrality to be low in arousal while those of happiness, surprise, fear, disgust, and anger are expected to be high in arousal (Russell, 1980).

Stimuli Material Creation

Method

Participants

Forty amateur actors (22 females; $M_{\text{age}} = 37.1$ years; $SD = 13.6$ years; range 18-70 years) were recruited from June to September 2022 via social media advertising. Participation was compensated CAD 25 for one hour of acting. Thirty-two participants were of Caucasian ethnicity (80%), one individual identifying as Black (2.5%), one from Morocco (2.5%), one as Caribbean (2.5%), two as Latin-American (5%) and two as indigenous/aboriginal (5%). One participant did not disclose their ethnicity.

All actors have consented to the images being disseminated, communicated, or made accessible to the public by any means and on all media, within the framework of the project and all related activities. All actors except two have consented to the secondary use of the images so

that they can be stored on the Open Science Framework (OSF) in a stimulus set intended for the scientific community, and they may be used by its members in the future, within the context of other research projects.

The inclusion criteria mandated that participants fall within the age range of 18 to 70 years and possess acting experience (as actors, improvisational performers, or theatre practitioners). Twenty-three participants had 5 years or less of acting experience (57.5%), eight had between 6 and 10 years of experience (20%), five participants had between 11 and 25 years of experience (12.5%), and two participants had more than 35 years of experience (5%). To obtain further information on participants' demographic details, supplementary materials are available on the Open Science Framework (OSF; <https://osf.io/frvwt/>).

Materials

The pictures were captured against a white background using a professional Nikon D800 camera mounted on a tripod positioned 2.65 meters away. Two Elinchrom D-Lite flashes were also set up, each positioned 2.10 meters from the background to ensure proper lighting. Half of the photoshoot sessions were held at the University of Quebec at Trois-Rivières (UQTR), while the remaining half occurred on the UQTR campus in Quebec City. We conducted the tests at two different sites to avoid recruitment difficulties. Subsequently, the images were standardized using Photoshop 2022 software (RRID:SCR_014199). Each image underwent background removal through precise clipping, following which the stimuli were placed against a uniform white background measuring 3500 x 4000 pixels. Furthermore, the height of the stimuli was standardized to 3000 pixels. To control a viewpoint effect, particularly when actors raised their arms (e.g., during expressions of happiness), we determined a custom height for each actor. This procedure was built from a previous study that specified minimal scale differences between

stimuli (Schindler et al., 2008). This adjustment ensured that there would be no discernible variations in height across different images featuring the same actor (see Figures 1 and 2).

Figure 1. *Examples of Congruent Stimuli*



Note. The size is adjusted to control the viewpoint effect.

Figure 2. *Examples of Incongruent Stimuli*



Note. The first emotion is body posture and the second is facial emotion (e.g., Neutrality – Disgust, body posture is neutrality and facial emotion is disgusting). The size is adjusted to control the viewpoint effect.

Procedure

We created a total of 1952 stimuli, with each actor producing 49 images. There are eight missing images for participants number five (three images), seven (three images), and eight (two images) due to a technical problem with the camera.

Participants were provided with standardized instructions to wear dark clothing, and shoes preferably black and to guide the production of their expressions. They were instructed to express the seven emotions (happiness, surprise, anger, fear, sadness, disgust, and neutrality) through their facial expressions and body postures, both in congruent (same facial and body expression) and incongruent (different facial and body expression) situations (see Figures 1 and 2). To do that, participants maintained their body postures and sequentially altered their facial expressions to convey the other six facial basic emotions (incongruent, e.g., continuing to express happiness with their body posture while displaying anger with their facial expressions). Each emotional expression was separated with a return to the congruent situation, where, for example, they would revert to a happy facial expression while maintaining the happy body posture.

Furthermore, each emotion was associated with a real-life scenario, such as anger linked to a conflict with someone, happiness associated with winning a lottery, sadness with the death of a loved one, surprise prompted by an unexpected job promotion, disgust evoked by something repulsive, fear inducted by a dangerous situation, and neutrality represented a relaxed posture devoid of any emotion. The scenarios were adapted from previous studies including the creation of the BESST (Thoma et al., 2013) and the BEAST image set (de Gelder & van den Stock, 2011). The order in which these scenarios were presented was counterbalanced across participants.

The validation of stimuli was performed in two separate experiments, the first one (part 1) with 76 participants and only a recognition task, the second (part 2) with 40 participants who had

to realize a recognition task and evaluate supplementary features (authenticity, arousal, intensity, and valence).

Validation of Stimuli: Part 1

Methods

Participants

Eighty-six participants (55 females; $M_{\text{age}} = 25.42$ years; $SD = 4.62$ years; range 18-36 years) were recruited from December 2022 to February 2023 through social media postings and from two undergraduate psychology courses offered by the Department of Psychology at UQTR. Each participant received CAD 10 as compensation for the one-hour session. Before the experiment, all participants completed *The Hospital Anxiety and Depression Scale* (HADS; Zigmond & Snaith, 1983), a questionnaire designed to assess mood. This step was essential, as the presence of anxiety or depression tendencies could potentially affect the processing of emotions (Peschard et al., 2014). The inclusion criterion for participation in the validation of stimuli was an age range of 18 to 35 years. Several exclusion criteria were considered when recruiting participants, including the presence of specific neurological disorders (e.g., traumatic brain injury), diagnosed psychological conditions (e.g., psychosis) or uncorrected visual impairment, and the use of medications or substances (e.g., drugs) that affect the nervous system. To obtain further information on participants' demographic details, supplementary materials are available on the Open Science Framework (OSF; <https://osf.io/frvwt/>).

Ten participants scored above 11 on the anxiety scale, indicating a tendency toward anxiety. However, all participants scored below 11 on the depression scale, indicating an absence of depressive tendencies during the experiment. Given the potential influence of anxiety levels on emotion perception (Peschard et al., 2014), data from participants with an anxiety score exceeding 11 were excluded from the analysis. This resulted in a final dataset comprising

76 participants (47 females; $M_{\text{age}} = 25.55$ years; $SD = 4.64$ years; range 18-36 years) with mean anxiety (mean score = 6.5) and depression (mean score = 3.5) scores below 11 (Zigmond & Snaith, 1983). These 76 participants were assigned among four groups (Group 1: 15 females; Group 2: 11 females; Group 3: 11 females; Group 4: 10 females), each consisting of 19 participants, with each group being exposed to a distinct set of stimuli. Further details regarding this are provided in the “Procedure” section.

The sample size was determined based on prior studies of similar stimulus creation (de Gelder & van den Stock, 2011; Thoma et al., 2013), and the observation that results remained consistent with the inclusion of additional participants.

Materials

For validation purposes, we presented only the facial expressions or the body postures from the created stimuli. In total, we used 280 body images and 1952 facial images for a total of 2232 stimuli. Specifically, we employed facial emotions in both congruent and incongruent situations, utilizing all 1952 facial images. For body postures, we selected the congruent image for each emotion for every actor, resulting in 280 body images. Considering that the body posture for each emotion did not change (e.g., the general posture, the orientation of the arms, etc.) as a function of congruence (e.g., participants maintained the body posture of anger and changed facial expression), we only used one exemplar of body posture for each emotion in the congruent condition.

To generate the facial emotion stimuli (see Figure 3), we employed Adobe Photoshop 2022 software to isolate the faces from the body. Each facial image was consistently cropped to the dimensions of 350 x 510 pixels. For the body posture stimuli, we included a gray oval to isolate the body expression while keeping the facial expressions intact in each of the stimuli.

Figure 3. *Examples of Face Emotion Stimuli*



Procedure

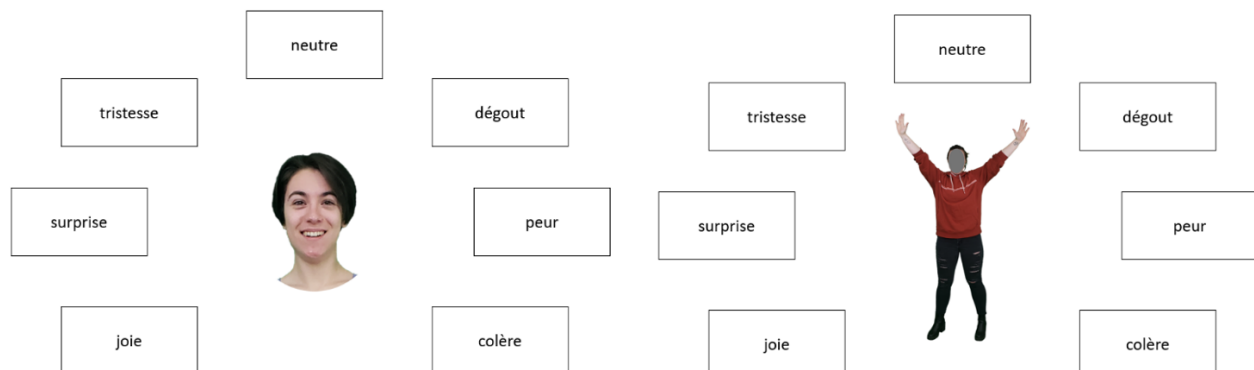
Participants were instructed to complete three questionnaires on LimeSurvey before the experiment: a socio-demographic questionnaire, an information and consent questionnaire, and the HADS (Zigmond & Snaith, 1983). Filling out the online questionnaires took a mean of ten minutes. During the experiment, participants were seated in a chair approximately 53 cm from a computer screen. The instructions were presented in French using the “Courier New” font at 18 pt size, displayed in black on a white background, and accompanied by a visual illustration of the task (see Figure 4).

We conducted the validation of 2232 stimuli (1952 face and 280 body stimuli) using a computer task implemented in E-PRIME 2.0 software. We divided the stimulus set into four subsets. Each subset consisted of 488 facial emotions stimuli and 70 body posture stimuli. A total of 558 stimuli were presented in each group of participants (nineteen participants/group). Participants' group assignment and stimulus subset assignment were done randomly. The session lasted a maximum of one hour for each participant. The experiment was divided into two blocks one for facial emotion recognition and the other for body posture recognition. The order of these blocks was counterbalanced among the participants and the order of facial and body stimuli was randomized. In the forced-choice task, participants were required to use the mouse to select the emotion that best corresponded, according to them, to either the facial or bodily emotion from the

seven emotions (happiness, anger, sadness, disgust, fear, surprise, neutrality). The images, which were sized at 49% x 22% for facial images and 75% x 40% for body images, were centered on the computer screen. The seven emotion options were evenly spaced around the image, as shown in Figure 4.

Each trial was displayed for 6000 milliseconds, and participants could end the trial by selecting an emotion. The positioning of the emotions on the screen was counterbalanced across participants. Before each block, participants completed a training session consisting of six trials. If a participant achieved less than 50% correct answers, they repeated the training. Finally, numerous breaks (every 122 stimuli) were offered to the participants during the session to mitigate the impact of fatigue on the task.

Figure 4. Example of a Trial for the Face (left) and Body (right) Expressions



Note. Seven basic emotions (joie = happiness; surprise = surprise; tristesse = sadness; neutre = neutrality; dégoût = disgust; peur = fear; colère = anger) were positioned around the image. The position of each emotion is counterbalanced between participants.

Measures of Interest

We assessed the quality of the responses of participants through three key parameters. First, we calculated Fleiss's Generalized kappa scores (Fleiss, 1971) for both facial and bodily stimuli. This score quantifies the consensus rate among participants regarding whether the stimuli for each emotion are considered representative of that emotion. This score has also been used in

other stimulus creation studies (de Gelder & van den Stock, 2011; Lopez et al., 2017). Second, we measured the internal consistency of participant's ratings using split-half reliability scores on accuracy data. This technique involves randomly dividing each of the four participant groups into two subgroups based on participant numbers (even and odd) and into two subgroups based on the gender of the participants (men and women). Next, we calculated the mean accuracy of each stimulus within each subgroup to perform Pearson's correlations and measure the Spearman-Brown coefficient. This method has previously been employed in other stimulus validation studies (Fernandes-Magalhaes et al., 2023; Ruiz-Padial et al., 2021). Third, we evaluated emotion recognition using two rates: "Raw hit rate" and "Unbiased hit rate". "Raw hit rate" assessed emotion categorization for each emotion, measuring the number of correct categorizations (i.e., when participants accurately selected the emotion genuinely expressed by the face or body) divided by the total number of stimuli for that emotion (van der Schalk et al., 2011). We also measured the "Unbiased hit rates" (Hu score; (Wagner, 1993) for each facial and bodily emotion. This score accounts for potential bias in the "Raw hit rate" due to participants over-attributing their emotional choices. For instance, if a participant used "anger" to categorize an angry face but also used "anger" for other emotions like happiness or disgust, this would lead to a high "Raw hit rate" of anger. However, the Hu score counters this bias by calculating the squared frequency of correct responses for a target emotion, divided by the product of the number of stimuli embodying that emotion and the overall frequency of choosing that emotion category (Wagner, 1993).

Statistical Analysis

Data were analyzed using JASP (version 0.18.1; RRID:SCR_015823), R (version 4.3.0), and the package ggplot2 (version 3.4.2; RRID:SCR_014601; (Wickham, 2016), tidyverse (version 1.2.0; RRID:SCR_019186; (Wickham et al., 2019), dplyr (version 1.1.2;

RRID :SCR_017102; (Wickham et al., 2023), emmeans (version 1.8.7; RRID :SCR_018734; (Russell et al., 2021). We applied the R function MAD (“Mean Absolute Deviation”; $b = 1.4826$ and $\text{threshold} = 3$) to detect and remove outlier data points and participants across all tasks (Leys et al., 2013, 2019).

To determine if some emotions were recognized more accurately than others, we examined whether there were significant differences in the recognition of each emotion (for both faces and bodies). To do that, we conducted a 7 (emotion) x 2 (gender of raters) x 2 (gender of actors) Mixed ANOVA on the arcsine transformation of the Hu score of face and body postures as recommended (Wagner, 1993) and already done in previous works (van der Schalk et al., 2011). The mixed ANOVA was performed with the Huynh-Feldt correction, as the sphericity condition was not respected (Haverkamp & Beauducel, 2017).

Finally, we aimed to identify instances where emotions were confused with one another, i.e. if a target emotion (e.g., anger) was systematically confused with another no-target emotion (e.g., disgust). To achieve this, we generated confusion matrices and performed repeated measure ANOVA on incorrect responses in categorizing each target emotion (for both facial and bodily expressions).

Results

Descriptive Statistics

The MAD function was applied to recognition task data, revealing one participant as an outlier for accuracy data and two for Hu score data. These participants were therefore removed from the analysis.

A total of 1952 facial stimuli and 280 body stimuli² have been subjected to validity measures. The dataset for facial emotions consisted of 37088 data (76 participants, 488 facial emotion stimuli per participant) with 425 missing values, while the dataset for body postures comprised 5320 data (76 participants, 70 body posture stimuli per participant) with 70 missing values. The means of accuracy, representing the proportion of emotions accurately categorized, was 0.75 (SD = 0.43) for facial stimuli and 0.68 (SD = 0.47) for body stimuli. Fleiss's Generalized kappa scores (Fleiss, 1971) were calculated for both facial and bodily stimuli across the four participant groups. The mean Fleiss's Generalized kappa score was 0.63 for facial stimuli and 0.53 for body ones. These scores indicate a substantial level of agreement among participants in the categorization of facial emotions and a moderate level of agreement for body postures across the seven emotions (Landis & Koch, 1977). Then, Bonferroni-adjusted *t*-tests revealed that facial emotions were recognized more accurately than body postures, with "Raw hit rates" ($t[74] = 6.066, p < 0.001, d = 0.70, CI [0.043, 0.085]$) and arcsine-transformed Hu scores ($t[73] = 3.555, p < 0.001, d = 0.413, CI [0.028, 0.1]$).

Face Stimuli

Reliability

We measured the internal consistency of participants' ratings by calculating split-half reliability scores. Pearson's correlations were significant ($ps < 0.001$) for all subgroups (even and odd-numbered participant subgroups, and gender subgroups) within the four stimulus groups. Additionally, we computed the Spearman-Brown coefficient for each comparison between even and odd-numbered participant subgroups. The Spearman-Brown coefficients were high for all four stimulus groups (Group1, $r = 0.810$; Group2, $r = 0.764$; Group3, $r = 0.841$; Group4,

² Body stimuli were created by taking the body expression of only congruent stimuli.

$r = 0.815$) within each subgroup. Finally, the Spearman-Brown coefficients for men and women subgroups were also high across all four stimulus groups (Group1, $r = 0.769$; Group2, $r = 0.788$; Group3, $r = 0.817$; Group4, $r = 0.832$).

“Raw Hit Rates” and “Unbiased Hit Rates”

Emotion recognition rates, as assessed by “Raw hit rates” and “Unbiased hit rates” (the Hu score), are presented in Tables 1 and 2. These rates are high for all emotions, apart from the “Unbiased hit rate” for sadness (0.48) and fear (0.34). For each emotion, both “Raw hit rates” and “Unbiased hit rates” scores are significantly above the chance level (see Tables 1 and 2; 1/7 response options; $\mu = 0.14$) and are three times above the chance level.

Then, when considering the Hu scores, we found that facial emotions of happiness are the ones with a better Hu score (“unbiased hit rate”; 0.91) while fear displays the lowest score (0.34; see Table 1). In terms of “Raw hit rate”, happiness (0.96), surprise (0.81), and neutrality (0.83) are emotions with the highest accuracy, while fear (0.49) records the lowest rate.

Furthermore, the Standard deviation (SD) of “Raw hit rate” and “Unbiased hit rate” is low for all emotions (see Table 1).

Taken together, these findings suggest that facial emotions are recognized with a high degree of accuracy and are rarely confused with each other.

Table 1. Summary Table of “Raw Hit Rate” and “Unbiased Hit Rates” for Facial Emotions

	“Raw Hit Rate” (+/- SD)	“Unbiased Hit Rate” (+/-SD)
Anger	0.71 (+/- 0.45) ***	0.56 (+/- 0.16) ***
Disgust	0.76 (+/- 0.43) ***	0.59 (+/- 0.14) ***
Happiness	0.96 (+/- 0.19) ***	0.91 (+/- 0.07) ***
Neutrality	0.83 (+/- 0.38) ***	0.66 (+/- 0.12) ***
Fear	0.49 (+/- 0.5) ***	0.34 (+/- 0.16) ***
Surprise	0.81 (+/- 0.39) ***	0.52 (+/- 0.13) ***
Sadness	0.68 (+/- 0.47) ***	0.48 (+/- 0.14) ***
Mean	0.75 (+/- 0.43)	0.58 (+/- 0.21)

Note. Table contains Mean of “Raw hit rate” (+/- Standard deviation), and Hu score/Unbiased hit rates (+/- Standard deviation) for each facial emotion. The asterisk (***) indicates that the “Raw hit rate” and “Hu score” significance at $ps < 0.001$ *** above chance level (0.14).

Table 2. *Confusion Matrix for the “Raw Hit Rate” of Facial Emotions*

		RESPONSE							
		Anger	Disgust	Happiness	Neutrality	Fear	Surprise	Sadness	Total
CORRECT	Anger	3705 (71.29)***	450 (8.66)	8 (0.15)	354 (6.81)	135 (2.60)	139 (2.67)	406 (7.81)	5197
	Disgust	366 (7.1)	3924 (76.15)***	23 (0.45)	24 (0.47)	199 (3.86)	93 (1.80)	524 (10.17)	5153
	Happiness	8 (0.15)	14 (0.27)	4999 (96.43)***	33 (0.64)	22 (0.42)	103 (1.99)	5 (0.1)	5184
	Neutrality	269 (5.23)	67 (1.30)	47 (0.91)	4248 (82.52)***	51 (0.99)	74 (1.44)	392 (7.61)	5148
	Fear	96 (1.85)	292 (5.62)	15 (0.29)	138 (2.66)	2568 (49.46)***	1910 (36.79)	173 (3.33)	5192
	Surprise	17 (0.33)	63 (1.22)	151 (2.92)	64 (1.24)	663 (12.82)	4175 (80.71)***	40 (0.77)	5173
	Sadness	361 (7)	408 (7.91)	71 (1.38)	565 (10.96)	162 (3.14)	104 (2.02)	3486 (67.60)***	5157

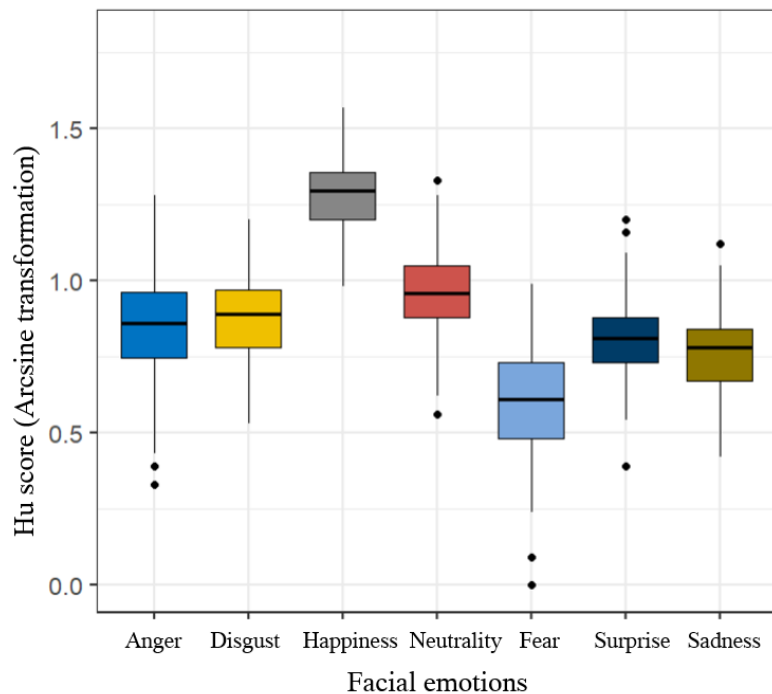
Note. The **bold** numbers represent the number (the “Raw hit rate” (%)) of categorizations for each of facial emotion and the other numbers represent categorization errors between each emotion. The asterisk (***) indicates that the “Raw hit rate” significance at $ps < 0.001$ *** above chance level (0.14). Confusion matrix was calculated by removing the empty responses.

Recognition of Facial Basic Emotions: Emotion and Gender Difference

A 7 (emotion) x 2 (gender of raters) x 2 (gender of actors) Mixed ANOVA on arcsine-transformed Hu score revealed a significant main effect of emotion ($F[4.345, 312.818] = 377.887, p < 0.001, \eta^2 = 0.629$), main effect of gender of actors ($F[1, 72] = 101.3, p < 0.001, \eta^2 = 0.044$), and an interaction effect between emotion and gender of actors ($F[4.944, 355.951] = 2.270, p = 0.048, \eta^2 = 0.002$). No other result was significant.

Facial expressions of happiness were recognized the most accurately, followed by neutrality, disgust, anger, surprise, sadness, and fear. Bonferroni-adjusted paired *t*-tests revealed a significant difference between all emotions ($ps < 0.001$) except for no significant difference between disgust and anger ($t[73] = -1.961, p = 1, d = -0.22, CI [-0.077, 0.017]$) and between anger and surprise ($t[73] = 2.037, p = 0.888, d = 0.229, CI [-0.016, 0.079]$). The analysis also showed that facial emotions expressed by female actors are better recognized than those expressed by male actors ($t[73] = 10.065, p < 0.001, d = 0.752, CI [0.083, 0.124]$). No other result was significant (see Figure 5).

Figure 5. Boxplot of the Arcsine Transformation of the Hu Score by Facial Emotions



Note. Hu scores for face emotions (anger, disgust, happiness, neutral, fear, surprise and sad). Significant difference between emotions ($ps < 0.001^{***}$) except between anger and disgust ($p = 1$) and between anger and surprise ($p = 0.888$).

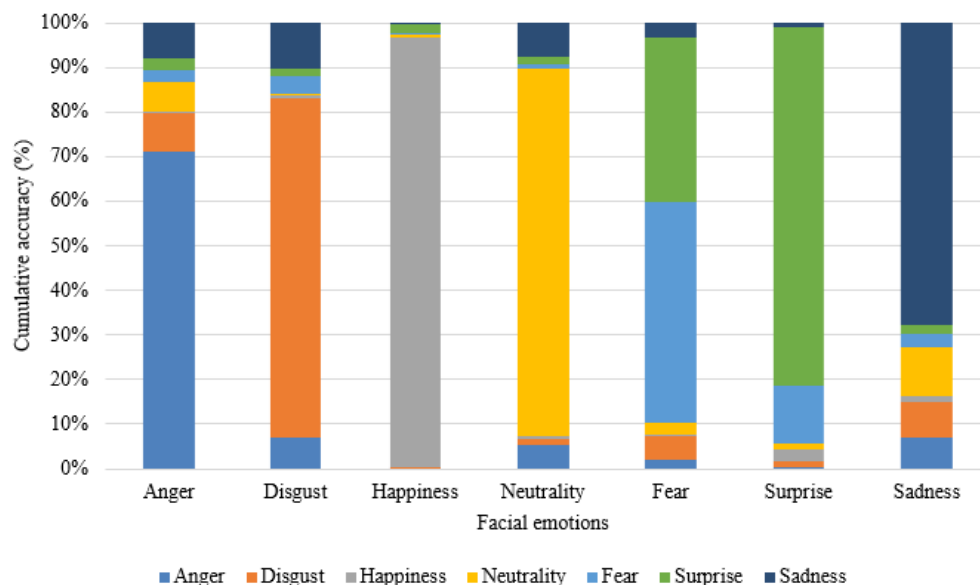
Confusion Between Facial Basic Emotions

We conducted repeated measure ANOVA on “Raw Hit Raw” to examine categorization errors within each emotion, aiming to determine whether the target emotion was confused with another emotion. For instance, concerning happiness stimuli, we assessed whether participants tended to categorize them as surprise significantly more often than as sadness or anger.

These analyses revealed that anger was significantly more frequently miscategorized as disgust, neutrality, and sadness than as other emotions ($ps < 0.001$). Similarly, disgust was more often confused with sadness than with other emotions ($ps < 0.001$). Happiness and fear were significantly more frequently miscategorized as surprise than as other emotions ($ps < 0.001$). In contrast, surprise was significantly more frequently confounded with fear than with other emotions ($ps < 0.001$), and sadness was more often confused with neutrality, anger, and disgust

($ps < 0.001$) than with other emotions ($ps < 0.001$). Finally, neutrality was more often confused with sadness and anger than with other emotions ($ps < 0.001$; see Figure 6).

Figure 6. *Cumulative of Accuracy (%) for Each Facial Emotion*



Note. Cumulative of accuracy (%) for each facial emotion. For each facial emotion, the graph gives the percentage (%) of response choice. We can see that certain emotions are mainly confused with each other, such as the facial emotion of fear (blue) with surprise (green). We can also see that for each facial emotion, the emotional target is mainly chosen in the response choices.

Body Stimuli

Reliability

Similarly to the facial stimuli, we assessed the internal consistency of participants' ratings by calculating split-half reliability scores. Pearson's correlations were significant ($ps < 0.001$) for all subgroups (even and odd-numbered participant subgroups and gender subgroups) within the four stimulus groups. Additionally, the Spearman-Brown coefficients were high for all four stimulus groups within each subgroup (*even and odd subgroup*: Group1, $r = 0.781$; Group2, $r = 0.755$; Group3, $r = 0.764$; Group4, $r = 0.685$; *Gender subgroup*: Group1, $r = 0.851$; Group2, $r = 0.811$; Group3, $r = 0.843$; Group4, $r = 0.704$).

“Raw Hit Rate” and “Unbiased Hit Rate”

Emotion recognition rates, as assessed by the “Raw hit rate” and the “Unbiased hit rate” are presented in Tables 3 and 4. For each emotion, both the “Raw hit rate” and the “Unbiased hit rate” scores are significantly above three times above the chance level (see Tables 3 and 4; 1/7 response options; $\mu = 0.14$) except for the “Raw hit rate” of disgust emotion (0.39) and the “Unbiased hit rates” for disgust (0.26), fear (0.35) and surprise (0.33).

Upon considering the Hu scores, we observed that neutral body posture exhibits the highest Hu score (the “Unbiased hit rate”; 0.73), while disgust displays the lowest score (0.23; see Figure 6). In terms of “Raw hit rate”, neutrality (0.97), anger (0.87), and happiness (0.73) in body postures show the highest “Raw hit rate”, whereas disgust (0.39) exhibits the lowest “Raw hit rate”.

Lower rates in both the “Raw hit rate” and the “Unbiased hit rate” may indicate that bodily postures such as disgust, surprise, and fear are recognized less accurately and are more frequently confused with other emotions. Nevertheless, these rates are still above the chance level (0.14).

Lastly, the Standard deviation (SD) of the “Raw hit rate” and the “Unbiased hit rate” is the lowest of all emotions (see Table 3).

Table 3. Summary Table of “Raw Hit Rate” and “Unbiased Hit Rates” for Body Postures

	“Raw Hit Rate” (+/- SD)	“Unbiased Hit Rate” (+/-SD)
Anger	0.87 (+/- 0.32) ***	0.71 (+/- 0.20) ***
Disgust	0.39 (+/- 0.49) ***	0.26 (+/- 0.22) ***
Happiness	0.73 (+/- 0.45) ***	0.59 (+/- 0.23) ***
Neutrality	0.97 (+/- 0.16) ***	0.75 (+/- 0.16) ***
Fear	0.66 (+/- 0.47) ***	0.35 (+/- 0.19) ***
Surprise	0.52 (+/- 0.5) ***	0.33 (+/- 0.25) ***
Sadness	0.61 (+/- 0.49) ***	0.54 (+/- 0.26) ***
Mean	0.673 (+/- 0.46)	0.504 (+/- 0.28)

Note. Table contains mean of “Raw hit rate” (+/- Standard deviation), and Hu score (+/- Standard deviation) of participants for each body posture. The asterisk (***) indicates that the “Raw hit rate” and “Hu score” significance at $ps < 0.001$ *** above chance level (0.14).

Table 4. *Confusion Matrix for the “Raw Hit Rate” of Body Postures*

		RESPONSE							Total
		Anger	Disgust	Happiness	Neutrality	Fear	Surprise	Sadness	
CORRECT	Anger	654 (88.38)***	11 (1.49)	30 (4.05)	28 (3.78)	3 (0.41)	7 (0.95)	7 (0.95)	740
	Disgust	25 (3.38)	294 (39.78)***	4 (0.54)	13 (1.76)	317 (42.90)	65 (8.80)	21 (2.84)	739
	Happiness	48 (6.43)	4 (0.54)	544 (72.92)***	2 (0.26)	23 (3.08)	125 (16.76)	0 (0)	746
	Neutrality	1 (0.13)	3 (0.40)	4 (0.54)	728 (97.46)***	0 (0)	1 (0.13)	10 (1.34)	747
	Fear	6 (0.81)	130 (17.52)	5 (0.67)	3 (0.40)	491 (66.17)***	94 (12.67)	13 (1.75)	742
	Surprise	87 (11.89)	35 (4.78)	101 (13.80)	47 (6.42)	52 (7.10)	387 (52.87)***	23 (3.14)	732
	Sadness	8 (1.09)	16 (2.17)	14 (1.90)	154 (20.92)	82 (11.14)	8 (1.09)	454 (61.68)***	736

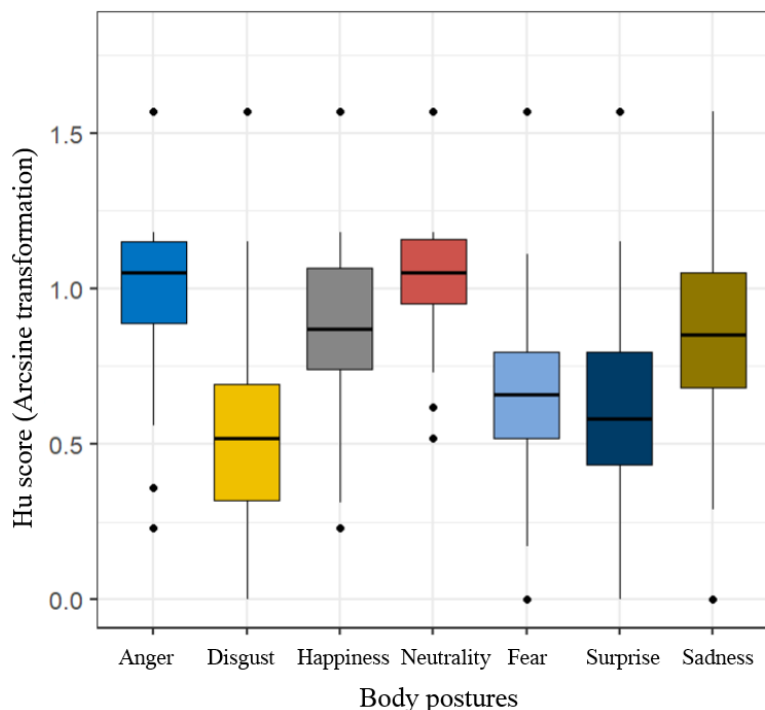
Note. The **bold** numbers represent the number (the “Raw hit rate” (%)) of categorizations for each of body posture and the other numbers represent categorization errors between each emotion. The asterisk (***) indicates that the “Raw hit rate” significance at $ps < 0.001$ *** above chance level (0.14). Confusion matrix was calculated by removing the empty responses.

Recognition of Basic Body Postures: Emotion and Gender Difference

A 7 (emotion) x 2 (gender of raters) x 2 (gender of actors) Mixed ANOVA on the arcsine transformation of the Hu score revealed a significant main effect of emotion ($F[5.033, 362.359] = 96.687, p < 0.001, \eta^2 = 0.361$), an interaction effect between emotion and gender of raters ($F[5.033, 362.359] = 2.481, p = 0.023, \eta^2 = 0.009$), and an interaction effect between emotion and gender of actors ($F[4.850, 349.193] = 96.687, p = 0.002, \eta^2 = 0.01$). No other result was significant.

Neutral body postures are recognized the most accurately, followed by anger, happiness, sadness, fear, surprise, and disgust. Bonferroni-adjusted paired *t*-tests revealed a significant difference between all emotions ($ps < 0.001$) except for no significant difference between neutrality and anger ($t[73] = -0.113, p = 1, d = -0.014, CI [-0.110, 0.102]$) and between disgust and surprise ($t[73] = -2.880, p = 0.088, d = -0.348, CI [-0.207, 0.006]$), between fear and surprise ($t[73] = 1.205, p = 1, d = 0.146, CI [-0.064, 0.148]$) and between happiness and sadness ($t[73] = 1.563, p = 1, d = 0.189, CI [-0.052, 0.161]$). No other result was significant (see Figure 7).

Figure 7. *Boxplot of the Arcsine Transformation of the Hu Score by Body Postures*

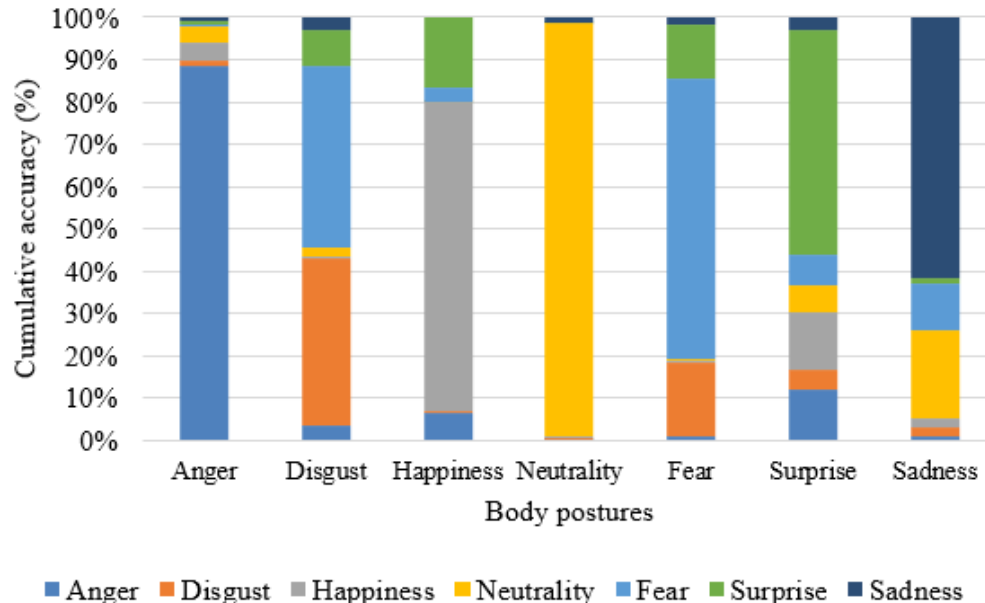


Note. Hu scores for body postures (anger, disgust, happiness, neutrality, fear, surprise, and sadness). Significant difference between emotions ($p < 0.001^{***}$) except between neutrality and anger ($p = 1$) and between disgust and surprise ($p = 0.088$), between fear and surprise ($p = 1$) and between happiness and sadness ($p = 1$).

Confusion Between Basic Body Postures

We conducted repeated measure ANOVA on “Raw hit raw” to examine categorization errors within each emotion, allowing us to determine if the target emotion is consistently confused with another emotion.

These analyses revealed that disgust was more often miscategorized as fear than as the other emotions ($p < 0.001$) and happiness was significantly more frequently miscategorized as surprise than as the other emotions ($p < 0.001$). Lastly, fear was significantly miscategorized more often as surprise and disgust than as other emotions ($p < 0.001$) while sadness was more frequently confused with neutrality than with other emotions ($p < 0.001$; see Figure 8).

Figure 8. *Cumulative of Accuracy (%) for Each Body Postures*

Note. For each body posture, the graph gives the percentage (%) in response choice. We can see that certain emotions are mainly confused with each other, such as the body posture of disgust with fear (light blue). We can also see that for each body posture, the emotional target is mainly chosen in the response choices.

Discussion

The mean “Raw hit rates” of 0.75 and “Unbiased hit rates” of 0.58 for facial basic emotion recognition, as well as 0.67 and 0.50 for basic body posture recognition, align with those reported in other facial emotion stimulus sets (Goeleven et al., 2008; Tottenham et al., 2009; van der Schalk et al., 2011) and body posture stimulus sets (de Gelder & van den Stock, 2011; Lopez et al., 2017; Thoma et al., 2013).

Regarding the recognition task, happiness is the most recognized facial emotion, followed by neutrality, disgust, anger, surprise, sadness, and fear. Concerning body postures, the most easily recognized emotion was a neutral body posture, followed by anger, happiness, sadness, fear, surprise, and disgust. These findings indicate a high level of recognition for the facial expression of happiness and poor recognition of fearful facial and bodily expressions, consistent

with prior research (de Gelder & van den Stock, 2011; Goeleven et al., 2008; Thoma et al., 2013; Tottenham et al., 2009) which attributes this to the familiarity of happiness compared to fear (Calvo et al., 2014). In everyday life, we are more likely to express and perceive emotions of happiness than of fear, making happiness more familiar than fear. Moreover, facial expressions of fear are expressed differently depending on the context (van den Stock et al., 2007). For example, an individual will not express fear in the same way when in a dark alley as when facing an aggressive animal. In contrast, happiness is expressed relatively similarly regardless of the context in which it is expressed. This variability in the expression of fear could explain the greater difficulty in recognising it compared to happiness, which is expressed more uniformly. Lastly, the lower recognition of the facial expressions of fear and sadness may also be related to the difference between voluntary and spontaneous emotional expressions (Zuckerman et al., 1976). It is possible that the actors encountered difficulties in expressing fear or sadness in a low-ecological, and controlled context. However, we believe that the use of scenarios should have minimized this effect (de Gelder & van den Stock, 2011).

Furthermore, disgust poses the greatest challenge in body posture recognition, while anger is consistently well recognized as previously demonstrated (Atkinson et al., 2004; Coulson, 2004; de Gelder & van den Stock, 2011; Lopez et al., 2017; Thoma et al., 2013). The high recognition of anger body postures could be attributed to the consistent portrayal of anger in our dataset, where actors predominantly express anger by clenching their fists and tensing their muscles, potentially making bodily expressions more easily recognizable.

The confusion between facial and body expressions of basic emotions found in our results has been well-documented in the literature (de Gelder & van den Stock, 2011; Lopez et al., 2017), particularly the confusion of facial expressions of fear and surprise (Jack et al., 2009; Matsumoto et al., 2011; Roy-Charland et al., 2014). Previous works have referred to the emotion

seed model (Susskind et al., 2007), and Facial Action Coding Systems (FACS; (Ekman & Friesen, 1976) to explain this confusion through the similarities in the numerous physical characteristics (Aviezer, Hassin, Bentin et al., 2008; Susskind et al., 2007) and muscle movements (Ekman & Friesen, 1976) between both facial expressions. Indeed, surprise and fear involve, among others, the raising of the inner and outer eyebrows, and the lifting of the eyelids. The activation of the same muscular movements leads to greater confusion in distinguishing between these emotions (Ekman & Friesen, 1976; Roy-Charland et al., 2014). On the contrary, certain emotions such as happiness are rarely confused with other emotions because they activate muscles that are specific to that emotion (e.g., happiness) and not to other emotions (Roy-Charland et al., 2014). Finally, the confusion between body postures of disgust and fear was also found in a previous study (Lopez et al., 2017) to be explained by the fact that these two emotions share the same action tendencies, involving the intention to protect oneself or move away from danger by adopting a defensive posture with outstretched arms (Frijda et al., 1989).

Finally, facial expressions expressed by female actors are better recognized than those expressed by males, indicating a gender effect of the actors on emotional recognition (Wells et al., 2016). This coincides with previous studies showing that women are better than men at communicating and expressing their emotions (Gregoric et al., 2014; Kret & de Gelder, 2012; LaFrance et al., 2003; McDuff et al., 2017; Wallbott, 1988). Many interpretations have been previously proposed to explain these differences in emotional expression between men and women, such as differences in brain structures (Cosgrove et al., 2007; Kret & de Gelder, 2012) or gender stereotypes (Wallbott, 1988). Another possible interpretation is that women are physically more expressive than men (Gregoric et al., 2014; LaFrance et al., 2003). Physically, when expressing their facial emotions, women express more facial action units (AUs), especially during positive valence emotions (McDuff et al., 2017), and contract muscles such as the

zygomaticus more intensely (Gregoric et al., 2014). This interpretation aligns with the hypothesis (see Part 2) that intensely expressed emotions are more easily recognizable than weakly intense emotions due to stronger muscle contractions related to emotional expression (Wingenbach et al., 2016).

Validation of Stimuli: Part 2

The main objective of this second part of the validation was to evaluate supplementary features (authenticity, arousal, intensity, and valence) expanding the perspectives of use for the stimulus set and strengthening the observations of the first part with more participants.

Methods

Participants

Forty-two participants (21 females; $M_{\text{age}} = 25.9$ years; ± 4.51 years; Range: 18-34 years) were recruited from February to April 2024 through social media postings. Each participant received CAD 60 as compensation for three one-hour sessions. The inclusion and exclusion criteria for this second part of the validation were the same as those for the first part. Two participants did not return after the first session and were thus excluded from the study. The total number of participants included in the study was 40 participants (20 females; $M_{\text{age}} = 26$ years; ± 4.59 years; Range: 18-34 years). Like the first validation (Part 1), the sample size was determined based on prior studies of similar stimulus creation (de Gelder & van den Stock, 2011; Thoma et al., 2013). Regarding the HADS scale (Zigmond & Snaith, 1983), no participants were excluded from the study based on depression and anxiety scores ($M_{\text{depression}} = 3.64 \pm 2.9$; $M_{\text{anxiety}} = 6.05 \pm 2.55$). These 40 participants were randomly assigned to two groups consisting of 20 participants each (Group 1: 11 females; Group 2: 9 females). To obtain further information on participants' demographic details, supplementary materials are available on the Open Science Framework (OSF; <https://osf.io/frvwt/>).

Materials

Facial and body expressions of basic emotions described in the first stimuli validation (Part 1) were also utilized in this second part of the validation (Part 2).

Procedure

For this second part of the validation, the experiment took place in the same location and utilized the same equipment (e.g., E-Prime 2.0) as the first validation (Part 1). The 40 participants were randomly divided into two groups (twenty participants/group). Each group evaluated 1116 stimuli (976 facial expressions and 140 body postures) distributed randomly across three one-hour sessions. Each session consisted of between 371 and 373 stimuli, including 46 body stimuli and 325 to 327 face stimuli. Before the initial session, participants completed socio-demographic information, consent forms, and HADS questionnaires (Zigmond & Snaith, 1983) on the LimeSurvey platform.

The three sessions were identical and spaced one week apart (e.g., every Monday for three weeks). The experiment began with a six-trial training session to familiarize participants with the task. Subsequently, participants completed, for each stimulus, a recognition task followed by rating scales assessing valence, arousal, intensity, and authenticity. The order of the four scales and the stimuli were counterbalanced across participants. The recognition task was identical to that of the first part of the validation except there was no time limit. As for the four scales, stimuli were centered on the screen (49% x 22% for facial images and 75% x 40% for body images), with questions and Likert scales (Courier New, Size 16) appearing below the image. Participants were instructed to (a) respond to a question about intensity (“How intense is this emotional expression?”; “Quelle est l’intensité de cette expression émotionnelle?”) on a Likert scale ranging from 0 (*none*), 1 (*weak*) to 9 (*strong*) (Dawel et al., 2017), (b) rate valence (“How positive or negative does this emotional expression seem to you?”; “À quel point votre ressenti

est positif ou négatif en voyant cette image?”), on a Likert scale from 1 (*very negative*) to 9 (*very positive*), (c) rate arousal (“How intense does this emotional expression make you feel?”; “Quelle est la force de votre ressenti en voyant cette image?”), on a Likert scale from 1 (*calm*) to 9 (*aroused*) (Goeleven et al., 2008; Lassalle et al., 2019; O'Reilly et al., 2016), and (d) rate authenticity (“How authentic does this emotional expression seem to you?”; “Dans quelle mesure cette expression émotionnelle vous semble-t-elle authentique?”), on a Likert scale from -7 (*completely false*) à +7 (*completely genuine*) including 0 (*I don't know*) (Dawel et al., 2017; Miolla et al., 2023).

Participants were instructed to respond using the mouse as quickly as possible based on their first impression. The next trial appeared once the participant made their selection. Multiple breaks (every 30 stimuli) were provided during the task to alleviate fatigue effects.

Measures and Statistical Analyses

The data resulting from the recognition task were analyzed in the same manner as the first part of validation, except with the addition of session as a variable in the Mixed ANOVA (7 (emotion) x 2 (gender of raters) x 2 (gender of actors) x 3 (session) Mixed ANOVA) on the arcsine transformation of the Hu score of face expression and body postures. The addition of the session variable allows us to observe the difference between the three sessions.

For arousal, valence, authenticity, and intensity features, the split-half reliability scores on accuracy data were also used to measure the internal consistency of participants' ratings. We also performed a Pearson correlation analysis to explore any associations between these features (Fernandes-Magalhaes et al., 2023). Finally, we computed the Mixed ANOVA (7 (emotion) x 2 (gender of raters) x 2 (gender of actors) x 3 (session)) on each feature (arousal, valence, authenticity, and intensity).

Results

Descriptive Statistics

The MAD function was applied to the entire dataset of each task, resulting in 2.67% of data points identified as outliers for the authenticity task, 2.40% for the intensity task, 0.04% for the valence task, and 0.66% for the arousal task. Similarly, outlier detection was performed across all participants, revealing one participant for the recognition task, one participant for the authenticity and the intensity task, another one participant for the valence task, and three for the arousal task. These participants were therefore removed from the analysis.

The mean accuracy was 0.72 (SD = 0.2) for facial stimuli and 0.63 (SD = 0.32) for body ones. The mean Fleiss's Generalized kappa score (Fleiss, 1971) was 0.63 for facial stimuli, and 0.56 for body ones. These scores indicate a substantial level of agreement among participants in categorizing facial emotions and a moderate level of agreement for body postures across the seven emotions (Landis & Koch, 1977). Subsequently, Bonferroni-adjusted *t*-tests revealed that facial emotions were recognized more accurately than body postures, with "Raw hit rates" ($t[38] = -10.251, p < .001, d = -1.642, CI [-0.110, -0.074]$). However, regarding arcsine-transformed Hu scores, no significant difference was found ($t[38] = -0.048, p = 0.962, d = -0.008, CI [-0.029, 0.028]$).

Facial Stimuli

Reliability

All Pearson's correlations were significant ($ps < 0.001$). The Spearman-Brown coefficients for split-half reliability, both subgroups by even odd participant numbers (Group1, $r = 0.794$; Group2, $r = 0.811$) and by gender (Group1, $r = 0.823$; Group2, $r = 0.857$), were high for two stimulus groups. In ratings of intensity, authenticity, valence, and arousal, the Spearman-Brown coefficients for split-half reliability, again both subgroups by even and odd participant

numbers (*Intensity*: Group1, $r = 0.895$; Group2, $r = 0.874$; *Valence*: Group1, $r = 0.820$; Group2, $r = 0.896$; *Arousal*: Group1, $r = 0.589$; Group2, $r = 0.629$; *Authenticity*: Group1, $r = 0.624$; Group2, $r = 0.702$) and by gender (*Intensity*: Group1, $r = 0.878$; Group2, $r = 0.880$; *Valence*: Group1, $r = 0.928$; Group2, $r = 0.908$; *Arousal*: Group1, $r = 0.928$; Group2, $r = 0.908$; *Authenticity*: Group1, $r = 0.632$; Group2, $r = 0.758$) were high for two stimulus groups.

“Raw Hit Rates” and “Unbiased Hit Rates”

Descriptive data in Tables 5 and 6 show that both “Raw hit rates” and “Unbiased hit rates” scores are significantly above the chance level (see Tables 5 and 6; 1/7 response options; $\mu = 0.14$) and are three times above the chance level for all emotion except for the fear (see Tables 5 et 6). Considering the Hu scores (“Unbiased hit rate”), facial emotions of happiness have the highest Hu score (0.87), while fear displays the lowest Hu score (0.29). Taken together, these findings suggest that facial emotions are recognized with a high degree of accuracy and are rarely confused with each other.

Table 5. Summary Table of “Raw Hit Rate” and “Unbiased Hit Rates” for Facial Emotions

	“Raw Hit Rate” (+/- SD)	“Unbiased Hit Rate” (+/-SD)
Anger	0.68 (+/-0.47) ***	0.52 (+/- 0.11) ***
Disgust	0.76 (+/-0.43) ***	0.58 (+/-0.14) ***
Happiness	0.96 (+/-0.20) ***	0.87 (+/-0.08) ***
Neutrality	0.76 (+/-0.43) ***	0.58 (+/-0.15) ***
Fear	0.39 (+/-0.49) ***	0.29 (+/-0.14) ***
Surprise	0.85 (+/-0.36) ***	0.52 (+/-0.08) ***
Sadness	0.67 (+/- 0.47) ***	0.45 (+/-0.10) ***
Mean	0.72 (+/- 0.45)	0.54 (+/- 0.21)

Note. Table contains Mean of “Raw hit rate” (+/- Standard deviation), and Hu score/Unbiased hit rates (+/- Standard deviation) for each facial emotion. The asterisk (***) indicates that the “Raw hit rate” and “Hu score” significance at $ps < 0.001$ *** above chance level (0.14).

Table 6. Confusion Matrix for the “Raw Hit Rate” of Facial Emotions

		RESPONSE							Total
		Anger	Disgust	Happiness	Neutrality	Fear	Surprise	Sadness	
CORRECT	Anger	3733 (68.37)***	469 (8.59)	7 (0.13)	465 (8.52)	92 (1.68)	209 (3.83)	485 (8.88)	5692
	Disgust	357 (6.61)	4118 (76.23)***	40 (0.74)	32 (0.59)	156 (2.89)	112 (2.07)	587 (10.87)	5632
	Happiness	4 (0.07)	17 (0.31)	5218 (95.92)***	97 (1.78)	13 (0.24)	87 (1.6)	4 (0.07)	5672
	Neutrality	430 (7.90)	57 (1.05)	117 (2.15)	4130 (75.91)***	47 (0.86)	91 (1.67)	569 (10.46)	5672
	Fear	165 (3.02)	368 (6.74)	20 (0.37)	136 (2.49)	2115 (38.74)***	2446 (44.80)	210 (3.85)	5694
	Surprise	27 (0.5)	86 (1.59)	238 (4.39)	87 (1.61)	368 (6.79)	4580 (84.50)***	34 (0.63)	5652
	Sadness	346 (6.36)	420 (7.72)	117 (2.15)	619 (11.38)	112 (2.06)	159 (2.92)	3668 (67.41)***	5674

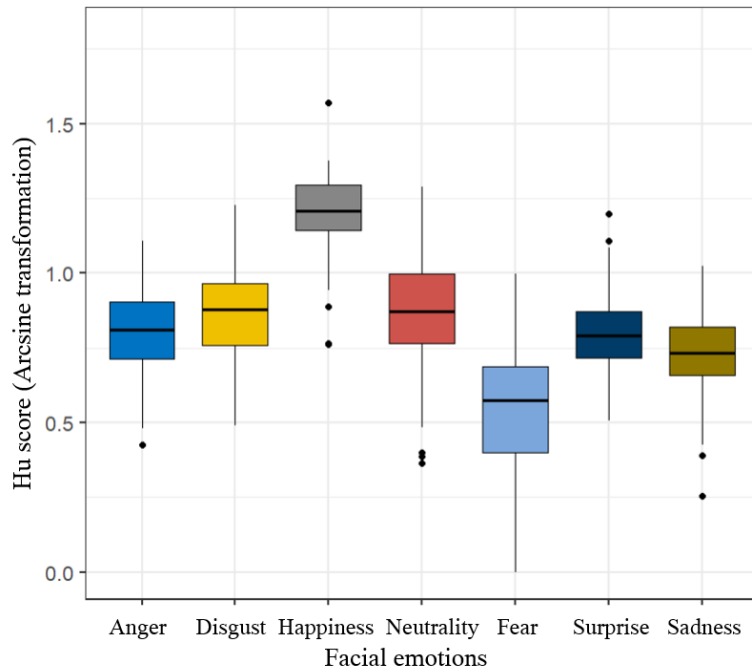
Note. The **bold** numbers represent the number (the “Raw hit rate” (%)) of categorizations for each of facial emotion and the other numbers represent categorization errors between each emotion. The asterisk (***) indicates that the “Raw hit rate” significance at $ps < 0.001$ *** above chance level (0.14). Confusion matrix was calculated by removing the empty responses.

Recognition of Facial Basic Emotions: Emotion and Gender Difference

A 7 (emotion) x 2 (genre of actors) x 2 (genre of raters) x 3 (session) Mixed ANOVA on the arcsine transformation of the Hu score revealed a significant main effect of emotion ($F[4.09, 151.348] = 227.865, p < 0.001, \eta^2 = 0.618$), genre of actors ($F[1, 37] = 164.502, p < 0.001, \eta^2 = 0.042$), and session ($F[2, 74] = 3.728, p = 0.029, \eta^2 = 0.002$). The analysis also showed an interaction effect between gender of actors and emotion ($F[6, 222] = 7.064, p < .001, \eta^2 = 0.003$) and an interaction between genre of actors, gender of raters and session ($F[2, 74] = 3.476, p = 0.036, \eta^2 = 0.0009$). No other results were significant.

Happiness was recognized the most accurately, followed by neutrality, disgust, anger, surprise, sadness, and fear. Bonferroni-adjusted paired t -tests revealed a significant difference between all emotions ($ps < 0.001$), except for surprise and anger ($t[38] = 0.136, p = 1, d = 0.019, CI [-0.057, 0.062]$) and for disgust and neutrality ($t[38] = -0.246, p = 1, d = -0.035, CI [-0.064, 0.055]$; see Figure 9). Additionally, Bonferroni-adjusted paired t -tests showed that all emotions expressed by female actors were recognized better than those expressed by male actors ($t[38] = 12.826, p < 0.001, d = 0.727, CI [0.084, 0.116]$), and that the recognition was better during the third session rather than the first ($t[38] = -2.721, p = 0.024, d = -0.213, CI [-0.056, -0.003]$). Finally, Bonferroni-adjusted paired t -tests indicated that female raters recognized female actors better than male raters during the second session ($t[38] = 4.429, p < 0.001, d = 0.770, CI [0.023, 0.188]$). No other result was significant (see Figure 9).

Figure 9. Boxplot of the Arcsine Transformation of the Hu Score by Facial Expression

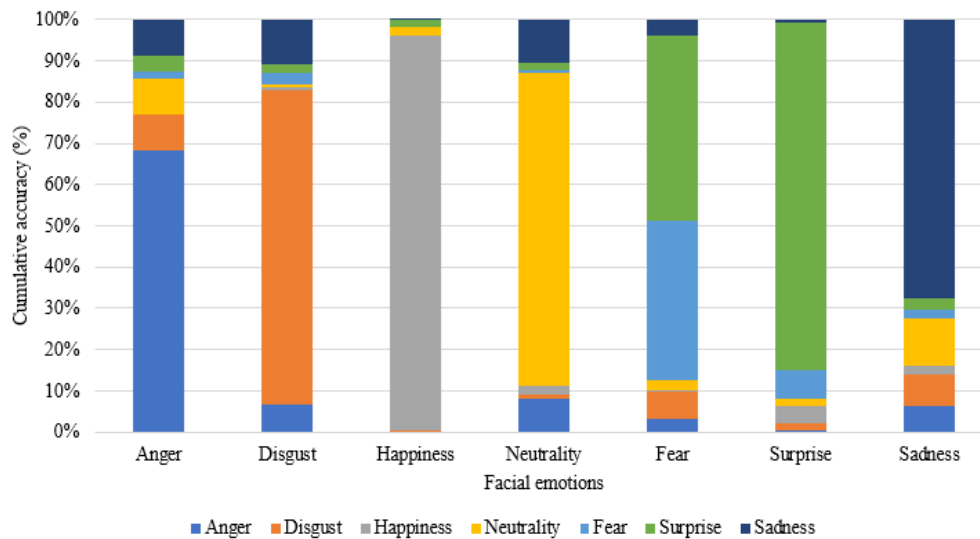


Note. Hu scores for facial emotions (anger, disgust, happiness, neutrality, fear, surprise, and sadness).

Confusion Between Facial Basic Emotions

Regarding repeated measure ANOVA on “Raw hit raw” to examine categorization errors within each emotion, these analyses revealed that anger was significantly more frequently miscategorized as sadness, neutrality, and disgust than as the other emotions ($ps < 0.001$). Additionally, disgust was significantly more frequently miscategorized as sadness than as the other emotions ($ps < 0.001$), and fear was more often miscategorized as surprise than as the other emotions ($ps < 0.001$; see Figure 10).

Figure 10. *Cumulative of Accuracy (%) for Each Facial Emotion*

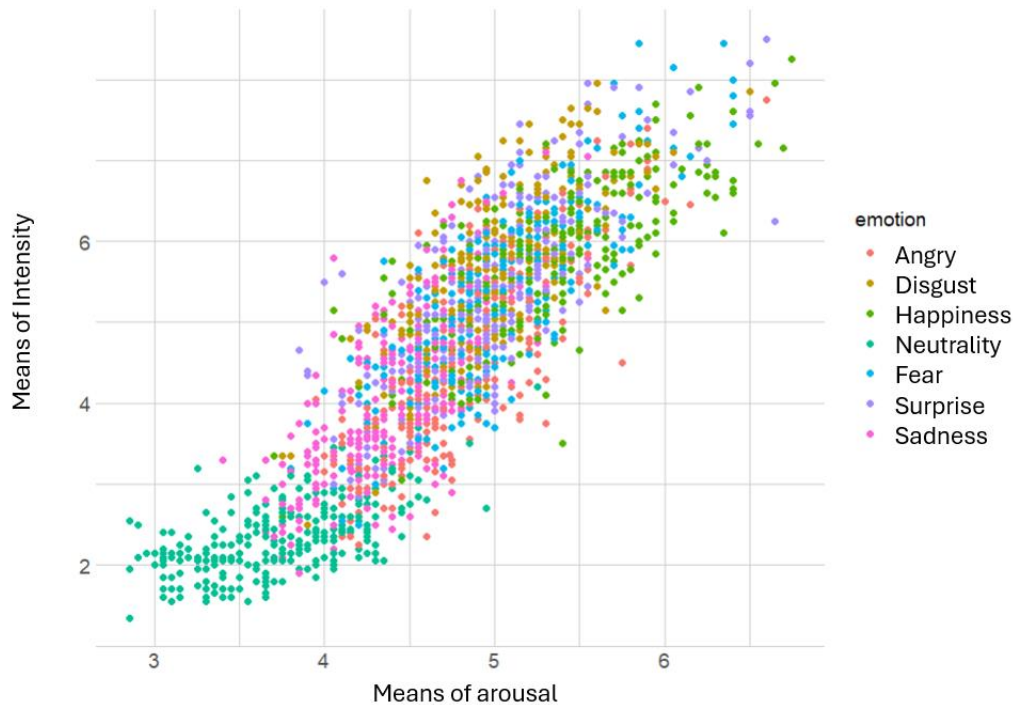


Note. For each facial emotions, the graph gives the percentage (%) in response choice. We can see that certain emotions are mainly confused with each other, such as the facial emotion of disgust with sadness. We can also see that for each facial emotions, the emotional target is mainly chosen in the response choices.

Valence, Arousal, Intensity, and Authenticity: Emotions and Gender Differences

Pearson correlations were computed between the four features (intensity, valence, authenticity, and arousal) of facial emotions to demonstrate the association between them. The analysis revealed a significant association between intensity and arousal of facial emotions ($r = 0.873, p < 0.001$). This association indicates a positive relationship, suggesting that a high level of arousal is correlated with a high level of intensity for facial emotions. No other results were significant (see Figure 11).

Figure 11. Scatterplot Representing the Mean Intensity as a Function of the Mean Arousal for Each Emotional Stimulus



7 (emotion) x 2 (genre of actors) x 2 (genre of raters) x 3 (session) Mixed ANOVAs on the ratings of valence, intensity, arousal, and authenticity revealed significant main effects of emotion, and gender of actors ($p_s < 0.001$). There were also significant main effects of the session for ratings of valence and intensity ($p_s < 0.022$). Additionally, there were an interaction effect between emotion and gender of actors for all ratings ($p_s < 0.001$), and an interaction between emotion and the gender of raters for ratings of valence ($F[1.668, 61.728] = 3.884$, $p = 0.033$, $\eta^2 = 0.018$). Finally, a significant interaction between gender of raters, gender of actors and emotion was found for ratings of authenticity ($F[6, 222] = 2.368$, $p = 0.031$, $\eta^2 = 0.0008$). No other results were significant.

Bonferroni-adjusted paired t -tests revealed significant differences within the four features (valence, intensity, arousal, and authenticity).

Valence. Happiness was rated as having the most positive valence compared to the other emotions ($ps < 0.001$). Anger, disgust, sadness, and fear were rated as having the most negative valence compared to the other emotions ($ps < 0.001$). Regarding the session, valence was assessed as more positive in session three compared to session one ($t[38] = -3.119, p = 0.008, d = -0.219, CI [-0.246, -0.030]$). Moreover, the female actors were rated as more negative ($t[38] = -2.760, p = 0.009, d = -0.07, CI [-0.076, -0.012]$) compared to male actors. Finally, disgust, fear, and sadness were evaluated as more negative when portrayed by female actors rather than male actors ($ps < 0.001$). On the other hand, happiness was evaluated as more positive when portrayed by female actors compared to male actors ($t[38] = 6.625, p < 0.001, d = 0.351, CI [0.104, 0.338]$). No other results were significant.

Arousal. Neutral facial expression was rated as generating the least arousal compared to the other emotions ($ps < 0.001$). Moreover, the female actors were rated as more arousing ($t[36] = 4.104, p < 0.001, d = 0.079, CI [0.039, 0.116]$) compared to male actors. Finally, arousal was greater when happiness, surprise, and sadness were expressed by female actors rather than male actors ($ps < 0.001$). No other results were significant.

Intensity. Happiness was rated as having the strongest intensity (with no difference from disgust, $p = 1$) compared to the other emotions ($ps < 0.001$). Neutral facial expression was rated as the least intense compared to the other emotions ($ps < 0.001$). Regarding the session, the intensity as less important in session three compared to session one ($t[38] = 2.822, p = 0.019, d = 0.238, CI [0.038, 0.539]$). Moreover, the female actors were rated as more intense ($t[38] = 8.06, p < 0.001, d = 0.179, CI [0.163, 0.272]$) compared to male actors. Finally, disgust, happiness, fear, surprise, and sadness were rated as more intense when portrayed by female actors compared to male actors ($ps < 0.001$). No other results were significant.

Authenticity. Happiness was rated as being the most authentic (with no difference from neutral, $p = 1$) compared to the other emotions ($ps < 0.001$). Moreover, the female actors were rated as more authentic ($t[38] = 5.335, p < 0.001, d = 0.177, CI [0.202, 0.449]$) compared to male actors. Finally, disgust, fear, and sadness were evaluated as more authentic when portrayed by female actors rather than male actors ($ps < 0.001$). No other results were significant (see Table 7).

Tables of ANOVA analyses for each measure are available in the supplementary material on the Open Science Framework (OSF; <https://osf.io/frvwt/>).

Body Stimuli

Reliability

All of Pearson's corrections were significant ($ps < 0.001$). The Spearman-Brown coefficients for subgroups by even and odd participant numbers (Group1, $r = 0.880$; Group2, $r = 0.866$) and by gender (Group1, $r = 0.914$; Group2, $r = 0.926$) revealed high split-half reliability for the two stimuli groups. In ratings of intensity, authenticity, valence, and arousal, the Spearman-Brown coefficients for subgroups by even and odd participant numbers (*Intensity*: Group1, $r = 0.879$; Group2, $r = 0.899$; *Valence*: Group1, $r = 0.709$; Group2, $r = 0.880$; *Arousal*: Group1, $r = 0.573$; Group2, $r = 0.730$; *Authenticity*: Group1, $r = 0.663$; Group2, $r = 0.588$) and by gender (*Intensity*: Group1, $r = 0.9$; Group2, $r = 0.877$; *Valence*: Group1, $r = 0.879$; Group2, $r = 0.850$; *Arousal*: Group1, $r = 0.879$; Group2, $r = 0.850$; *Authenticity*: Group1, $r = 0.592$; Group2, $r = 0.733$) revealed high split-half reliability for two stimuli groups.

Table 7. Means and SD for Arousal, Valence, Intensity, and Authenticity of Each Stimuli Emotions by Gender Actors

Emotions	Arousal			Valence		
	Men	Women	Total	Men	Women	Total
Angry	4.86 (1.75)	4.77 (1.73)	4.81 (1.74)	3.78 (1.13)	3.79 (1.09)	3.79 (1.11)
Disgust	4.93 (1.66)	5.04 (1.7)	4.99 (1.68)	3.70 (1.16)	3.55 (1.15)	3.61 (1.15)
Happiness	5.24 (1.82)	5.41 (1.87)	5.34 (1.85)	6.43 (1.51)	6.65 (1.48)	6.55 (1.5)
Neutrality	3.79 (1.8)	3.73 (1.79)	3.76 (1.8)	4.72 (0.99)	4.82 (1.08)	4.77 (1.04)
Fear	4.92 (1.68)	5.02 (1.72)	4.97 (1.7)	4.18 (1.15)	3.87 (1.19)	4.01 (1.18)
Surprise	4.84 (1.68)	4.99 (1.72)	4.93 (1.71)	4.94 (1.28)	4.98 (1.33)	4.96 (1.31)
Sadness	4.34 (1.69)	4.48 (1.71)	4.42 (1.7)	4.01 (1.13)	3.78 (1.12)	3.88 (1.13)
Total	4.70 (1.78)	4.78 (1.82)	4.74 (1.80)	4.54 (1.49)	4.49 (1.58)	4.52 (1.54)

Emotions	Intensity			Authenticity		
	Men	Women	Total	Men	Women	Total
Angry	4.57 (2.29)	4.48 (2.24)	4.52 (2.26)	1.40 (3.68)	1.50 (3.69)	1.46 (3.68)
Disgust	5.49 (1.95)	5.76 (1.99)	5.64 (1.98)	1.42 (3.74)	1.79 (3.73)	1.62 (3.74)
Happiness	5.62 (2.06)	5.98 (1.99)	5.82 (2.03)	3.09 (3.58)	3.39 (3.45)	3.26 (3.51)
Neutrality	2.41 (2.46)	2.34 (2.47)	2.37 (2.47)	3.23 (2.93)	3.29 (2.94)	3.27 (2.93)
Fear	5.11 (2.06)	5.28 (2.11)	5.20 (2.09)	1.59 (3.48)	2.19 (3.37)	1.92 (3.44)
Surprise	5.04 (2.04)	5.49 (2.01)	5.29 (2.03)	1.39 (3.70)	1.64 (3.74)	1.53 (3.73)
Sadness	3.81 (2.23)	4.24 (2.27)	4.05 (2.26)	1.14 (3.78)	1.67 (3.82)	1.43 (3.81)
Total	4.58 (2.40)	4.79 (2.45)	4.7 (2.43)	1.9 (3.66)	2.21 (3.62)	2.1 (3.64)

“Raw Hit Rates” and “Unbiased Hit Rates”

Descriptive data in Tables 8 and 9 show that both the “Raw hit rates” and “Unbiased hit rates” scores for each emotion are significantly above the chance level, except for disgust ($p = 0.06$; see Tables 8 and 9; 1/7 response options; $\mu = 0.14$) and are three times above the chance level, except for disgust, fear, and sadness (see Tables 8 et 9). Considering the Hu scores, neutrality emerges with the highest Hu score (“Unbiased Hit Rate”; 0.70) while disgust exhibits the lowest score (0.10; see Table 8). Taken together, these findings suggest that facial emotions are recognized with a high degree of accuracy and are rarely confused with each other, except for disgust.

Table 8. Summary Table of “Raw Hit Rate” and “Unbiased Hit Rates” for Body Posture

	“Raw Hit Rate” (+/- SD)	“Unbiased Hit Rate” (+/-SD)
Anger	0.90 (+/-0.29) ***	0.68 (+/-0.19) ***
Disgust	0.12 (+/-0.33)	0.10 (+/-0.12)
Happiness	0.69 (+/-0.46) ***	0.58 (+/-0.23) ***
Neutrality	0.96 (+/-0.20) ***	0.70 (+/-0.18) ***
Fear	0.83 (+/-0.38) ***	0.38 (+/-0.10) ***
Surprise	0.48 (+/-0.50) ***	0.35 (+/-0.24) ***
Sadness	0.45 (+/- 0.50) ***	0.41 (+/-0.27) ***
Mean	0.63 (+/-0.38)	0.5 (+/-0.28)

Note. Table contains Mean of “Raw hit rate” (+/- Standard deviation), and Hu score/Unbiased hit rates (+/- Standard deviation) for each facial emotion. The asterisk (***) indicates that the “Raw hit rate” and “Hu score” significance at $ps < 0.001$ *** above chance level (0.14).

Table 9. *Confusion Matrix for the “Raw Hit Rate” of Body Postures*

		RESPONSE							
		Anger	Disgust	Happiness	Neutrality	Fear	Surprise	Sadness	Total
CORRECT	Anger	700 (90.56) ***	1 (0.13)	31 (4.01)	27 (3.49)	10 (1.29)	4 (0.52)	0 (0)	773
	Disgust	37 (4.78)	93 (12.02)	8 (1.03)	27 (3.49)	536 (69.25)	59 (7.62)	14 (1.81)	774
	Happiness	82 (10.61)	3 (0.39)	532 (68.82) ***	11 (1.42)	37 (4.79)	107 (13.84)	1 (0.13)	773
	Neutrality	4 (0.52)	1 (0.13)	6 (0.78)	740 (95.73) ***	0 (0)	0 (0)	22 (2.85)	773
	Fear	22 (2.84)	27 (3.49)	6 (0.78)	14 (1.81)	640 (82.69) ***	59 (7.62)	6 (0.78)	774
	Surprise	149 (19.28)	14 (1.81)	79 (10.22)	78 (10.09)	73 (9.44)	373 (48.25) ***	7 (0.91)	773
	Sadness	13 (1.68)	6 (0.78)	19 (2.46)	194 (25.10)	188 (24.32)	8 (1.03)	345 (44.63) ***	773

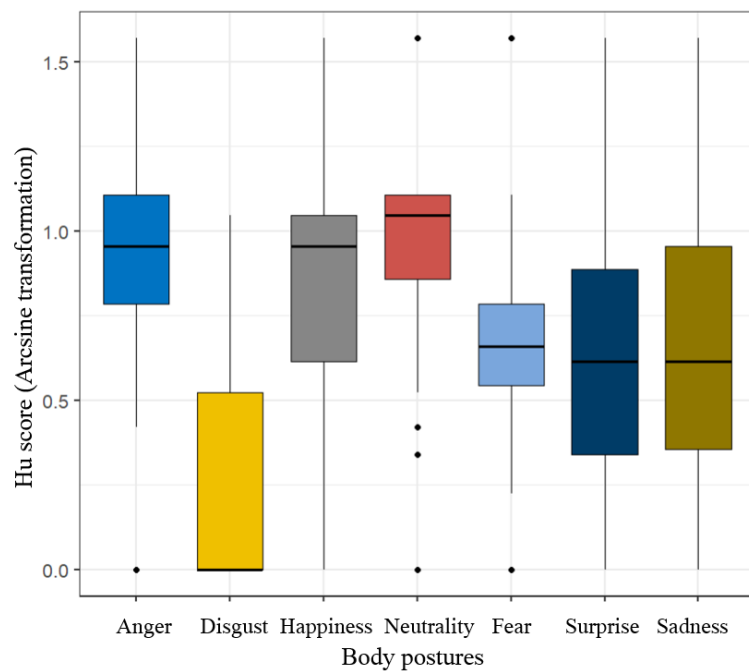
Note. The **bold** numbers represent the number (the “Raw hit rate” (%)) of categorizations for each of body postures and the other numbers represent categorization errors between each emotion. The asterisk (***) indicates that the “Raw hit rate” significance at $ps < 0.001$ *** above chance level (0.14). Confusion matrix was calculated by removing the empty responses.

Recognition of Basic Body Postures: Emotion and Gender Difference

A 7 (emotion) x 2 (genre of actors) x 2 (genre of raters) x 3 (session) Mixed ANOVA on the arcsine transformation of the Hu score revealed a significant main effect of emotion ($F[4.565, 150.640] = 112.994, p < 0.001, \eta^2 = 0.422$), an interaction effect between session and gender of raters ($F[2, 66] = 3.563, p = 0.034, \eta^2 = 0.003$), and between emotion and gender of actors ($F[5.162, 170.356] = 4.064, p < 0.001, \eta^2 = 0.007$). Finally, an interaction between the gender of raters, gender of actors, session, and emotion was also found ($F[9.310, 307.217] = 2.088, p = 0.029, \eta^2 = 0.008$). No other result was significant.

Neutrality is recognized the most accurately, followed by anger, fear, happiness, surprise, sadness, and disgust. Bonferroni-adjusted paired *t*-tests revealed a significant difference between all emotions ($ps < 0.021$) except for no significant difference between happiness, and anger ($t[38] = 2.374, p = .039, d = 0.292, CI [-0.029, 0.227]$), between neutrality and anger ($t[38] = -0.964, p = 1, d = -0.119, CI [-0.169, 0.088]$), and between fear, surprise and sadness ($ps = 1$). Bonferroni-adjusted paired *t*-tests showed that disgust was better recognized when the actors were females rather than male actors ($t[38] = 3.812, p = 0.016, d = -0.437, CI [-0.902, 0.284]$). No other result was significant (see Figure 12).

Figure 12. *Boxplot of the Arcsine Transformation of the Hu Score by Body Postures*

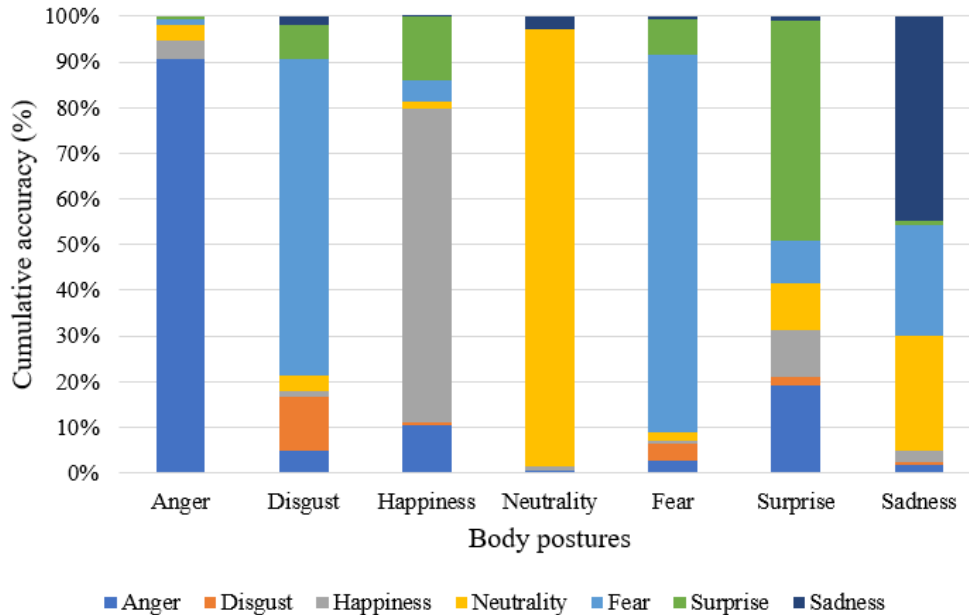


Note. Hu scores for body postures (anger, disgust, happiness, neutrality, fear, surprise, and sadness).

Confusion Between Basic Body Postures

Concerning repeated measure ANOVA on “Raw hit raw” to examine categorization errors within each emotion, these analyses revealed that disgust was significantly more frequently miscategorized as fear than as the other emotions ($ps < 0.001$), and sadness was more often miscategorized as neutrality and fear than as the other emotions ($ps < 0.001$; see Figure 13).

Figure 13. *Cumulative of Accuracy (%) for Each Body Postures*

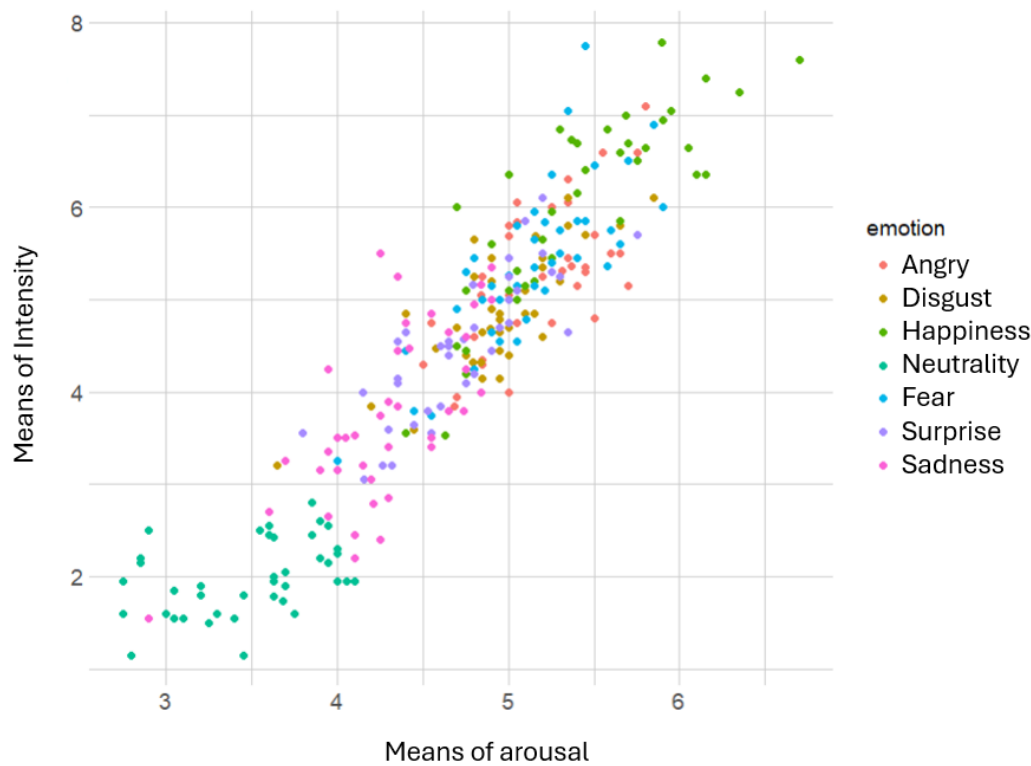


Note. For each body postures, the graph gives the percentage (%) in response choice. We can see that certain emotions are mainly confused with each other, such as the body posture of disgust with fear. We can also see that for each body posture, the emotional target is mainly chosen in the response choices except for the disgust.

Intensity, Valence, Arousal, and Authenticity: Emotions and Gender Differences

Pearson correlations were computed between the four features to show the association between them. The analysis revealed a significant association between intensity and arousal of body postures ($r = 0.817, p < 0.001$). No other significant results were observed (see Figure 14).

Figure 14. Scatterplot Representing the Mean Intensity as a Function of the Mean Arousal for Each Emotional Stimulus



7 (emotion) x 2 (genre of actors) x 2 (genre of raters) x 3 (session) Mixed ANOVAs on the ratings of valence, intensity, arousal, and authenticity indicated a significant main effect of emotion ($p_s < 0.001$). Additionally, significant interactions were observed between the gender of actors and emotion for authenticity ($F[6, 204] = 3.237, p = 0.005, \eta^2 = 0.004$) and intensity ($F[5.714, 182.851] = 4.504, p < 0.001, \eta^2 = 0.002$), as well as between the gender of raters and emotion for arousal ($F[2.656, 92.959] = 4.778, p = 0.005, \eta^2 = 0.026$) and valence ($F[2.636, 94.881] = 3.399, p = 0.026, \eta^2 = 0.018$). Furthermore, a significant main effect of the gender of actors ($F[1, 34] = 5.039, p = 0.031, \eta^2 = 0.001$) and an interaction between the gender of actors and raters ($F[1, 34] = 8.929, p = 0.005, \eta^2 = 0.003$) for authenticity were observed. No other significant results were found.

Bonferroni-adjusted paired *t*-tests revealed significant differences within the four features (valence, intensity, arousal, and authenticity).

Valence. Happiness was the body posture rated as having the most positive valence compared to the other emotions ($ps < 0.001$). On the other hand, anger, disgust, fear, and sadness were rated as having the most negative valence compared to the other emotions ($ps < 0.001$). No other significant results were found.

Arousal. Neutral body posture was rated as generating the least arousal compared to the other emotions ($ps < 0.001$). No other significant results were found.

Intensity. Happiness was the body posture rated as having the strongest intensity compared to the other emotions ($ps < 0.001$). On the other hand, neutral body posture was rated as the least intense compared to the other emotions ($ps < 0.001$). No other significant results were found.

Authenticity. Neutral body posture was rated as the most authentic compared to the other emotions ($ps < 0.001$). Moreover, the female actors were rated as more authentic ($t[38] = 2.245$, $p = 0.031$, $d = 0.08$, CI [0.019, 0.379]) compared to the male actors. Finally, the male raters rated the female actors as more authentic than the male actors ($t[38] = 3.7$, $p = 0.005$, $d = 0.187$, CI [0.113, 0.815]). No other significant results were found (see Table 10).

Tables of ANOVA analyses for each measure are available in the supplementary material on the Open Science Framework (OSF; <https://osf.io/frvwt/>).

Table 10. Means and SD for Arousal, Valence, Intensity, and Authenticity of Each Stimuli Emotions by Gender Actors

Emotions	Arousal			Valence		
	Men	Women	Total	Men	Women	Total
Angry	5.21 (1.86)	5.12 (1.77)	5.16 (1.81)	3.64 (1.77)	3.75 (1.25)	3.70 (1.27)
Disgust	4.9 (1.78)	5 (1.65)	4.96 (1.71)	3.85 (1.65)	3.89 (1.20)	3.87 (1.20)
Happiness	5.40 (1.89)	5.41 (1.81)	5.40 (1.85)	5.94 (1.81)	5.79 (1.69)	5.86 (1.70)
Neutrality	3.43 (1.75)	3.54 (1.85)	3.49 (1.81)	5.11 (1.85)	5.04 (0.87)	5.07 (0.87)
Fear	5.11 (1.69)	5.15 (1.76)	5.13 (1.73)	3.83 (1.76)	3.73 (1.17)	3.78 (1.20)
Surprise	4.78 (1.74)	4.69 (1.67)	4.73 (1.70)	4.86 (1.67)	4.68 (1.22)	4.76 (1.19)
Sadness	4.23 (1.76)	4.35 (1.68)	4.30 (1.71)	4.25 (1.68)	4.13 (1.23)	4.18 (1.27)
Total	4.72 (1.89)	4.75 (1.84)	4.74 (1.86)	4.50 (1.84)	4.43 (1.47)	4.46 (1.47)

Emotions	Intensity			Authenticity		
	Men	Women	Total	Men	Women	Total
Angry	3.75 (1.29)	5.29 (2.14)	5.22 (2.12)	0.52 (2.10)	0.35 (4.21)	0.42 (4.21)
Disgust	3.89 (1.21)	4.89 (2.03)	4.91 (2.03)	1.40 (2.04)	1.46 (3.66)	1.43 (3.68)
Happiness	5.79 (1.72)	6.10 (2.09)	6.01 (2.09)	1.20 (2.09)	1.41 (4.18)	1.32 (4.09)
Neutrality	5.04 (0.88)	2.01 (2.58)	1.99 (2.54)	3.18 (2.52)	3.18 (3.12)	3.18 (3.10)
Fear	3.73 (1.22)	5.48 (1.95)	5.39 (1.98)	1.05 (2.01)	1.44 (4.00)	1.27 (3.88)
Surprise	4.68 (1.16)	4.59 (2.07)	4.51 (2.03)	1.19 (2.00)	0.94 (3.59)	1.05 (3.53)
Sadness	4.13 (1.30)	3.53 (2.15)	3.75 (2.22)	1.11 (2.25)	2.01 (3.65)	1.60 (3.51)
Total	4.56 (2.40)	4.53 (2.45)	4.54 (2.48)	1.38 (3.66)	1.54 (3.62)	1.47 (3.81)

Discussion

Overall, the results of the second part of validation replicated those of first part. In terms of intensity, valence, arousal, and authenticity, several results emerged. Firstly, neutral facial and body expressions were rated as the least intense and generating the lowest arousal (Goeleven et al., 2008) whereas happiness facial and body expressions were rated as the most intense rather than other emotions. This finding is consistent with the dimensional model, where neutrality is associated with low arousal and happiness is associated with high arousal (Russell, 1980). Furthermore, intensities of emotional expressions (facial and bodily) seem to be positively correlated with emotional arousal, consistently with other results (O'Reilly et al., 2016).

Secondly, facial and bodily expressions of happiness were consistently rated as the most positive emotions, while expressions of disgust, anger, sadness, and fear were perceived as more negative compared to other emotions. This aligns with the dimensional model where happiness is inherently positive, and other emotions are typically associated with negative valence (Russell, 1980).

Thirdly, facial and bodily expressions are generally evaluated as moderately authentic (with a range from 0.42 to 3.18 for body posture and 1.43 to 3.27 for facial expressions on a scale ranging from -7 to 7). However, none of the emotions has an average authenticity score below 0 indicating that none of them are considered fake (Dawel et al., 2017). Moreover, the authenticity scores of the ECIFBSS stimuli are like those stimulus sets (Dawel et al., 2017; Ong et al., 2023). The authenticity score of certain facial and bodily emotions, such as fear, surprise, sadness, anger, and disgust, seems to be lower than other emotions like happiness and neutrality (below 1.9). One possible explanation is that the distinction between a genuine and a fake expression appears to be difficult for the human eye to perceive (Fernandes-Magalhaes et al., 2023). Additionally, in this study, the evaluation of the authenticity of the stimuli was conducted on isolated facial and bodily emotions, which likely increased the difficulty. Indeed, in everyday life, facial emotions are perceived within an emotional context, such as body posture and the emotional scene, which aids in recognizing the expressed emotion (Abo Foul et al., 2018; Karaaslan et al., 2020; Kret et al., 2013; Lecker et al., 2020; Mondloch et al., 2013; Willis et al., 2011). Emotional information from other information channels is likely necessary to more accurately assess authenticity. Moreover, the isolated perception of bodily expression may not provide sufficient information to evaluate authenticity, as facial emotion seems to be richer in emotional information and is perceived more prominently (Lecker et al., 2020).

Fourthly, facial expressions of happiness and neutrality, along with neutral body postures, were rated as the most authentic (Dawel et al., 2017). One plausible explanation for the high authenticity of happy expressions could be the frequency of daily occurrences where individuals might express false happiness, such as when posing for photographs. Conversely, expressions like anger are less frequently feigned, resulting in a perception of lower authenticity (e.g., if a parent pretends to be an angry tiger, he wants his child to realize that this is not authentic; (Dawel et al., 2017).

Finally, facial and bodily expressions of basic emotions conveyed by female actors were consistently perceived as more authentic compared to the male actors, suggesting a similar mechanism applies to the recognition of facial and bodily expressions. Moreover, facial basic emotions displayed by women were evaluated as more intense and arousing compared to those expressed by men. This finding could explain why basic emotions exhibited by women were better recognized rather than men in both the first and second validations. Regarding emotional intensity, an intensely expressed emotion is more easily recognizable than a weakly intense emotion (Wingenbach et al., 2016). Indeed, when an individual expresses an emotion intensely, the muscles associated with that emotion contract strongly, making the facial features forming the expression more apparent and thus facilitating its recognition (Wingenbach et al., 2016).

General Discussion

The goal of this study was to provide a novel stimulus set of basic static emotional expressions comprising both faces and bodies in congruent and incongruent conditions. To our knowledge, such a stimulus set does not currently exist, and previous studies have relied on facial basic emotions from one stimulus set and body postures from another one. This methodology generates a potential bias in emotion perception due to the suboptimal quality of stimuli. In response, we created a novel dataset of stimuli, the ECIFBSS, which includes a total of 1952

stimuli featuring both facial expressions and body postures of basic emotions, either in congruent pairs or incongruent combinations. The creation of the ECIFBSS involved photographing 40 actors as they expressed seven basic emotions (anger, disgust, happiness, neutrality, fear, surprise, and sadness) with their faces and bodies. The sample of 40 actors aimed to be diverse in terms of age, gender, ethnicity, and acting skills to create a varied sample for the expression of emotions. The first and second validation results align with our initial hypotheses. As expected, facial emotions are better recognized than body postures. Moreover, the recognition of emotions varied significantly. For instance, happiness displayed a higher recognition rate for facial emotions, while anger exhibited higher recognition among body postures. Lastly, certain emotions were systematically confused with others, such as the confusion between surprise and fear in facial expressions and between sadness and neutrality in body postures. Finally, the facial expression of happiness is perceived as the most authentic emotion, and there is a positive correlation between the intensity and arousal for both facial and bodily expressions.

Overall, both facial emotions and body postures are well recognized. However, the recognition rate of facial emotions and body postures in the ECIFBSS stimulus set is subject to certain biases. First, during the creation of the stimuli, simultaneously expressing two different emotions with the face and body posture (incongruent condition) can be challenging and may limit the range of body postures that actors can effectively express (e.g., a downcast head for sadness or a hand on the face for surprise). These methodological challenges may result in reduced “Raw hit rates” and “Unbiased hit rate”. To facilitate emotional expression, actors were initially instructed to express congruent emotions with both their faces and bodies, then modify only the facial expression (incongruent condition), and finally return to a congruent condition. Second, the forced choice task during the validation stimuli, which involved choosing from seven emotional options (happiness, surprise, anger, disgust, fear, sadness, and neutrality), may have

contributed to increased confusion compared to studies using two (Thoma et al., 2013) or four alternative forced choice tasks (de Gelder & van den Stock, 2011). However, these rates could be inflated when forced choices direct participants toward predefined answers (Russell, 1994). Thirdly, the use of verbal labels in the emotional recognition task may have influenced the recognition. Some studies have suggested that bodily postures are better recognized without verbal label support (van den Stock et al., 2007). Lastly, there is an imbalance between the number of female and male raters in the first validation (47 females; 29 males) whereas there are balanced samples in the second validation (20 females and 20 males). It seems unlikely that this difference had an influence on our observations, especially since the results concerning the raters' gender were similar across both validations, indicating little to no influence of this imbalance on the recognition of facial and bodily expressions. In contrast, the ECIFBSS offers valid **static** emotional stimuli designed to encompass a wide range of basic emotions. Notably, the actors were given the freedom to authentically express various emotions, unbound by scripted scenarios. The approach has yielded a stimulus set of emotional stimuli characterized by expressions closely mirroring real-life emotional displays.

Advantages and Limits of the ECIFBSS

The ECIFBSS is an innovative stimulus set offering many stimuli designed to study emotional processing as close as possible to reality. This stimulus set presents several advantages. Firstly, the use of professional equipment during the stimulus creation ensures high-quality images in full color, a critical factor as image quality can significantly affect emotion perception (Wallbott, 1992). Secondly, the stimulus set exhibits a wide range of variables, including age (18 to 70 years old), gender (similar number of men and women), ethnicity, and in terms of emotions displayed (anger, disgust, neutrality, happiness, fear, surprise, and sadness). This rich variability opens interesting possibilities for future research. The ECIFBSS can be

effectively utilized to explore the perception of basic emotions by controlling and manipulating variables such as age, gender, and ethnicity. Furthermore, the inclusion of many basic emotions facilitates the examination of emotional processing in a more realistic and complex manner. In addition, the ECIFBSS provides both facial and bodily expressions of basic emotions in congruence and incongruence, allowing for the study of emotions involving facial and postural emotional information without the need for extensive photo editing, which could potentially compromise quality and realism. Finally, the ECIFBSS has correct internal consistency and provides robust validation through various features such as intensity, authenticity, valence, and arousal, enabling a wide application of this stimulus set.

These advantages highlight the potential utility of the ECIFBSS in future research. Firstly, studies focusing solely on basic emotions expressed by facial emotions or body postures will be able to utilize the images isolating these two channels of emotional information available in the ECIFBSS. Additionally, studies investigating the interaction between facial and bodily expressions of basic emotions can leverage the ECIFBSS without the need for manipulation via software such as Photoshop, while controlling various parameters such as the gender of the actors. Finally, data on intensity, authenticity, valence, and arousal expand the potential applications of the ECIFBSS, for example, to examine the implications of the intensity of facial and bodily expressions in the emotional perception.

Nonetheless, some limitations are inherent to this stimulus set. First, despite being based on real-life scenarios, the images were captured in laboratory settings to control many variables, such as lighting, and were not obtained from genuine real-life contexts. Additionally, the actors had to express incongruent facial and bodily emotional expressions, which could also reduce the authenticity of the emotions portrayed. This might impact the way actors express their emotions, as previous research has identified distinctions between voluntary and spontaneous emotions

(Zuckerman et al., 1976). Nevertheless, as de Gelder and van den Stock (2011) also suggested, the expertise of actors and the use of scenarios to elicit emotions should have reduced this impact. Furthermore, in the second validation part, we evaluated the authenticity of facial and bodily expressions, indicating a correct level of authenticity in alignment with other databases (Dawel et al., 2017; Ong et al., 2023). Second, despite instructions from the experimenters, some actors subtly tilted their heads when expressing certain facial emotions. This introduces a limitation, as the perception of emotions varies between tilted and untilted expressions. This aspect could be considered in future studies by using both tilted and untilted emotion stimuli expressed by the same actor. Third, like other databases (Dawel et al., 2017; Goeleven et al., 2008; Thoma et al., 2013; Tottenham et al., 2009), the ECIFBSS was not created based on the Facial Action Coding System (FACS; Ekman & Friesen, 1976) for facial basic emotions, partly because of the tedious nature of this methodology (Goeleven et al., 2008). We favored a methodology using direct instructions to facilitate participants' focus on the recognition of basic emotions (Goeleven et al., 2008).

Lastly, the facial and body expressions in the ECIFBSS stimulus set are derived from static images rather than dynamic media (e.g., videos). This may be seen as a limitation. Indeed, in everyday life, we mostly perceive emotions dynamically, and the processing of dynamic and static emotions differs. Dynamic emotions are notably facilitated by perceptible movements in expressions (Ambadar et al., 2005; Cunningham & Wallraven, 2009). Therefore, it is essential in future studies to increase the number of high-quality dynamic emotional stimuli. However, despite this limitation, the ECIFBSS still makes a significant contribution to the scientific community, particularly in the study of congruence between facial and bodily expressions. The quality of static stimuli currently used in the study of interactions between facial and bodily expressions has been noted to be very poor (Abo Foul et al., 2018; Aviezer et al., 2011; Li, 2021),

impacting the results obtained from these studies. The availability of validated, high-quality stimuli would improve future research in this field. Furthermore, a previous study on the influence of facial and bodily expressions using dynamic stimuli (Nelson & Mondloch, 2017) achieved similar results to a study using static stimuli (Lecker et al., 2020), indicating potentially similar mechanisms in the interaction of statically and dynamically perceived facial and body expressions. Future studies should further explore the influences between facial and bodily expressions at both static and dynamic levels with better quality stimuli.

Conclusion

In alignment with the current research trends focusing on studying emotion perception as closely as possible to real-life situations by integrating emotions from multiple channels of emotional information, such as facial expressions and body postures, the creation of the ECIFBSS emerges as a highly pertinent development. Notably, it represents the pioneering stimulus set to offer validated stimuli featuring both facial basic emotions and body postures in congruent and incongruent contexts. Furthermore, this study demonstrates that accurate recognition of facial emotions and body postures produces results in emotional recognition that align with previous research. The ECIFBSS consists of a stimulus set of visually high-quality and diverse stimuli, characterized by variations in age, gender, number of actors, and the range of included emotions. This diversity equips future studies with a valuable resource for deeper exploration into the integration of multiple simultaneously presented channels of emotional information and the interactions that underlie them.

Declaration

Acknowledgements

We thank Cindy Levesque for photography materials and the Quebec City campus of the UQTR for the provision of an adapted room for the realization of this project.

Funding

This work was supported by “Natural Sciences and Engineering Research Council of Canada” (NSERC) Discovery Grant (Grant numbers: RGPIN-2020-06706).

Conflicts of interest

The authors declare that they have no conflict of interest, financial, or otherwise.

Ethics approval

Approval was obtained from the ethic committee of University of Quebec at Trois-Rivières. The ethical certification number is CERPPE-22-06-07.03.

Consent to participate

Participant in creation and validations of stimuli read an information and consent form and provided their written consent. All data were rendered pseudonymized. All participants consented to take part in the study and agreed for the results to be used for this study and published in a scientific journal. The participants were volunteers and could stop the experiment at any time.

Consent for publication

All actors agreed in writing that their images being disseminated, communicated, or made accessible to the public by any means and on all media, within the framework of the project and all related activities. All actors except two (image from these two participants are therefore not present on OSF) have consented to the secondary use of the images so that they can be stored on

the Open Science Framework (OSF) in a stimulus set intended for the scientific community, and they may be used by its members in the future, within the context of other research projects.

Availability of data and materials

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The datasets and stimuli generated during and/or analysed during the current study are available in the Open Science Framework (OSF) repository,

<https://osf.io/frvwt/>, <https://osf.io/tyq4u/>, <https://osf.io/qpyec/>, <https://osf.io/2f3jb/>.

Code availability

R script and code is available at OSF <https://osf.io/frvwt/>.

Author's contributions

Anne-Sophie Puffet: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing-original draft, Writing-reviewing, and editing. *Simon Rigoulot*: Conceptualization, Funding acquisition, Resources, Supervision, Writing-reviewing and editing, Investigation, Methodology, Project administration, Validation.

Open Practices Statement

The data and materials for all experiments are available at <https://osf.io/frvwt/>, <https://osf.io/tyq4u/>, <https://osf.io/qpyec/>, <https://osf.io/2f3jb/> and this study's design and its analysis were not pre-registered.

References

- Abo Foul, Y., Eitan, R., & Aviezer, H. (2018). Perceiving emotionally incongruent cues from faces and bodies: Older adults get the whole picture. *Psychology and Aging, 33*(4), 660-666. <https://doi.org/10.1037/pag0000255>
- Albohn, D. N., Brandenburg, J. C., Kveraga, K., & Adams, R. B., Jr. (2022). The shared signal hypothesis: Facial and bodily expressions of emotion mutually inform one another. *Attention, Perception, & Psychophysics, 84*(7), 2271-2280. <https://doi.org/10.3758/s13414-022-02548-6>
- 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. <https://doi.org/10.1111/j.0956-7976.2005.01548.x>
- 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. <https://doi.org/10.1068/p5096>
- Aviezer, H., Bentin, S., Dudarev, V., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion, 11*(6), 1406-1414. <https://doi.org/10.1037/a0023578>
- Aviezer, H., Ensenberg, N., & Hassin, R. R. (2017). The inherently contextualized nature of facial emotion perception. *Current Opinion in Psychology, 17*, 47-54. <https://doi.org/10.1016/j.copsyc.2017.06.006>
- Aviezer, H., Hassin, R. R., Bentin, S., & Trope, Y. (2008). Putting facial expressions back in context. In I. N. A. J. Skowronski (Ed.), *First impressions* (pp. 255-286). Guilford Press.

- 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. <https://doi.org/10.1111/j.1467-9280.2008.02148.x>
- Balconi, M., & Carrera, A. (2011). Cross-modal integration of emotional face and voice in congruous and incongruous pairs: The P2 ERP effect. *Journal of Cognitive Psychology, 23*(1), 132-139. <https://doi.org/10.1080/20445911.2011.473560>
- Calvo, M. G., Gutiérrez-García, A., Fernández-Martín, A., & Nummenmaa, L. (2014). Recognition of facial expressions of emotion is related to their frequency in everyday life. *Journal of Nonverbal Behavior, 38*(4), 549-567. <https://doi.org/10.1007/s10919-014-0191-3>
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research, 1242*, 126-135. <https://doi.org/10.1016/j.brainres.2008.04.023>
- Cosgrove, K. P., Mazure, C. M., & Staley, J. K. (2007). Evolving knowledge of sex differences in brain structure, function, and chemistry. *Biological Psychiatry, 62*(8), 847-855. <https://doi.org/10.1016/j.biopsych.2007.03.001>
- Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior, 28*, 117-139. <https://doi.org/10.1023/B:JONB.0000023655.25550.be>
- Cunningham, D. W., & Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. *Journal of Vision, 9*(13), Article 7. <https://doi.org/10.1167/9.13.7>

- Dawel, A., Wright, L., Irons, J., Dumbleton, R., Palermo, R., O'Kearney, R., & McKone, E. (2017). Perceived emotion genuineness: Normative ratings for popular facial expression stimuli and the development of perceived-as-genuine and perceived-as-fake sets. *Behavior Research Methods*, *49*(4), 1539-1562. <https://doi.org/10.3758/s13428-016-0813-2>
- de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3475-3484. <https://doi.org/10.1098/rstb.2009.0190>
- de Gelder, B., & van den Stock, J. (2011). The Bodily Expressive Action Stimulus Test (BEAST). Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in Psychology*, *2*, Article 181. <https://doi.org/10.3389/fpsyg.2011.00181>
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*(3), 550-553. <https://doi.org/10.1037/0033-295X.99.3.550>
- Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 45-60). John Wiley & Sons Ltd. <https://doi.org/10.1002/0470013494.ch3>
- Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology & Nonverbal Behavior*, *1*(1), 56-75. <https://doi.org/10.1007/BF01115465>
- Falagiarda, F., & Collignon, O. (2019). Time-resolved discrimination of audio-visual emotion expressions. *Cortex*, *119*, 184-194. <https://doi.org/10.1016/j.cortex.2019.04.017>
- Fernandes-Magalhaes, R., Carpio, A., Ferrera, D., van Ryckeghem, D., Pelaez, I., Barjola, P., De Lahoz, M. E., Martin-Buro, M. C., Hinojosa, J. A., van Damme, S., Carretie, L., & Mercado, F. (2023). Pain E-motion Faces Database (PEMF): Pain-related micro-clips for emotion research. *Behavior Research Methods*, *55*(7), 3831-3844. <https://doi.org/10.3758/s13428-022-01992-4>

- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378-382. <https://doi.org/10.1037/h0031619>
- Focker, J., Gondan, M., & Roder, B. (2011). Preattentive processing of audio-visual emotional signals. *Acta Psychologica*, 137(1), 36-47. <https://doi.org/10.1016/j.actpsy.2011.02.004>
- Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, 57(2), 212-228. <https://doi.org/10.1037/0022-3514.57.2.212>
- Gerdes, A. B., Wieser, M. J., & Alpers, G. W. (2014). Emotional pictures and sounds: A review of multimodal interactions of emotion cues in multiple domains. *Frontiers in Psychology*, 5, Article 1351. <https://doi.org/10.3389/fpsyg.2014.01351>
- Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: A validation study. *Cognition & Emotion*, 22(6), 1094-1118. <https://doi.org/10.1080/02699930701626582>
- Gregoric, B., Barbir, L., Celic, A., Ilakovac, V., Hercigonja-Szekeres, M., Perkovic Kovacevic, M., & Heffer, M. (2014). Recognition of facial expressions in men and women. *Medicina Fluminensis: Medicina Fluminensis*, 50(4), 454-461. <https://hrcak.srce.hr/129788>
- Haverkamp, N., & Beauducel, A. (2017). Violation of the sphericity assumption and its effect on type-i error rates in repeated measures ANOVA and Multi-Level Linear Models (MLM). *Frontiers in Psychology*, 8, Article 1841. <https://doi.org/10.3389/fpsyg.2017.01841>
- 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. <https://doi.org/10.1016/j.cub.2009.07.051>

- Karaaslan, A., Durmuş, B., & Amado, S. (2020). Does body context affect facial emotion perception and eliminate emotional ambiguity without visual awareness? *Visual Cognition*, 28(10), 605-620. <https://doi.org/10.1080/13506285.2020.1846649>
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., & Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: An event-related fMRI study. *Neuroimage*, 37(4), 1445-1456. <https://doi.org/10.1016/j.neuroimage.2007.06.020>
- Kret, M. E., & de Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia*, 50(7), 1211-1221. <https://doi.org/10.1016/j.neuropsychologia.2011.12.022>
- Kret, M. E., & de Gelder, B. (2013). When a smile becomes a fist: The perception of facial and bodily expressions of emotion in violent offenders. *Experimental Brain Research*, 228(4), 399-410. <https://doi.org/10.1007/s00221-013-3557-6>
- Kret, M. E., Stekelenburg, J. J., Roelofs, K., & de Gelder, B. (2013). Perception of face and body expressions using electromyography, pupillometry and gaze measures. *Frontiers in Psychology*, 4, Article 28. <https://doi.org/10.3389/fpsyg.2013.00028>
- LaFrance, M., Hecht, M. A., & Paluck, E. L. (2003). The contingent smile: A meta-analysis of sex differences in smiling. *Psychological Bulletin*, 129(2), 305-334. <https://doi.org/10.1037/0033-2909.129.2.305>
- Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 33(2), 363-374. <https://doi.org/10.2307/2529786>

- Lassalle, A., Pigat, D., O'Reilly, H., Berggen, S., Fridenson-Hayo, S., Tal, S., Elfstrom, S., Rade, A., Golan, O., Bolte, S., Baron-Cohen, S., & Lundqvist, D. (2019). The EU-Emotion Voice Database. *Behavior Research Methods*, *51*(2), 493-506.
<https://doi.org/10.3758/s13428-018-1048-1>
- Lecker, M., Dotsch, R., Bijlstra, G., & Aviezer, H. (2020). Bidirectional contextual influence between faces and bodies in emotion perception. *Emotion*, *20*(7), 1154-1164.
<https://doi.org/10.1037/emo0000619>
- Leys, C., Delacre, M., Mora, Y. L., Lakens, D., & Ley, C. (2019). How to classify, detect, and manage univariate and multivariate outliers, with emphasis on pre-registration. *International Review of Social Psychology*, *32*(1). <https://doi.org/10.5334/irsp.289>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764-766.
<https://doi.org/10.1016/j.jesp.2013.03.013>
- Li, X. (2021). Recognition characteristics of facial and bodily expressions: Evidence from ERPs. *Frontiers in Psychology*, *12*, Article 680959. <https://doi.org/10.3389/fpsyg.2021.680959>
- Liu, T., Pinheiro, A., Zhao, Z., Nestor, P. G., McCarley, R. W., & Niznikiewicz, M. A. (2012). Emotional cues during simultaneous face and voice processing: Electrophysiological insights. *PLoS One*, *7*(2), Article e31001. <https://doi.org/10.1371/journal.pone.0031001>
- Lopez, L. D., Reschke, P. J., Knothe, J. M., & Walle, E. A. (2017). Postural communication of emotion: Perception of distinct poses of five discrete emotions. *Frontiers in Psychology*, *8*, Article 710. <https://doi.org/10.3389/fpsyg.2017.00710>

- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122-1135.
<https://doi.org/10.3758/s13428-014-0532-5>
- Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion*, *5*(1), 119-124.
<https://doi.org/10.1037/1528-3542.5.1.119>
- Matsumoto, D., Sung Hwang, H., Harrington, N., Olsen, R., & King, M. (2011). Facial behaviors and emotional reactions in consumer research. *Acta de investigación psicológica*, *1*(3), 441-453. <https://doi.org/10.22201/fpsi.20074719e.2011.3.199>
- McDuff, D., Kodra, E., Kaliouby, R. E., & LaFrance, M. (2017). A large-scale analysis of sex differences in facial expressions. *PLoS One*, *12*(4), Article e0173942.
<https://doi.org/10.1371/journal.pone.0173942>
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, *102*(45), 16518-16523. <https://doi.org/10.1073/pnas.0507650102>
- Mende-Siedlecki, P., Qu-Lee, J., Lin, J., Drain, A., & Goharзад, A. (2020). The Delaware Pain Database: A set of painful expressions and corresponding norming data. *Pain Reports*, *5*(6), Article e853. <https://doi.org/10.1097/PR9.0000000000000853>
- Miolla, A., Cardaioli, M., & Scarpazza, C. (2023). Padova Emotional Dataset of Facial Expressions (PEDFE): A unique dataset of genuine and posed emotional facial expressions. *Behavior Research Methods*, *55*(5), 2559-2574.
<https://doi.org/10.3758/s13428-022-01914-4>

- Mondloch, C. J., Nelson, N. L., & Horner, M. (2013). Asymmetries of influence: Differential effects of body postures on perceptions of emotional facial expressions. *PLoS One*, *8*(9), Article e73605. <https://doi.org/10.1371/journal.pone.0073605>
- Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Visual Cognition*, *25*(4-6), 563-574. <https://doi.org/10.1080/13506285.2017.1301615>
- Novak, L. R., Gitelman, D. R., Schuyler, B., & Li, W. (2015). Olfactory-visual integration facilitates perception of subthreshold negative emotion. *Neuropsychologia*, *77*, 288-297. <https://doi.org/10.1016/j.neuropsychologia.2015.09.005>
- O'Reilly, H., Pigat, D., Fridenson, S., Berggren, S., Tal, S., Golan, O., Bolte, S., Baron-Cohen, S., & Lundqvist, D. (2016). The EU-Emotion Stimulus Set: A validation study. *Behavior Research Methods*, *48*(2), 567-576. <https://doi.org/10.3758/s13428-015-0601-4>
- Ong, J. H., Leung, F. Y. N., & Liu, F. (2023). The Reading Everyday Emotion Database (REED): A set of audio-visual recordings of emotions in music and language. *Lang Resources & Evaluation*, *59*, 27-49. <https://doi.org/10.1007/s10579-023-09698-5>
- Paulmann, S., & Pell, M. D. (2011). Is there an advantage for recognizing multi-modal emotional stimuli? *Motivation and Emotion*, *35*(2), 192-201. <https://doi.org/10.1007/s11031-011-9206-0>
- Peschard, V., Maurage, P., & Philippot, P. (2014). Towards a cross-modal perspective of emotional perception in social anxiety: Review and future directions. *Frontiers in Human Neuroscience*, *8*, Article 322. <https://doi.org/10.3389/fnhum.2014.00322>
- Pourtois, G., Debatisse, D., Despland, P. A., & de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Cognitive Brain Research*, *14*(1), 99-105. [https://doi.org/10.1016/S0926-6410\(02\)00064-2](https://doi.org/10.1016/S0926-6410(02)00064-2)

- Reschke, P. J., & Walle, E. A. (2021). The unique and interactive effects of faces, postures, and scenes on emotion categorization. *Affective Science*, 2(4), 468-483.
<https://doi.org/10.1007/s42761-021-00061-x>
- Roy-Charland, A., Perron, M., Beaudry, O., & Eady, K. (2014). Confusion of fear and surprise: A test of the perceptual-attentional limitation hypothesis with eye movement monitoring. *Cognition and Emotion*, 28(7), 1214-1222.
<https://doi.org/10.1080/02699931.2013.878687>
- Ruiz-Padial, E., Pastor, M. C., Mercado, F., Mata-Martin, J. L., & Garcia-Leon, A. (2021). MATTER in emotion research: Spanish standardization of an affective image set. *Behavior Research Methods*, 53(5), 1973-1985. <https://doi.org/10.3758/s13428-021-01567-9>
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178. <https://doi.org/10.1037/h0077714>
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, 115(1), 102-141.
<https://doi.org/10.1037/0033-2909.115.1.102>
- Russell, A., Buerkner, P., Herve, M., Love, J., & Singmann, H. (2021). Package 'emmeans'. *R package version*.
- Schindler, K., van Gool, L., & de Gelder, B. (2008). Recognizing emotions expressed by body pose: A biologically inspired neural model. *Neural Networks*, 21(9), 1238-1246.
<https://doi.org/10.1016/j.neunet.2008.05.003>
- Seubert, J., Kellermann, T., Loughhead, J., Boers, F., Brensinger, C., Schneider, F., & Habel, U. (2010). Processing of disgusted faces is facilitated by odor primes: A functional MRI study. *Neuroimage*, 53(2), 746-756. <https://doi.org/10.1016/j.neuroimage.2010.07.012>

- Song, S., Wu, M., & Feng, C. (2022). Early influence of emotional scenes on the encoding of fearful expressions with different intensities: An event-related potential study. *Frontiers in Human Neuroscience*, *16*, Article 866253. <https://doi.org/10.3389/fnhum.2022.866253>
- Susskind, J. M., Littlewort, G., Bartlett, M. S., Movellan, J., & Anderson, A. K. (2007). Human and computer recognition of facial expressions of emotion. *Neuropsychologia*, *45*(1), 152-162. <https://doi.org/10.1016/j.neuropsychologia.2006.05.001>
- Talwar, S., Barbero, F. M., Calce, R. P., & Collignon, O. (2023). Automatic brain categorization of discrete auditory emotion expressions. *Brain topography*, *36*(6), 854-869. <https://doi.org/10.1007/s10548-023-00983-8>
- Thoma, P., Soria Bauser, D., & Suchan, B. (2013). BESST (Bochum Emotional Stimulus Set): A pilot validation study of a stimulus set containing emotional bodies and faces from frontal and averted views. *Psychiatry Research*, *209*(1), 98-109. <https://doi.org/10.1016/j.psychres.2012.11.012>
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242-249. <https://doi.org/10.1016/j.psychres.2008.05.006>
- Tsantani, M., Podgajicka, V., Gray, K. L. H., & Cook, R. (2022). How does the presence of a surgical face mask impair the perceived intensity of facial emotions? *PLoS One*, *17*(1), Article e0262344. <https://doi.org/10.1371/journal.pone.0262344>
- van den Stock, J., & de Gelder, B. (2014). Face identity matching is influenced by emotions conveyed by face and body. *Frontiers in Human Neuroscience*, *8*, Article 53. <https://doi.org/10.3389/fnhum.2014.00053>

- van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, 7(3), 487-494. <https://doi.org/10.1037/1528-3542.7.3.487>
- van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion*, 11(4), 907-920. <https://doi.org/10.1037/a0023853>
- Vroomen, J., Driver, J., & Gelder, B.D. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cognitive, Affective & Behavioral Neuroscience*, 1(4), 382-387. <https://doi.org/10.3758/CABN.1.4.382>
- Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, 17, 3-28. <https://doi.org/10.1007/BF00987006>
- Wallbott, H. G. (1988). Big girls don't frown, big boys don't cry – Gender differences of professional actors in communicating emotion via facial expression. *Journal of Nonverbal Behavior*, 12(2), 98-106. <https://doi.org/10.1007/BF00986928>
- Wallbott, H. G. (1992). Effects of distortion of spatial and temporal resolution of video stimuli on emotion attributions. *Journal of Nonverbal Behavior*, 16, 5-20. <https://doi.org/10.1007/BF00986876>
- Wells, L. J., Gillespie, S. M., & Rotshtein, P. (2016). Identification of emotional facial expressions: Effects of expression, intensity, and sex on eye gaze. *PLoS One*, 11(12), Article e0168307. <https://doi.org/10.1371/journal.pone.0168307>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. <https://ggplot2.tidyverse.org>

- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), Article 1686.
<https://doi.org/10.21105/joss.01686>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). dplyr: A grammar of data manipulation. <https://github.com/tidyverse/dplyr>
- Willis, M. L., Palermo, R., & Burke, D. (2011). Judging approachability on the face of it: The influence of face and body expressions on the perception of approachability. *Emotion*, 11(3), 514-523. <https://doi.org/10.1037/a0022571>
- Wingenbach, T. S., Ashwin, C., & Brosnan, M. (2016). Validation of the Amsterdam Dynamic Facial Expression Set--Bath Intensity Variations (ADFES-BIV): A set of videos expressing low, intermediate, and high intensity emotions. *PLoS One*, 11(1), Article e0147112. <https://doi.org/10.1371/journal.pone.0147112>
- Zhang, M., Liu, T., Jin, Y., He, W., Huang, Y., & Luo, W. (2019). The asynchronous influence of facial expressions on bodily expressions. *Acta Psychologica*, 200, Article 102941. <https://doi.org/10.1016/j.actpsy.2019.102941>
- Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67(6), 361-370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>
- Zimmer, U., Wendt, M., & Pacharra, M. (2022). Enhancing allocation of visual attention with emotional cues presented in two sensory modalities. *Behavioral and Brain Functions*, 18(1), Article 10. <https://doi.org/10.1186/s12993-022-00195-3>

Zuckerman, M., Hall, J. A., DeFrank, R. S., & Rosenthal, R. (1976). Encoding and decoding of spontaneous and posed facial expressions. *Journal of Personality and Social Psychology*, 34(5), 966-977. <https://doi.org/10.1037/0022-3514.34.5.966>

Transition vers le Chapitre 3

Le premier article propose des stimuli d'expressions émotionnelles faciales et corporelles en situation de congruence et d'incongruence. Ces stimuli, de haute qualité, pallient les limites des études antérieures qui utilisaient des images en noir et blanc issues d'ensembles de stimuli variés. L'ECIFBSS sera utile tant pour la suite des études de cette thèse que pour la communauté scientifique.

La prochaine étude exploitera ces stimuli validés, permettant de se rapprocher des conditions de perception de la vie quotidienne, afin d'examiner l'automatisme de l'effet contextuel entre expressions faciales et corporelles en manipulant la charge cognitive. Pour cela, nous avons demandé à vingt-huit participants de réaliser une tâche de reconnaissance émotionnelle à partir d'expressions faciales et corporelles congruentes ou incongruentes, tout en effectuant une double tâche de mémorisation sous faible ou forte charge cognitive. L'activité cérébrale a été enregistrée par EEG, et les taux de reconnaissance ainsi que les temps de réponse ont été mesurés.

Chapitre 3

Article 2 – The Role of Cognitive Load in Automatic Integration of Emotional Information from Face and Body

The Role of Cognitive Load in Automatic Integration of Emotional Information from Face and Body¹

*Anne-Sophie Puffet^{1,2}, and Simon Rigoulot^{1,2}

¹Department of Psychology, University of Quebec at Trois-Rivières, Canada.

²Research group CogNAC (Cognition, Neurosciences, Affect and Behaviour).

Author Note

*Anne-Sophie Puffet <https://orcid.org/0000-0001-9352-2526>; anne-sophie.puffet@uqtr.ca.

Simon Rigoulot <https://orcid.org/0000-0002-9069-6149>; simon.rigoulot@uqtr.ca.

Statements and Declarations

Funding This work was supported by “Natural Sciences and Engineering Research Council of Canada” (NSERC) Discovery Grant (Grant numbers: RGPIN-2020-06706).

Competing interest The authors have no competing interests to declare that are relevant to the content of this article.

Author’s contribution:

A-S.P: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing-original draft, Writing-reviewing, and editing.

S.R: Conceptualization, Funding acquisition, Resources, Supervision, Writing-reviewing and editing, Investigation, Methodology, Project administration.

¹ Les normes de publication pour la revue *Scientific Reports* sont disponibles à <https://www.nature.com/srep/author-instructions/submission-guidelines>

Abstract

We perceive emotions daily through facial expressions, often accompanied by a body posture that provide additional emotional context. Congruent facial and bodily expressions (conveying the same emotion) enhance emotional recognition compared to incongruent ones, suggesting interaction between these channels. Although behavioral evidence suggests that this integration occur automatic, its underlying neural mechanisms remains unclear. This study investigated the automaticity of facial and bodily expressions integration by manipulating cognitive load. Twenty-eight participants completed an emotion recognition task with congruent or incongruent facial and bodily expressions while performing a memory task under low or high cognitive load. EEG recordings captured brain activity, and emotion recognition accuracy and reaction times were measured. Results revealed that congruent expressions improved recognition, with bodily expressions exerting a stronger influence on facial expression recognition than vice versa. Early neural responses (P100, N100, P250, N250) were stronger during facial expression focus, while later responses reflected attention to body expressions. Bayesian analyses provided strong evidence for the absence of significant interaction between congruence and cognitive load, supporting the automaticity of integration. These findings suggest that emotional expressions are integrated automatically, independent of cognitive resources, and emphasize the differential influence of bodily expressions over facial expressions in shaping emotional perception.

Keywords: Cognitive load; contextual effect; facial emotion; bodily expression; automaticity

Introduction

In everyday life, we primarily perceive emotions through facial expressions, which serve as rapid and easily recognizable indicators of a person's emotional state (Ekman, 1992; Tracy & Robins, 2008). However, facial expressions are sometimes ambiguous, less common and more difficult to identify, particularly when emotions are concealed by social norms or complex feelings such as embarrassment or shame. For example, a smile may conceal underlying nervousness, complicating the accurate recognition of emotions based solely on facial cues. In such situations, additional emotional context becomes essential. Surrounding cues, such as body language (Gu et al., 2013), background scenes (Kret & de Gelder, 2010), vocal tone (Focker et al., 2011), and even olfactory cues (Seubert et al., 2010), provide critical information that enhances the accurate recognition of emotions conveyed through facial expressions.

Emotions are typically perceived through multiple sensory channels simultaneously, which helps emotional recognition, particularly when these cues are congruent (i.e., expressing the same emotion; (Belin et al., 2012). For example, in a horror movie, visual and auditory fear cues combine to amplify the emotional experience. In contrast, pairing a horror scene with cheerful or absent music can make it harder to perceive fear conveyed in the film. Coherently with these observations, numerous studies have showed that congruent audio-visual cues improve emotional recognition, while incongruent cues make it more difficult (Jessen & Kotz, 2011; Paulmann et al., 2009; Seubert et al., 2010; van den Stock et al., 2007). This interaction also occurs within a single sensory modality, such as vision, where multiple emotional cues may be perceived simultaneously. For instance, facial expressions often appear alongside bodily expressions (Lecker et al., 2020; Mondloch et al., 2013), emotional background scene (e.g., a disgusted context such as garbage; (Righart & de Gelder, 2006; van den Stock et al., 2014), or even a combination of all these cues (Reschke & Walle, 2021). When facial and bodily

expressions are combined, emotion facial recognition is enhanced and faster when these cues are congruent, while incongruent bodily cues bias facial expression recognition (Aviezer et al., 2011; Meeren et al., 2005; Mondloch et al., 2013; Reschke & Walle, 2021). Other studies have manipulated attentional focus of participants by instructing them to focus either on the face to recognize the expressed emotion (face focus condition), or on the body posture to recognize the emotion conveyed (body focus condition). These studies have shown a bidirectional interaction between the two emotional channels, that is, just as body expressions influence the recognition of facial emotions, facial expressions also affect the recognition of bodily emotions (Kret et al., 2013; Lecker et al., 2020). However, facial expressions seem to have a weaker influence on the recognition of bodily expressions than bodily expressions have on the recognition of facial expressions. According to Lecker et al. (2020), this asymmetry in the influence of these emotional signals is likely because individuals are naturally inclined to focus on facial expressions as the primary channel of emotional information. In everyday interactions, person tend to rely on the face to recognize emotion of others, while bodily expressions are typically processed as contextual cues. This default attentional bias facilitates the emergence of contextual effects when facial and bodily cues are combined. However, when participants are instructed to focus specifically on body posture, this strategy is less intuitive. As a result, bodily expressions are more likely to be interpreted in isolation, without being integrated into the broader emotional context (Lecker et al., 2020). These interactions, often described as the congruency effect (Lecker et al., 2020; Mondloch et al., 2013) or the contextual effect (Lecker et al., 2020), highlight the critical role of facial and bodily expression integration in shaping emotional recognition.

In the literature, it remains unclear whether the integration of facial and bodily expressions occurs automatically or requires voluntary effort. While some studies suggest automatic processing of emotional signal integration (Aviezer et al., 2011; Karaaslan et

al., 2020), others show that focusing attention on either facial or bodily expressions influences this integration, contradicting the hypothesis of automaticity (Lecker et al., 2020). However, the (lack of clear) definition of automaticity makes difficult to investigate this concept and any definitive conclusion is hard to reach as automaticity probably involves multiple aspects. In this regard, three criteria have been proposed to evaluate the automaticity of cognitive processes: efficiency, unconsciousness, and speed (Moors & De Houwer, 2006). While the criterion of speed has been less extensively explored, one study manipulated this criterion and suggested that even when participants had very limited time to respond and were exposed to stimuli for only a brief moment, bodily expressions still influenced the recognition of facial expressions. This suggests that the integration of facial and bodily expressions occurs rapidly (Meeren et al., 2005). Additionally, the unconsciousness criterion evaluates whether a process operates without requiring attention or awareness (Moors & De Houwer, 2006). Evidence suggests that even when facial and bodily expressions are presented subliminally (e.g., 33ms), bodily expressions influence the perception of facial expressions, indicating that this integration occurs outside conscious awareness (Karaaslan et al., 2020). The criterion of efficiency evaluates whether a process requires minimal cognitive or attentional resources (Moors & De Houwer, 2006). This is often tested using a dual-task protocol, where participants perform a primary task while simultaneously engaging in a secondary task that demands cognitive resources (e.g., a digit memorization task, target detection task, etc.). If the performance on the primary task remains unaffected by the cognitive load imposed by the secondary task, the process is considered efficient (Moors & De Houwer, 2006). To date, only one study has assessed the efficiency of integrating facial and bodily expressions (Aviezer et al., 2011). Participants were asked to recognize emotional facial expressions that were either congruent or incongruent with bodily expressions while performing a secondary task involving memorization of either a complex

sequence (e.g., 183K65) or a simpler sequence (just the letter, e.g., 183K65). By comparing high and low cognitive load conditions, researchers evaluated whether the contextual effect of bodily expressions on facial expression recognition was influenced by cognitive load. The findings suggests no impact of cognitive load on this effect, indicating that the integration process is efficient (Aviezer et al., 2011). Overall, these results suggest that the integration of facial and bodily expressions likely involves automatic processing, supported by evidence for both unconsciousness and efficiency.

Using electrophysiological methods, the temporal dynamics of emotional stimuli processing can be examined through event-related potentials (ERPs). Evidence suggests that the integration of facial and bodily expressions unfolds in three-stages (Gu et al., 2013). The first stage involves the automatic and rapid processing of bodily expressions, with a negativity bias favoring threatening stimuli. This is reflected by an increase of P100 amplitude, a positive component localized at occipital electrode sites, known to be sensitive to visual stimulations and their psychophysical parameters (such as color, visual field position, contrast, spatial frequencies, etc.), when fearful bodily expressions are perceived compared to happy ones during, combined face-body perception, regardless of whether attention is focused on the face or on the body (Gu et al., 2013). This early response suggests that bodily expressions are processed automatically, even without visual awareness. In contrast, the authors did not observe significant differences in P100 amplitude between fearful and happy facial expressions, suggesting that the negativity bias operates differently for facial and bodily emotional cues (Gu et al., 2013). The second stage is characterized by the rapid detection of incongruency between facial and bodily expressions (Gu et al., 2013). This process is reflected in increased N200 component amplitude, a negative component localized at fronto-central electrode sites, known to be sensitive to conflict detection (Bago et al., 2018), for incongruent stimuli compared to congruent ones (Chen et al., 2022; Gu et

al., 2013). Conflict detection can occur as early as 100ms, as indicated by the increase of P100 amplitude during incongruent conditions (Li, 2021; Meeren et al., 2005). These findings indicate that facial and bodily expressions are integrated quickly and automatically (Meeren et al., 2005). The final stage involves selective attention and detailed evaluation of emotional stimuli. At this stage, processing depends on the focus of attention. When attention is directed toward facial expressions, P300 amplitude is greater for congruent conditions than for incongruent ones (Gu et al., 2013; Li, 2021). Conversely, when attention is directed toward bodily expressions, the amplitude of P300, a positive component localized at centro-parietal electrode sites and sensitive to probability of occurrence of stimuli (Polich, 1990), but also to their affective congruence and emotional valence (Luo et al., 2010), is higher for fearful expressions than for happy ones, regardless of congruence (Gu et al., 2013). This suggests that attentional focus modulates the integration of facial and bodily expressions during later stages of processing (Gu et al., 2013; Lecker et al., 2020). In summary, the simultaneous perception of facial and bodily expressions leads to distinct early emotional processing compared to isolated facial (Luo et al., 2010; Smith et al., 2003) or bodily expressions (de Gelder et al., 2015). Bodily expressions appear to dominate early processing when both cues are presented together (Gu et al., 2013). Additionally, the rapid detection of incongruence highlights the automatic integration of facial and bodily expressions (Chen et al., 2022; Gu et al., 2013; Li, 2021; Meeren et al., 2005). At later stages, attentional focus further modulates emotional processing, as reflected in P300 variations (Gu et al., 2013).

All of these studies have highlighted the integration of facial and bodily expressions both behaviorally and at the neural level. However, it remains uncertain whether this integration meets the criterion of efficiency for automatic processing, particularly in relation to brain activity (Moors & De Houwer, 2006). While Aviezer et al. (2011) suggest automatic integration by manipulating cognitive load at the behavioral level, no study has yet investigated this process

using ERP components to examine the chronology of integration. Furthermore, the large majority of studies in this field are subject to several limitations: (1) First, the use of low-quality stimuli restricts the ecological validity of their findings and does not allow for a realistic assessment of emotional perception. (2) Second, their conclusions regarding the absence of significant effects rely on frequency analyses, which are not suitable for confirming the absence of effects. In contrast, Bayesian analyses can estimate the likelihood of an effect occurring (Lee & Wagenmakers, 2013; van den Bergh et al., 2020, 2023; van Doorn et al., 2021). (3) Moreover, many studies frame bodily expressions only as an emotional context that influences the recognition of facial emotions (Chen et al., 2022; Meeren et al., 2005), often overlooking the bidirectional nature of this contextual effect. Specifically, the influence of facial expressions on the recognition of bodily emotions remains understudied (Gu et al., 2013; Lecker et al., 2020; Zhang et al., 2019).

The objective of this study is to investigate the automaticity of integrating facial and bodily expressions, with a focus on the efficiency criterion (Moors & De Houwer, 2006). To achieve this, participants will perform a task involving the recognition of either facial or bodily expressions (i.e., face focus vs. body focus conditions) while simultaneously engaging in a digit memorization task (i.e., memorization of seven different or similar digits). The manipulation of attentional focus on either facial or bodily expressions is used to observe the bidirectionality of the contextual effect (Lecker et al., 2020) during the manipulation of automaticity criterion (high or low cognitive load). The attentional focus does not represent a criterion of automaticity. The study will utilize ERP component measures, high-quality congruent and incongruent facial and bodily expression stimuli (Puffet & Rigoulot, 2025), and Bayesian statistical analyses to address gaps identified in previous research. Three hypotheses will be tested using both behavioral and EEG measures.

- *Hypothesis 1:* we expect a classical contextual effect between facial and bodily cues, i.e., facilitated emotion recognition in the congruent condition, and hindered recognition in the incongruent condition. At behavioral level, we should observe higher accuracy and faster response times for emotionally congruent pairs compared to incongruent expressions (Gu et al., 2013; Lecker et al., 2020). At the brain level, we expect incongruent expressions to elicit larger amplitudes in early ERP components (P100, N100, P200, N200) reflecting an early detection of conflict (Gu et al., 2013; Meeren et al., 2005).
- *Hypothesis 2:* we hypothesize that attentional focus/task instructions will modulate the contextual effect between facial and bodily cues (Lecker et al., 2020). We expect that the influence of facial expressions on the recognition of bodily expressions (in the body focus condition) will be less pronounced than the influence of bodily expressions on the recognition of facial emotions (in the face focus condition). In terms of ERP components, we expect that when attention is directed to the face, congruent expressions may be associated with enhanced P300 amplitudes, consistent with facilitated extraction of emotional information and more efficient allocation of sustained attention (Gu et al., 2013; Li, 2021).
- *Hypothesis 3:* we hypothesize that there will be no influence of cognitive load on other factors (Aviezer et al., 2011), suggesting the automaticity of these processes. It should lead to a high level of confidence in the alternative hypothesis, i.e., no differences in the contextual effect with respect to cognitive load for behavioral (accuracy of recognition and reaction times) and brain measures (amplitudes of ERPs of interest, P100, N100, P200, N200).

Methods

Participants

The sample size for this study was estimated using the *SuperPower* package (version 0.2.0; power = 0.95, standard deviation = 2) in R Studio (version 4.3.0; (Lakens & Caldwell, 2021). Thirty-five right-handed participants (25 females; $M_{\text{age}} = 26 \pm 5$ years; range 18-35 years) were recruited between June 2023 and November 2023 through social networks and a specific University platform for advertising experiments. Each participant received CAD 30 as compensation for participating in a three-hour session. The inclusion criterion was an age range of 18 to 35 years. Exclusion criteria included neurological disorders (e.g., traumatic brain injury), diagnosed psychological conditions (e.g., psychosis), uncorrected visual or auditory impairments, and the use of medications or substances (e.g., drugs) affecting the nervous system.

Participants completed *The Hospital Anxiety and Depression Scal* (HADS; Means of Anxiety score = 6.77 ± 3.3 ; Means of depression score = 2.94 ± 1.97) to assess mood (Zigmond & Snaith, 1983). This screening was run because anxiety or depression tendencies influence emotion processing (Peschard et al., 2014). Two participants were excluded due to high HADS anxiety scores (17 and 12, respectively; (Zigmond & Snaith, 1983).

Demographic information, including date of birth, gender, dominant hand, native language, and education level, was also collected (see <https://osf.io/5at6g/> for details). All participants provided their informed consent. Moreover, the study was approved by the UQTR ethics committee (certificate number: CERPPE-22-06-07.03) and all methods were performed in accordance with the relevant guidelines and regulations.

Materials

Stimuli were sourced from the “Validation of the Emotionally Congruent and Incongruent Face-Body Static Set” (ECIFBSS; (Puffet & Rigoulot, 2025). This dataset includes 1952 images

of facial and bodily expressions (49 images/actors) presented in both congruent and incongruent situations. For this study, we selected 84 stimuli: 42 congruent (14 images per emotion: happiness, anger, and sadness) and 42 incongruent (seven images for each incongruent combination of happiness, anger, and sadness). We selected the images so that facial and bodily expressions had a relatively high and similar average recognition accuracy (94% for facial expressions and 92% for bodily expressions; OSF; <https://osf.io/5at6g/>). The emotions of anger, happiness, and sadness have been selected because they are commonly employed in studies on contextual effects between facial and body cues and have been shown to elicit such effects (Abo Foul et al., 2018; Kret et al., 2013).

The ECIFBSS provided the evaluation of intensity, valence, and arousal (see Table 1). Multiple ANOVAs were performed to explore differences in these emotional dimensions (valence, arousal, and intensity) as a function of emotion (happiness, sadness, and anger), congruence (congruent vs. incongruent), and mode of expression (facial vs. bodily expressions).

Table 1. Mean and Standard Deviation of Valence, Arousal and Intensity for Facial and Bodily Emotions

	Valence	Arousal	Intensity
Facial emotion			
Happiness	6.73 (+/- 0.68)	5.52 (+/- 0.58)	5.90 (+/- 1.06)
Anger	3.55 (+/- 0.34)	5.12 (+/- 0.45)	4.95 (+/- 1.08)
Sadness	3.77 (+/- 0.24)	4.44 (+/- 0.32)	4.18 (+/- 0.63)
Happiness	6.53 (+/- 0.42)	5.58 (+/- 0.52)	6.45 (+/- 0.90)
Bodily emotion			
Anger	3.57 (+/- 0.31)	5.38 (+/- 0.37)	5.62 (+/- 0.72)
Sadness	3.93 (+/- 0.34)	4.38 (+/- 0.31)	4.14 (+/- 0.76)

In terms of valence, the ANOVA revealed a significant main effect of emotion ($F[2, 156] = 948.348, p < 0.001, \eta^2 = 0.918$), and a three-way interaction between emotion, mode of expression and, congruence ($F[2, 156] = 3.346, p = 0.038, \eta^2 = 0.003$). Bonferroni-adjusted paired t -tests showed that happiness was perceived as more positive than anger ($t[156] = -39.411, p < 0.001, d = -7.448, CI [-3.255, -2.886]$), and sadness ($t[156] = 35.755, p < 0.001, d = 6.757, CI [2.601, 2.970]$). Given the significant three-way interaction between emotion, congruence, and mode of expression, follow-up ANOVAs were conducted separately for each emotion to further examine the interaction between congruence and mode of expression. This analysis for sadness showed a significant main effect of mode of expression ($F[1, 52] = 4.106, p = 0.048, \eta^2 = 0.069$). Bonferroni-adjusted paired t -tests showed that sad faces were perceived as more negative than sad bodies ($t[52] = 2.026, p = 0.048, d = 0.542, CI [0.002, 0.315]$). No other significant effects were found ($ps > 0.05$).

In terms of arousal, the ANOVA showed a significant main effect of emotion ($F[2, 156] = 108.979, p < 0.001, \eta^2 = 0.552$) as well as a Congruence * Mode of expression interaction ($F[1, 156] = 9.668, p = 0.002, \eta^2 = 0.024$). Bonferroni-adjusted paired t -tests showed that happiness ($t[156] = 14.239, p < 0.001, d = 2.691, CI [0.952, 1.332]$) and anger

($t[156] = 0.496, p < 0.001, d = 1.984, CI [0.652, 1.031]$) were perceived as more arousing than sadness. Simple effects analysis also showed a significant effect of mode of expression during incongruent condition ($F[1, 156] = 9.714, p = 0.002$) and a significant effect of congruence when emotions were expressed by the body ($F[1, 156] = 4.902, p = 0.028$) and by the face ($F[1, 156] = 4.766, p = 0.031$). No other significant effects were found ($ps > 0.05$).

In terms of intensity, the ANOVA revealed a significant main effect of emotion ($F[2, 156] = 76.604, p < 0.001, \eta^2 = 0.459$), and mode of expression ($F[1, 156] = 8.667, p = 0.004, \eta^2 = 0.026$), as well as Congruence * Mode of expression interaction ($F[1, 156] = 7.716, p = 0.006, \eta^2 = 0.023$). Bonferroni-adjusted paired t -tests showed that happiness was rated as more intense than anger ($t[156] = -5.464, p < 0.001, d = -1.033, CI [-1.277, -0.505]$), and anger more intense than sadness ($t[156] = 6.886, p < 0.001, d = 1.301, CI [0.737, 1.509]$). Moreover, bodily expressions were rated as more intense than facial expressions ($t[156] = 2.944, p = 0.004, d = 0.454, CI [0.129, 0.665]$). Simple effects analysis also showed a significant effect of mode of expression in the incongruent condition ($F[1, 156] = 16.369, p < 0.001$), and a significant effect of congruence when emotions were expressed by the body ($F[1, 156] = 5.182, p = 0.024$). No other significant effects were found ($ps > 0.05$).

Given that brightness and contrast of images influence ERP components amplitudes (Puce et al., 2013), we controlled these variables across stimuli (OSF; <https://osf.io/5at6g/>). Brightness and contrast values were measured using ImageJ software. The mean brightness was 232.67 (+/- 2.13) for congruent conditions and 232.03 (+/- 2.13) for incongruent conditions. The mean of contrast was 62 (+/- 3.9) for congruent conditions and 63.13 (+/- 3.33) for incongruent conditions. Two one-way ANOVAs were conducted to examine potential differences in brightness and contrast between congruent and incongruent conditions. Results could not reveal significant effect of congruency on either brightness or contrast ($ps > 0.154$).

Finally, we used a set of 12 different and 12 identical digits to create sequences of seven random numbers. Two parameters were enforced during sequence generation: (1) no consecutive identical numbers were included (e.g., 4497425), and (2) ascending (e.g., 1234567) or descending (e.g., 7654321) sequences were avoided (OSF; <https://osf.io/5at6g/>). Given that this method differs slightly from previous studies (Aviezer et al., 2011; Lima et al., 2019; Tracy & Robins, 2008), we conducted several pilot tests to ensure that our methodology effectively manipulated cognitive load of participants. These pilot tests helped us determine the best approach for our study.

Experimental Design

The experimental design followed a within-subject approach, manipulating four independent variables: attentional focus (on face emotion or bodily expression), cognitive load (high or low cognitive load), congruency (congruent or incongruent between facial and bodily expressions), and type of emotional expression (happiness, sadness, and anger). The study consisted of a total of 672 trials, evenly divided into 336 trials per attentional focus condition (on face emotion or bodily expression). Within each attentional focus condition, there were 168 trials per congruency condition (congruency and incongruency), 168 trials per cognitive load condition (high and low), and 28 trials per emotion (happiness, sadness, or anger) in congruent situations, and 14 trials per pairs of incongruent emotions (e.g., angry face with sadness body).

General Procedure

Prior to the experiment, participants completed three online questionnaires via LimeSurvey software: a socio-demographic questionnaire, the HADS (Zigmond & Snaith, 1983), and an informed consent and information questionnaire.

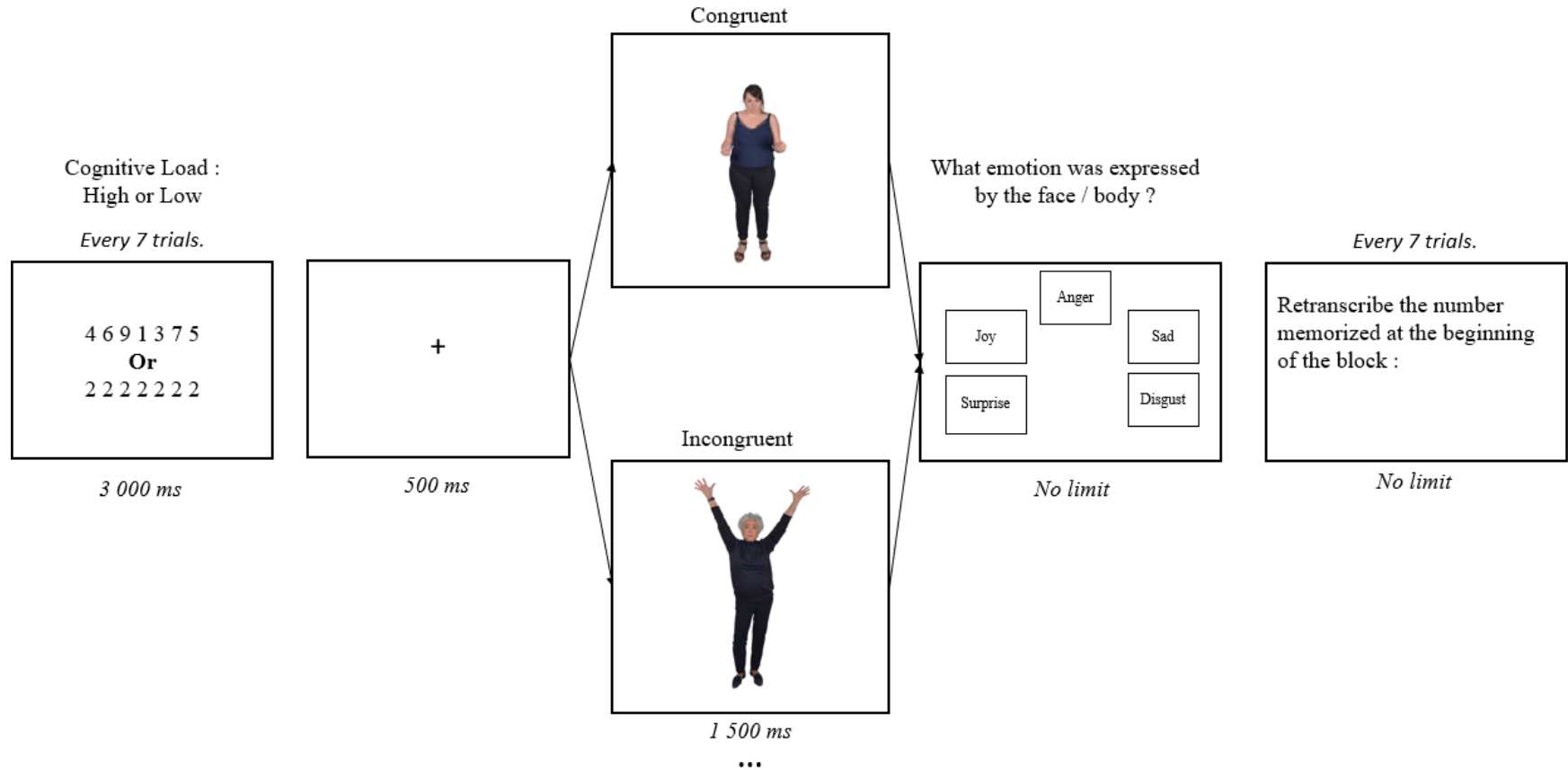
The experiment lasted approximately three hours, including one hour for the task and two hours for encephalography (EEG) equipment setup, control of impedances, and quality of signal.

It was conducted in a dedicated room equipped with an EEG system at the University of Quebec at Trois-Rivières (UQTR). Participants were seated comfortably on a chair inside a Faraday cage at 60 cm from a computer screen. The experiment task was programmed using E PRIME 2.0 software.

The task began with on-screen instructions, followed by two identical parts, counterbalanced in order across participants. In one part, participants focused on recognizing facial expressions, while in the other, they focused on bodily expressions. Each part started with a training session of two blocks, followed by an experimental session consisting of 48 randomized blocks. At the beginning of each block, participants were presented with a sequence of seven digits to memorize, displayed for 3000ms at the center of the screen (“Courier New” font; bold; point size 25). These sequences were either different (e.g., 4916274; representing high cognition load) or identical (e.g., 1111111; representing low cognition load). Following this, seven congruent and incongruent images were shown consecutively for 1500ms each. Each image was preceded by a 500ms fixation cross. Participants were instructed to identify the emotion expressed in either the facial or bodily expressions based on their initial impression, without time constraints. Participants had to choose, as quickly as possible, among five emotion labels: the three target emotions (anger, happiness, and sadness), plus two extra choices, disgust and surprise. Disgust was included because it is often confused with anger in incongruent face-body pairings (Aviezer et al., 2008), which can amplify contextual effects. Surprise was also included as it is often misperceived as happiness in both facial and bodily expressions (Puffet & Rigoulot, 2025), thereby also potentially amplifying contextual effect between these emotions. Given that surprise is an ambiguous emotion, which could be evaluated as positive or negative, it is possible that this choice might introduce variability and potentially affect the experimental outcomes. However, we believe that this influence would be minimal given that surprise was not among our

target emotions but only a supplementary choice of response. After completing the emotion recognition task for all seven images in the block, participants transcribed the digit sequence they had memorized at the beginning of the block. EEG and behavioral data (accuracy and response time) were recorded simultaneously throughout the experiment (see Figure 1).

Figure 1. *Schema of Block Structure*



Note. The block was constituted of the presentation of digits, a fixation cross, a picture of emotionally congruent or incongruent face and body expressions, a prompt for emotional identification. Every seven trials, participant had to write the digits they had memorized.

Statistics Analysis

The behavioral and EEG data were analyzed using MATLAB (R2021a), JASP (version 0.18.1; RRID:SCR_015823), and R (version 4.3.0). The package ggplot2 (version 3.4.2; RRID:SCR_014601; (Wickham, 2016), tidyverse (version 1.2.0; RRID:SCR_019186; (Wickham et al., 2019), dplyr (version 1.1.2; RRID :SCR_017102; (Wickham et al., 2023), and emmeans (version 1.8.7; RRID :SCR_018734; (Russell et al., 2021) were used. The analyses conducted are available on the OSF platform (<https://osf.io/5at6g/>).

Behavioral Analysis

First, we analyzed data from the memorization task. A paired *t*-test and a Bayesian *t*-test were performed on digit memorization scores to examine differences between cognitive load levels. Additionally, a Pearson correlation was conducted between digit memorization and emotion recognition accuracy to assess their association. Further analyses were conducted on trials where participants achieved a 100%-digit memorization rate. Trials with incorrect digit memorization were excluded from further analysis.

Secondly, we analyzed the emotion recognition task. Behavioral data were assessed using two metrics: the accuracy rate of correct categorization and response time. For accuracy, we calculated the *Unbiased hit rates* (Hu score; (Wagner, 1993) for emotion recognition. This metric adjusts for potential response bias that could inflate accuracy score. For instance, a participant who categorizes all facial expressions as “anger” may achieve a high accuracy score for anger, but this does not indicate true recognition. The Hu score corrects this bias by considering the squared frequency of correct responses for a target emotion, divided by the product of the total number of stimuli representing that emotion and the overall frequency of selecting that emotion category (Wagner, 1993). To identify and exclude outliers in both Hu scores and response times, we used the *Mean Absolute Deviation* (MAD) function in R (Leys et al., 2013, 2019).

We conducted classical repeated measures ANOVAs on the arcsine-transformed Hu score as recommended in the literature (Goeleven et al., 2008; Wagner, 1993) and response times, incorporating three factors: congruency, attentional focus, and cognitive load. Since classical repeated measures ANOVAs do not allow for direct support of the null hypothesis (i.e., no difference between conditions), we complemented these with repeated measures Bayesian ANOVAs using the same factors. For Bayesian analyses, we used default prior options for effects ($r = 0.5$ for the fixed effects) as recommended (van den Bergh et al., 2020; van Doorn et al., 2021). The Bayesian ANOVAs produced two types of models: the null model and the alternative model. Alternative models included one or more factors and interactions between them. JASP software calculated Bayes factors (BF₁₀) for each model, quantifying the strength of evidence in favor of each alternative model relative to the null model or the best-performing model (van den Bergh et al., 2023). In other words, a BF₁₀ ratio indicates how much more likely the observed data are under an alternative model (i.e., the most likely model) compared to the null model. Evidence for the null hypothesis is indicated by BF₁₀ values between 1 and 1/3 (weak evidence), between 1/3 and 1/10 (moderate evidence), and below 1/10 (strong evidence). Conversely, B₁₀ values between 1 and 3 (weak evidence), 3 and 10 (moderate evidence), and above 10 (strong evidence) support the alternative hypothesis (Lee & Wagenmakers, 2013; van Doorn et al., 2021). For example, a BF₁₀ of 30 indicates that the data are thirty times more likely to occur under the alternative model than the null model (van den Bergh et al., 2020). Additionally, we calculated *inclusion Bayes factor* (BF_{inclusion}) for each factor (congruency, focus, and cognitive load) to quantify the evidence for including these factors and their interactions within the set of models. This approach allowed us to assess the predictive value of each factor in the dataset (van den Bergh et al., 2020).

ERP Recording, Preprocessing, and Analysis

BrainVision software was used to record EEG data. EEG activity was recorded from 64 scalp electrodes, with Fpz serving as the ground electrode. Data were sampled at a rate of 500Hz per channel for offline analysis. Additionally, the vertical electrooculogram (VEOG) and horizontal electrooculogram (HEOG) were recorded using three electrodes: below the left eye (Fp2), and on the left (FT9) and right (FT10) sides of the eyes. Two electrodes (TP9 and TP10) were placed on the earlobes as reference electrodes. All electrode impedances were kept below 25 k Ω .

Data pre-processing was performed using the EEGLab (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes in MATLAB, following the pipeline outlined by Lopez-Calderon and Luck (Lopez-Calderon & Luck, 2014). The data were re-referenced offline to the average and filtered with a 0.1 Hz high-pass filter. Independent Component Analysis (ICA) was performed by the first author (ASP) using AMICA (Palmer et al., 2012). The signals were segmented into epochs ([-200ms; 700ms]) around the target stimuli. A low-pass filter with a cutoff at 30 Hz was applied, followed by two rounds of artifact detection. The first round used the “moving window peak-to-peak threshold” method from ERPLAB, with a threshold of 100uV, a window size of 200ms, and a window step of 100ms. The second round targeted eye movements at electrodes 5, 25, and 30, using the “step-like artifacts” detection method, with a threshold of 30uV, a window size of 200ms, and a window step of 50ms. After artifact rejection, an average of 96.5% of data was retained for further analysis (Face focus: 97.57%; body focus: 95.46%; Congruency: 96.72%; Incongruency: 96.42%; High cognitive load: 96.53%; Low cognitive load: 96.52%). As in the behavioral analyses, EEG analyses were only conducted on trials in which participants achieved a 100%-digit memorization rate. Trials with incorrect digit memorization were excluded. On average, 77.71% of the data ($M = 539.5$ trials)

was retained for analysis (Face focus: 77.91% [$M = 268.25$ trials]; body focus: 77.1% [$M = 271.25$ trials]; Congruency: 78.35% [$M = 271.46$ trials]; Incongruency: 77.14% [$M = 268.04$ trials]; High cognitive load: 65.64% [$M = 227.25$ trials]; Low cognitive load: 89.91% [$M = 312.25$ trials]).

The Factorial Mass Univariate Toolbox (FMUT) was used to analyze the EEG data (Fields & Kuperberg, 2020). This analysis involves iteratively conducting thousands of inferential statistical tests and applying multiple comparison corrections across all electrodes within a temporal window of interest (Durston & Itier, 2021; Fields & Kuperberg, 2020; Hudson et al., 2021; Jaspers-Fayer et al., 2022; McCrackin & Itier, 2021). Mass univariate analyses provide an exploratory approach to identifying effects without requiring *a priori* assumptions and offer greater power for detecting effects compared to traditional spatiotemporal averaging methods (Fields & Kuperberg, 2020; Groppe et al., 2011).

We first conducted an exploratory mass univariate analysis on all electrodes and time points from -100ms pre-stimulus to the end of the epoch (700ms). Based on this exploratory analysis and previous work (Gu et al., 2013; Meeren et al., 2005; van Dillen & Derks, 2012; Zhang et al., 2015), we identified two early time windows of interest at fronto-central sites (F1, F2, F3, F5, F6, FC1, FC2, FC3, FC5, Fz, and FCz) during the P100 (70-140ms) and N250 (150 – 318ms) and at parieto-occipital sites (PO7, PO8, PO4, PO3, P5, P6, P7, P8, O1, O2, and Oz) during the P100 (70-140ms) and P250 (150 – 318ms). We also identified one later time window in the centro-parietal (Cz, C1, C2, C3, C4, CPz, CP1, CP2, CP3, CP4, Pz, P1, P2, P3, and P4) and occipital sites (O1, O2, and OZ; 200 – 700ms).

Mass univariate analysis used an alpha level of 0.05 and was performed with three within-subjects factors: cognitive load, focus, and congruence. In FMUT, ANOVAs were conducted using 100,000 permutations for each data point, with correction for multiple comparisons via

permutation test, as recommended (Durstun & Itier, 2021; Fields & Kuperberg, 2020; Hudson et al., 2021). Finally, we also conducted repeated-measures Bayesian ANOVAs on the amplitude of components averaged over the same time windows and electrode sites as the FMUT analysis, using the three factors: cognitive load, focus, and congruence.

Results

The behavioral and EEG analyses included 28 participants (22 females; $M_{\text{age}} = 27 \pm 5$ years; range 18-35 years) with mean HADS score of 6.68 (± 3.45) for anxiety and 3.04 (± 1.84) for depression because five participants were excluded from the data analysis due to a low number of successful trials.

Memorization Task

The mean digit memorization accuracy was 0.803 (± 0.36), with 0.929 (± 0.074) for the low cognition load condition and 0.676 (± 0.185) for the high cognition load condition.

A paired-samples *t*-test on accuracy revealed a significant difference between high and low cognitive load ($t[27] = -7.065$, $p < 0.001$, $d = -1.335$, CI [-0.326, -0.180]). Digits were better memorized in the low cognitive load condition compared to the high cognitive load condition. Similarly, a Bayesian paired-samples *t*-test on accuracy provided strong evidence in favor of H1 ($\text{BF}_{10} = 113651.451$), indicating a high likelihood that digits in the low cognitive load condition were better retained than those in the high cognitive load condition.

Pearson correlations were computed between memorization accuracy and emotion recognition accuracy to assess their association. The analysis revealed no significant correlation ($r = 0.008$, $p = 0.281$).

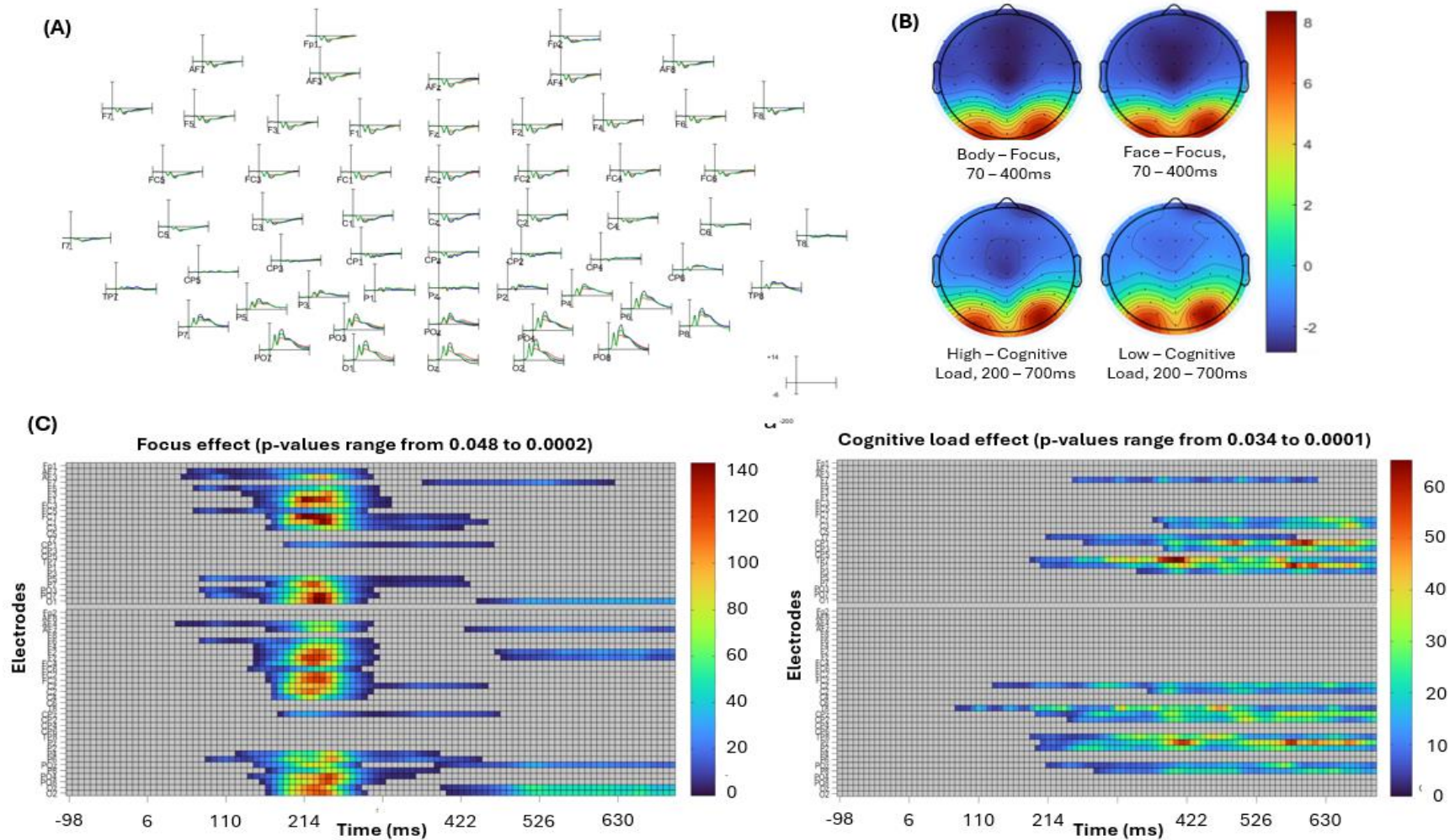
Preliminary Analyses: Behavioral and EEG Data

We used the MAD function (Leys et al., 2013, 2019) to identify outlier trials and participants. For Hu scores, no participants or trials were excluded due to Hu score outliers. For

reaction times, no participants and 11.4% of the trials were excluded due to response time outliers.

For ERP data, the exploratory analysis conducted with FMUT across all electrodes within the time window of -100ms to 700ms revealed a significant effect of focus at fronto-central (F1, F2, F3, F5, F6, FC1, FC2, FC3, FC5, Fz, and FCz) and parieto-occipital (PO7, PO8, PO4, PO3, P5, P6, P7, P8, O1, O2, and Oz) sites between 70ms to 300ms, and at occipital sites (O1, O2, and OZ) from 400 to 700ms, with a sustained potential around 500ms (p-values ranged from 0.048 to 0.0002). Additionally, a significant effect of cognitive load was found at the centro-parietal (Cz, C1, C2, C3, C4, CPz, CP1, CP2, CP3, CP4, Pz, P1, P2, P3, and P4) sites between 200ms and 700ms, with a sustained potential around 500ms (p-values ranged from 0.034 to 0.0001). No other effects were significant ($ps > 0.05$; see Figure 2).

Figure 2. Plot ERP Waveforms of All Electrodes (A), Grand Averaged ERP for Poz and Oz Electrodes (B), and FMUT map of Focus and Cognitive Load Effects (C)



Note. For grand averaged ERPs for all electrodes (A) the cerebral activity was elicited between -100 and 700ms post-stimulus. For C graph: coloured sections correspond to the significant F-values as indicated by the colour bar.

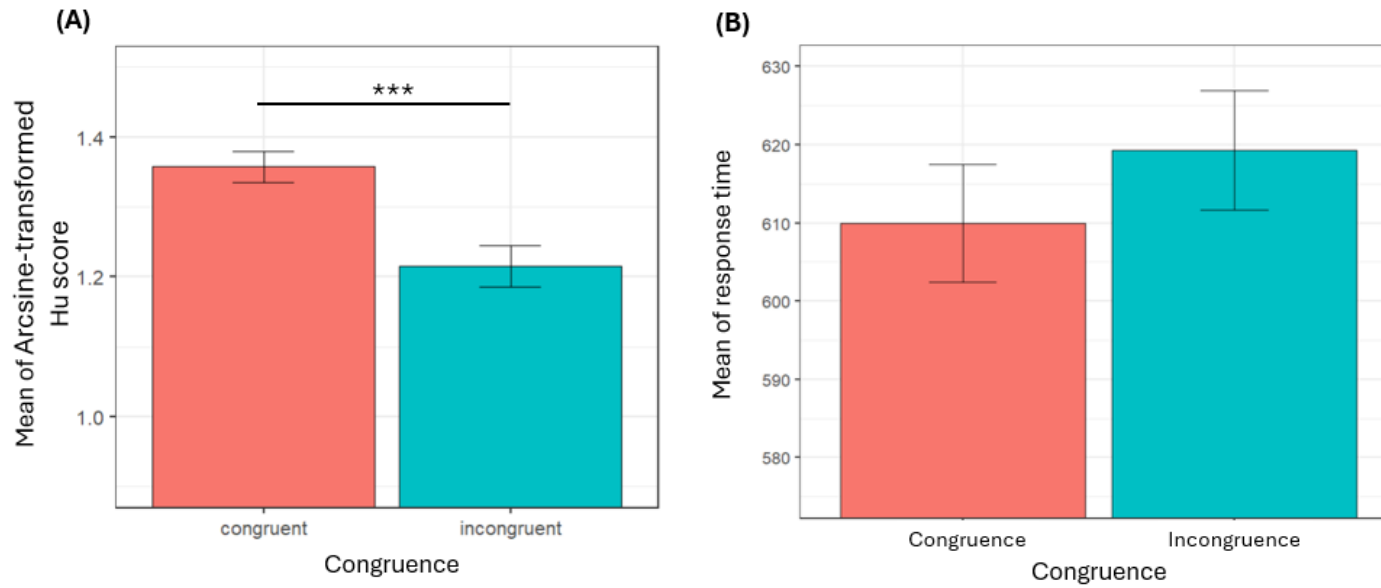
Based on these exploratory analyses, more specific analyses using the FMUT were conducted in specific scalp zones and time windows (see section about processing of ERP data). At the fronto-central sites, the N100 and N250 time-windows were examined. At the parieto-occipital sites, P100 and P250 time-windows were analyzed. Finally, the occipital sites were analyzed during a late-stage time-window.

Hypothesis 1: Contextual Effect

The 2 (Focus) * 2 (Cognitive Load) * 2 (Congruence) repeated measures ANOVA on the arcsine-transformed Hu scores revealed a significant main effect of congruence ($F[1, 27] = 82.452, p < 0.001, \eta^2 = 0.307$). Bonferroni-adjusted paired t -tests on congruence main effect showed that emotions were better recognized during congruent conditions rather than incongruent conditions ($t[27] = 9.08, p < 0.001, d = 0.848, CI [0.110, 0.174]$; see Figure 3). The Bayesian repeated measures ANOVA indicated strong evidence that the best model, i.e., the one with the largest Bayes Factor, includes the main effect of congruence, focus, and the Focus * Congruence interaction ($BF_{10} = 9.281e+7$). $BF_{inclusion}$ showed strong evidence for including the congruence factor ($BF_{inclusion} = 1.159e+7$).

The same repeated measures ANOVA on reaction times did not show any effect related to congruency ($p = 0.095$). Bayesian analyses provided moderate evidence against the inclusion of the congruence factor ($BF_{inclusion} = 0.08$; see Figure 3).

Figure 3. Histogram of Means of Arcsine Transformation (A) and Response Time (B) for the Congruence



Note. *** $p < 0.001$.

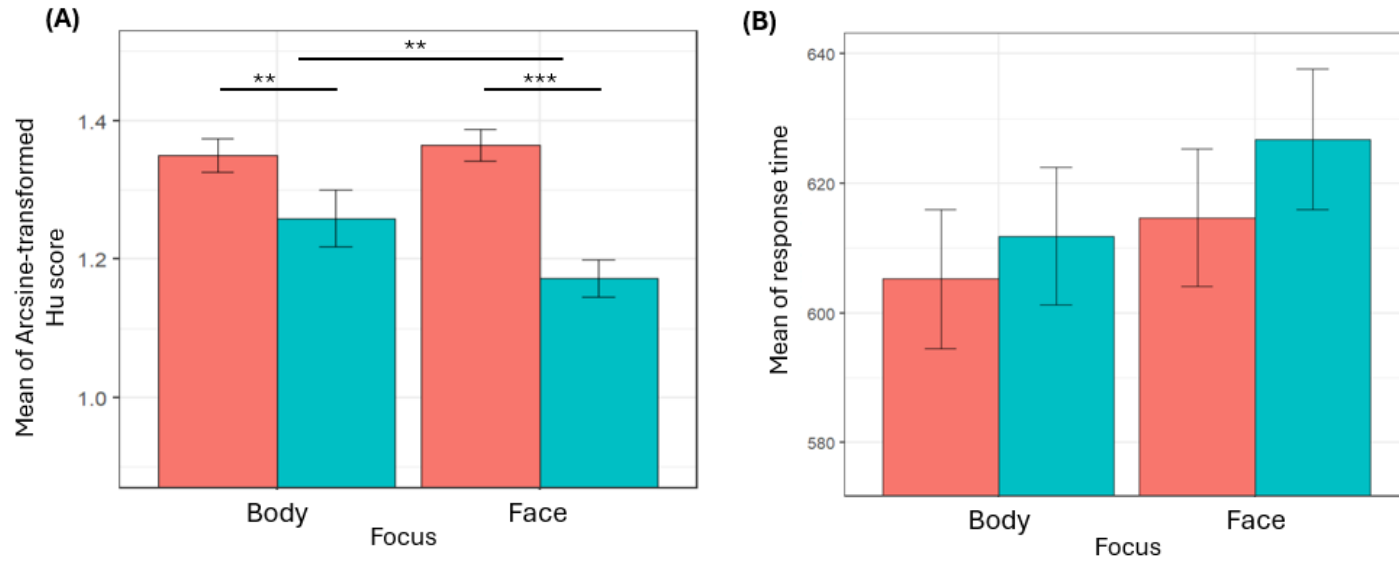
Finally, analyses FMUT on ERPs did not show an effect of congruence ($ps > 0.065$) and Bayesian analyses revealed weak evidence against the inclusion of the congruence factor across ERP components (*N100*: $BF_{inclusion} = 0.243$; *P100*: $BF_{inclusion} = 0.256$; *N250*: $BF_{inclusion} = 0.458$; *P250*: $BF_{inclusion} = 0.387$; *The sustained potential*: $BF_{inclusion} = 0.249$).

Hypothesis 2: Attentional Focus Effect

The repeated measures ANOVA on arcsine transformation to the Hu Score revealed a significant Focus * Congruence interaction ($F[1, 27] = 8.46, p = 0.007, \eta^2 = 0.039$). No main effect of focus was found ($p = 0.110$). Simple effects analysis on focus and congruence interaction showed a significant effect of focus during incongruence conditions ($F[1, 27] = 5.870, p = 0.022$). The simple effects analysis also showed a significant effect of congruence when participants focused on the face ($F[1, 27] = 71.603, p < 0.001$) and on the body posture ($F[1, 27] = 14.569, p < 0.001$; see Figure 4). The Bayesian analyses revealed moderate evidence for including the Congruence * Focus interaction ($BF_{inclusion} = 8.381$), and weak evidence against the inclusion of the focus factor ($BF_{inclusion} = 0.923$).

For reaction times, the repeated measures ANOVA did not show any effect related to main effect of focus ($p = 0.578$) and Congruence * Focus interaction ($p = 0.571$). The Bayesian analyses indicated strong evidence that the best model includes the main effects of cognitive load and focus ($BF_{10} = 19.340$). The Bayesian analyses revealed weak evidence against the inclusion of the focus factor ($BF_{inclusion} = 0.601$), and the Focus * Congruence interaction ($BF_{inclusion} = 0.601$; see Figure 4).

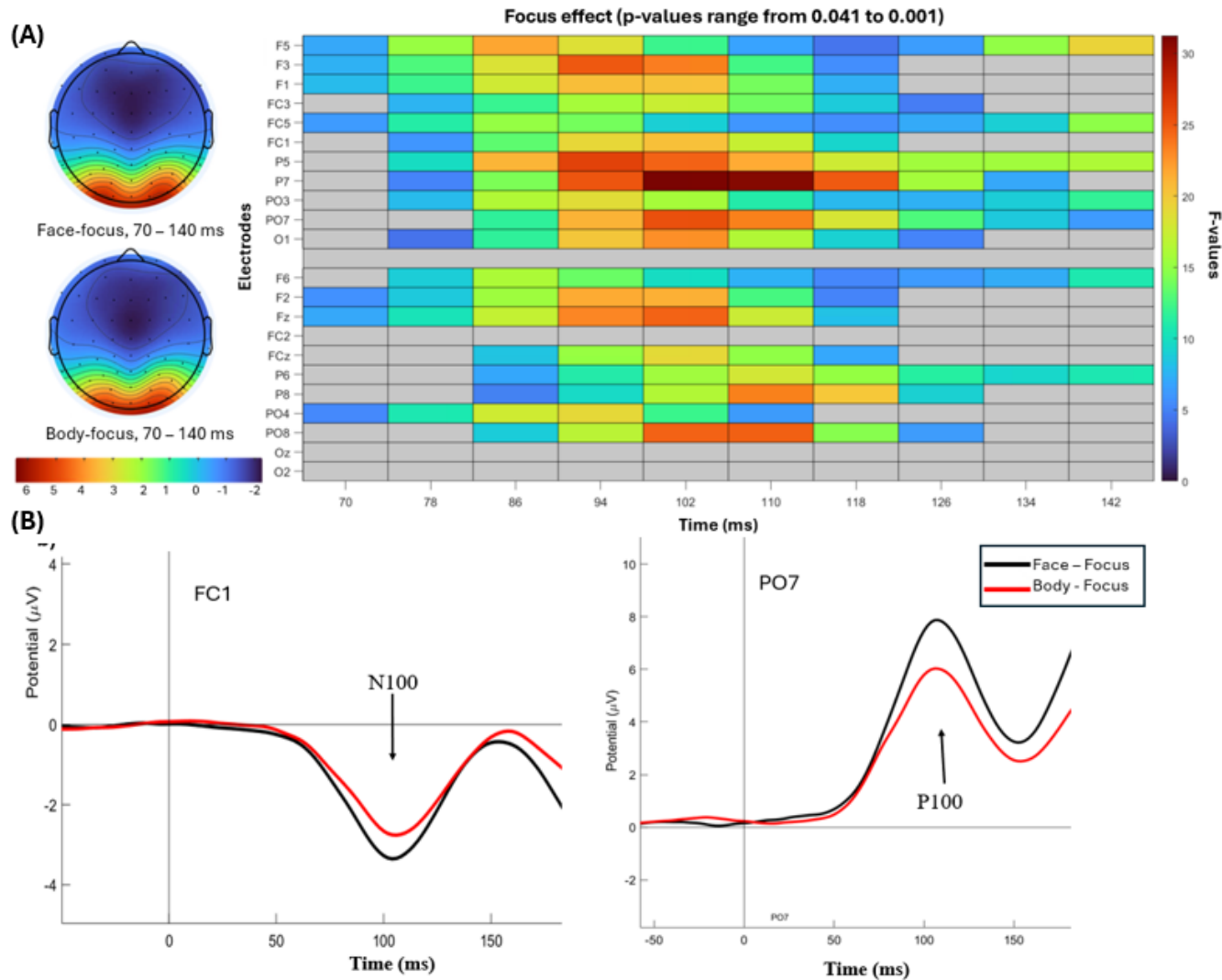
Figure 4. Histogram of Means of Arcsine Transformation (A) and Response Time (B) for Interaction between Congruence and Focus Type



Note. ** $p < 0.01$. *** $p < 0.001$.

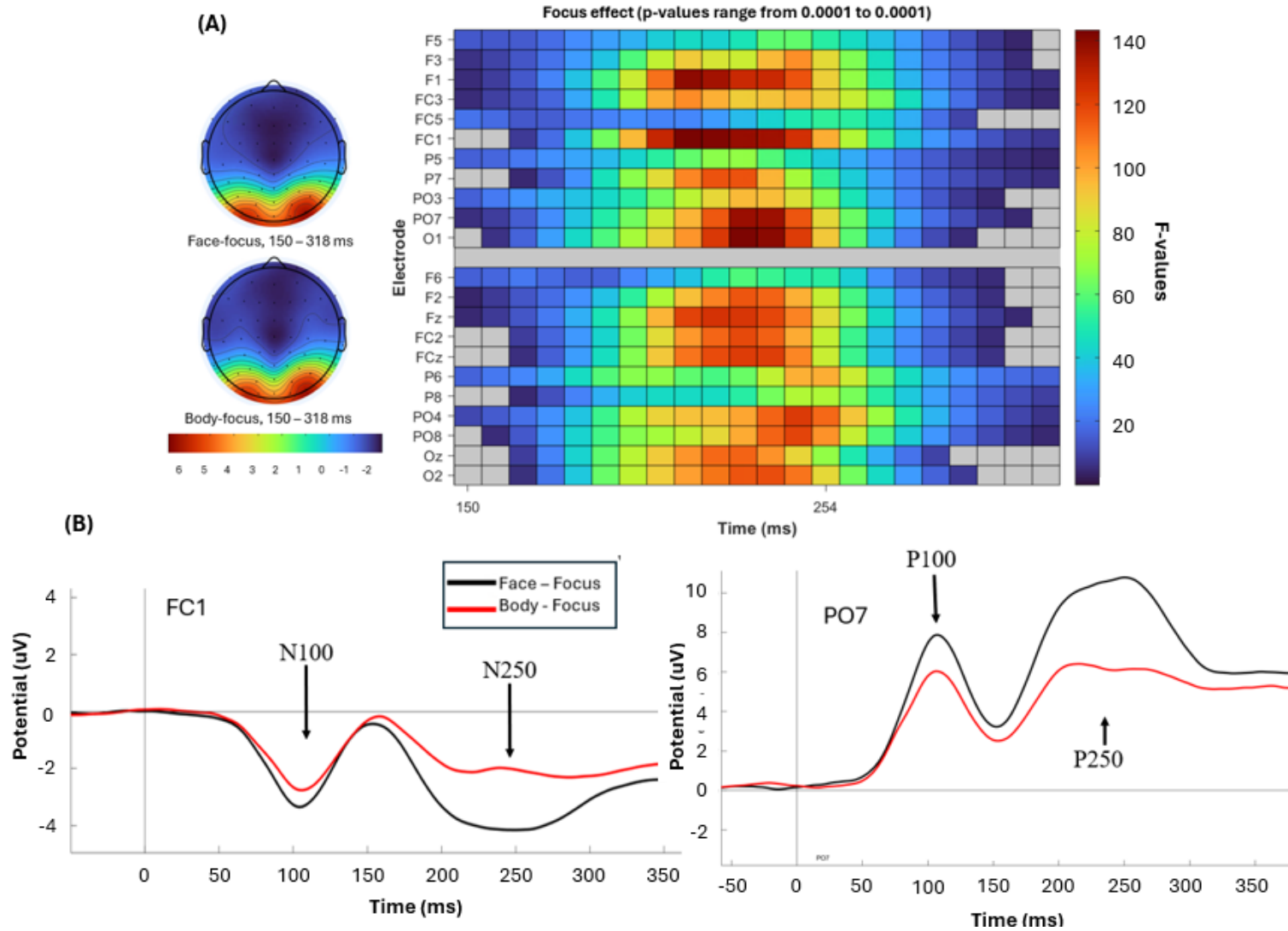
Finally, analyses FMUT on ERPs found a significant focus effect ($ps < 0.041$) at the fronto-central sites (F1, F2, F3, F5, F6, FC1, FC2, FC3, FC5, Fz, and FCz) for the N100 component, at the parieto-occipital sites (PO7, PO8, PO4, PO3, P5, P6, P7, P8, O1, O2, and Oz) for the P100 component, at the fronto-central sites (F1, F2, F3, F5, F6, FC1, FC2, FC3, FC5, Fz, and FCz) for the N250 component, at the parieto-occipital sites (PO7, PO8, PO4, PO3, P5, P6, P7, P8, O1, O2, and Oz) for the P250 component, and at the occipital sites (O1, O2, and Oz) for a sustained potential during the late-stage time-window (400-700ms). No other significant effects were observed ($ps > 0.065$; see Figures 5, 6, and 7). The Bayesian repeated measures ANOVA indicated strong evidence that the best model for the N100, P100, P250, N250, and the sustained potential (after 400ms post-stimuli) includes the main effect of focus (*N100*: $BF_{10} = 114.105$; *P100*: $BF_{10} = 64.880$; *N250*: $BF_{10} = 6.985e+7$; *P250*: $BF_{10} = 1.407e+7$; *The sustained potential*: $BF_{10} = 3941.140$). The analyses revealed strong evidence for including the focus factor (*N100*: $BF_{inclusion} = 115.267$; *P100*: $BF_{inclusion} = 66.916$; *N250*: $BF_{inclusion} = 6.966e+7$; *P250*: $BF_{inclusion} = 1.333e+7$; *The sustained potential*: $BF_{inclusion} = 3815.977$).

Figure 5. Scalp map, FMUT map and Grand Averaged ERP on the Fronto-Central, and Parieto-Occipital Sites



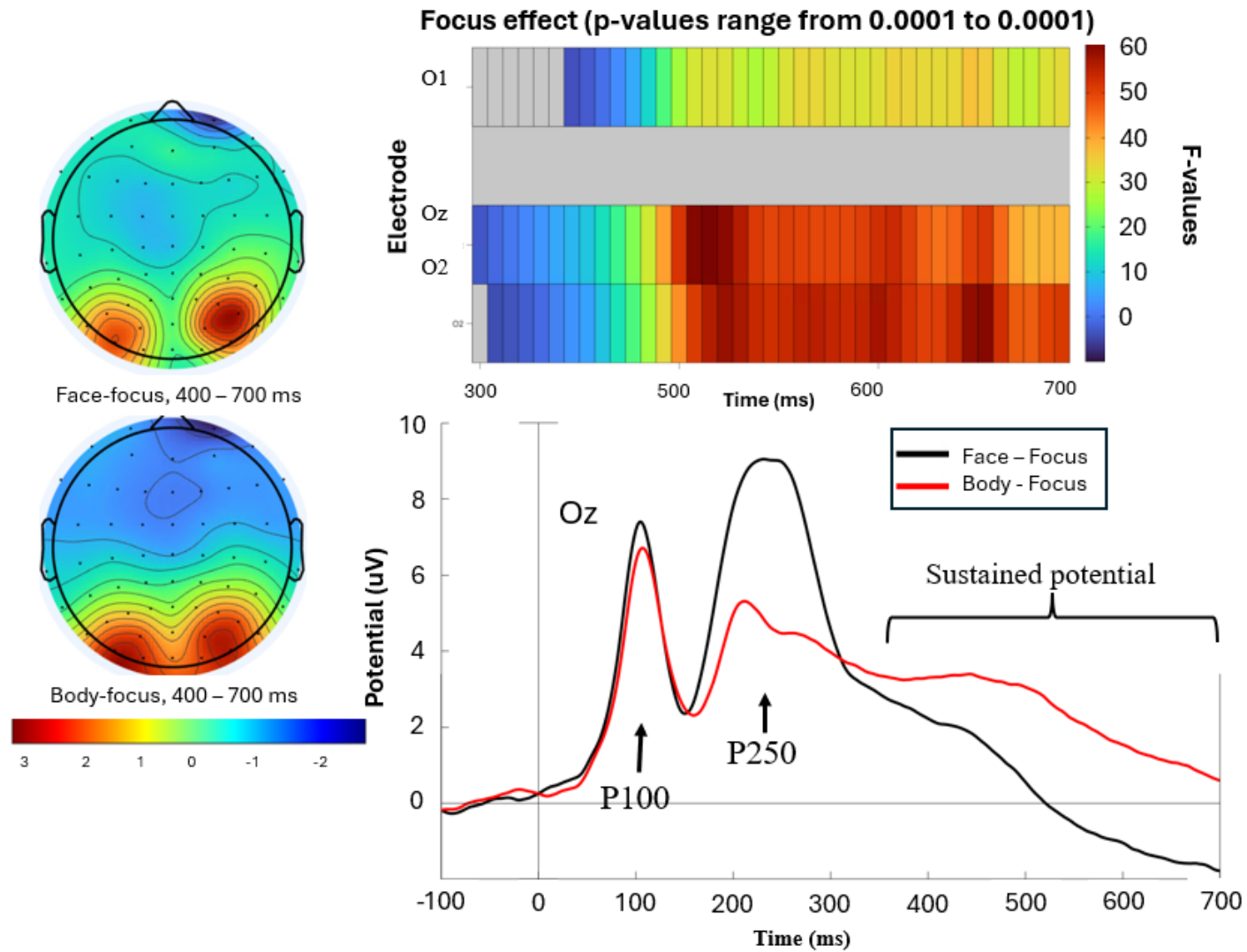
Note. (A) scalp and FMUT map for face-focus and body-focus condition between 70-140ms. (B) Grand averaged ERP for FC1 and PO7 electrodes for face-focus and body-focus conditions.

Figure 6. Scalp map, FMUT map and Grand Averaged ERP on the Fronto-Central, and Parieto-Occipital Sites



Note. (A) scalp and FMUT map for face-focus and body-focus condition between 150-318ms. (B) Grand averaged ERP for FC1 and PO7 electrodes for face-focus and body-focus conditions.

Figure 7. Scalp map, FMUT map and Grand Averaged ERP on Occipital Sites



Note. Scalp and FMUT map for face-focus and body-focus conditions between 400-700ms. Grand averaged ERP for Oz electrode for face and body focus conditions.

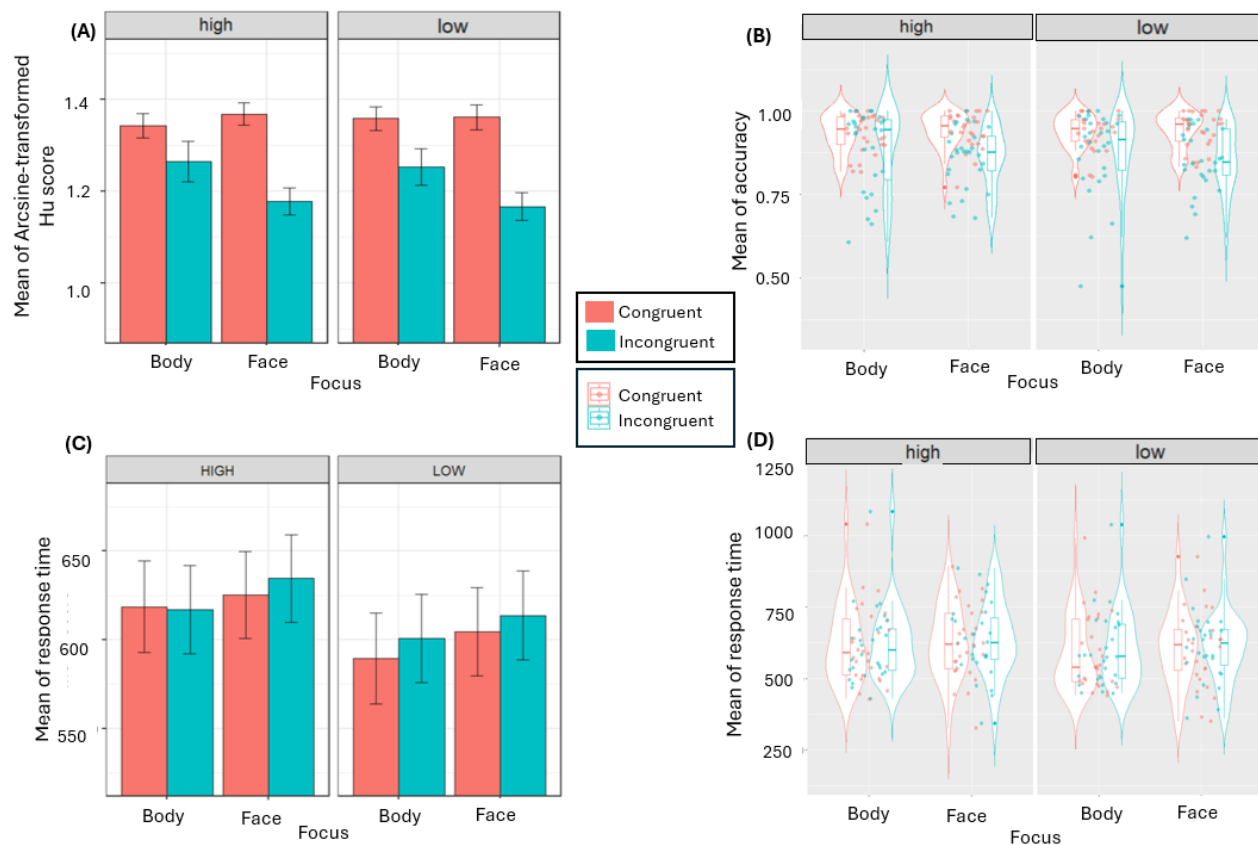
Hypothesis 3: Cognitive Load Effect

The repeated measures ANOVA on arcsine transformation to the Hu Score found no significant main effect of cognitive load ($p = 0.761$) or Congruence * Cognitive Load interaction ($p = 0.455$). No other comparisons, including those with cognitive load factor, were significant ($ps > 0.455$). The Bayesian repeated measures ANOVA indicated weak evidence against the inclusion of the Congruence * Cognitive Load interaction ($BF_{inclusion} = 0.311$) and moderate evidence against the inclusion of the cognitive load factor ($BF_{inclusion} = 0.192$).

The same repeated measures ANOVA on reaction times revealed a significant main effect of cognitive load ($F[1, 27] = 7.937, p = 0.009, \eta^2 = 0.025$). No other comparisons, including Cognitive Load, were significant ($ps > 0.095$; see Figure 8). Bonferroni-adjusted paired t -tests showed that response times were longer during high rather than low cognitive load ($t[27] = 2.817, p = 0.009, d = 0.164, CI [5.908, 37.582]$). The Bayesian analyses revealed moderate evidence for including the cognitive load factor ($BF_{inclusion} = 3.486$), and moderate evidence against including the Congruence * Cognitive Load interaction ($BF_{inclusion} = 0.116$).

Finally, analyses FMUT on ERPs did not show significant Congruence * Cognitive Load interaction ($ps > 0.065$). The Bayesian repeated measures ANOVA on ERPs revealed weak to moderate evidence against including the Congruence * Cognitive Load factor ($N100$: $BF_{inclusion} = 0.265$; $P100$: $BF_{inclusion} = 0.452$; $N250$: $BF_{inclusion} = 0.992$; $P250$: $BF_{inclusion} = 0.863$; *The sustained potential*: $BF_{inclusion} = 1.019$).

Figure 8. Histogram of Congruence in Function of Focus Type for Each Cognitive Load Conditions for Means of Arcsine Transformation (A) and for Response Times (C), and Violin Plot of Accuracy Means in Congruence in Function of Cognitive Load and Focus for Means of Arcsine Transformation (B) and Response Times (D)



Discussion

In our daily lives, we perceive emotions through both facial and bodily expressions, and these emotional cues interact to enhance the recognition of congruent emotions while making the recognition of incongruent ones more challenging (Lecker et al., 2020). However, it remains unclear whether these interactions occur automatically or voluntarily. Using high-quality stimuli, appropriate statistical methods (e.g., Bayesian statistics), and brain activity data with EEG analyses, we observed a contextual effect between facial and bodily expressions, with both facial and bodily expressions being recognized more accurately and quickly in congruent conditions compared to incongruent ones. Furthermore, bodily expressions had a greater influence on the

recognition of facial expressions than the reverse. At the cerebral level, we observed two stages of emotional processing, depending on attentional focus. In the early stages (N100, P100, P250, N250), greater brain activity was associated with attention directed at facial expressions rather than bodily expressions. In contrast, later stages (after 400ms) indicated that sustained attention was more strongly linked to focus on bodily expressions than facial expressions. Finally, both behavioral and neural findings suggest that there was no interaction between congruence and cognitive load, implying that the contextual effect was processed automatically. We discuss these effects in the following sections.

The Contextual Effect Between Face and Bodily Cues

Consistent with previous behavioral research, our findings suggest that recognition of facial and bodily expressions is enhanced in congruent conditions and becomes biased in incongruent ones (Lecker et al., 2020; Mondloch et al., 2013; Willis et al., 2011). Specifically, we observed improved recognition accuracy in congruent conditions compared to incongruent ones. This supports the notion that emotional recognition is facilitated by the integration of co-occurring emotional cues from both facial and bodily expressions (Kret et al., 2013; Meeren et al., 2005). This contextual effect supports the hypothesis of a common and integrated processing of facial and bodily emotional signals (de Gelder, 2006; de Gelder et al., 2015).

However, in contrast to earlier studies (Gu et al., 2013; Li, 2021), the behavioral findings and hypothesis of the present study, we did not observe significant differences in neural activity between congruent and incongruent conditions at early (P100, N100, P200, N200) and later (P300) stages for congruent compared to incongruent conditions during the face focus condition. Two non-exclusive explanations may account for this discrepancy. First, the high accuracy rate (90% correct) in the emotion recognition task suggests that participants found the task relatively easy, likely due to the extended stimuli presentation time (1500ms). While shorter presentation

times can introduce artifacts or increase task difficulty (Lecker et al., 2020), longer durations often result in ceiling effects, minimizing observable differences in brain activity (Kret & de Gelder, 2013). Thus, while the accuracy data confirm the impact of facial and bodily cues on recognition, the minimal ambiguity of the stimuli and prolonged exposure time may have diminished neural differences between congruent and incongruent conditions.

Second, the inclusion of a concurrent memory task may have influenced the neural processing of emotional context. Although we varied cognitive load (low vs. high), our study lacked a baseline condition without a memory task. This omission prevents direct comparison of contextual effects with and without cognitive load. Performing a dual task, regardless of cognitive load level, may have uniformly influenced brain activity associated with emotional processing. Supporting this idea, a study by Cao et al. (2022) on the impact of cognitive load on processing congruent and incongruent emotional background scenes and facial expression found a significant contextual effect on N170 amplitude under no-load conditions, but this effect disappeared under cognitive load conditions (Cao et al., 2022). These results suggest that the impact of emotional background scenes on facial expression processing primarily occurs when cognitive resources are not taxed by an additional task. This indicates that cognitive load can modulate emotional integration (Cao et al., 2022). In the present study, it is possible that the contextual effects at the neural level were diminished not by differences in cognitive load but by the overall influence of performing a secondary task. Future research should include a no-task baseline condition to clarify how dual-task demands impact the integration of facial and bodily emotional cues at the neural level.

The Role of Attentional Focus on the Facial and Bodily Expressions Integration

The contextual effect between facial and bodily expressions is modulated by attentional focus and operates bidirectionally. Bodily expressions influence the recognition of facial

expressions and vice versa, suggesting that bodily expressions are not merely contextual cues for interpreting facial expressions but can also serve as primarily targets for emotional recognition (Kret et al., 2013; Lecker et al., 2020). Notably, our findings indicate that facial emotion recognition is more strongly influenced by bodily expressions than the reverse (Lecker et al., 2020). This asymmetry may stem from a natural tendency to prioritize facial cues over body cues in everyday social interactions, with bodily expressions functioning as complementary rather than primary channels for emotion recognition (Aviezer et al., 2012; Lecker et al., 2020). Another possible explanation lies in the visual salience and consistency of bodily expressions, which often feature distinctive, easily recognizable cues, such as clenched fists in anger. The high reliability of these visual markers could reduce the reliance on facial expressions for emotional interpretation (Gu et al., 2013). Finally, another possible explanation lies in the intensity and arousal differences between facial and bodily expressions in the incongruent condition. Bodily expressions were perceived as more intense and more arousing than facial expressions. Given that higher emotional intensity and arousal are associated with improved recognition (Morningstar et al., 2021), the greater intensity of bodily expressions may have increased their influence on the recognition of less intense facial expressions. To our knowledge, the role of emotional intensity in contextual effects between facial and bodily cues remains unexplored, underscoring the need for further research.

At the neural level, the temporal processing of facial and bodily expressions follows a similar trajectory during both early and late stages, as evidenced by the consistent engagement of neural components such as P100, N100, P250, and N250. This activity in early neural encoding suggests substantial shared mechanisms for processing facial and bodily emotional cues, consistent with studies highlighting the involvement of common regions, such as the fusiform cortex (de Gelder, 2006; Stekelenburg & de Gelder, 2004).

However, during the early stages (100 to 250ms), attentional focus on facial expressions elicits greater brain activity, as evidenced by higher amplitudes in P100, N100, P250, and N250 compared to when attention is directed toward bodily expressions. The heightened activation around 100ms (P100 et N100) for facial-focused attention aligns with prior findings (Meeren et al., 2005; Stekelenburg & de Gelder, 2004) and may reflect the sensitivity of these components to the physical properties of facial stimuli, such as color, contrast (Puce et al., 2013; Yang et al., 2020), and low-frequency spatial cues, such as whether the mouth is open or closed (Pourtois et al., 2005). These differences in P100 and N100 amplitudes likely reflect the sensitivity of these components to facial features rather than being specific to emotional processing (Stekelenburg & de Gelder, 2004). Finally, our findings diverge from studies reporting greater early activation for bodily expressions compared to facial expressions (Li, 2021; Zhang et al., 2015). This discrepancy may arise from methodological differences, particularly in the types of emotional stimuli used. While previous studies predominantly examined negative emotions (Li, 2021; Zhang et al., 2015), our study included both positive and negative emotions. This broader emotional range aligns with findings of Meeren et al. (2005), who observed enhanced early brain activity for facial expressions across emotional valences, rather than for bodily expressions. The differences observed between this present study and previous ones could also be attributed to the stimuli used. Unlike earlier research, we employed high-quality, color stimuli designed to closely mimic real-life emotional perception. Additionally, we controlled the contrast and brightness of the stimuli to ensure these factors did not influence the EEG signals. Future research should further investigate how different emotional valences affect the neural mechanisms underlying the processing of facial and bodily expressions, while also utilizing higher-quality stimuli to obtain data that more accurately reflects real-world emotional processing.

In the later stages of processing (from 400ms), attentional focus on bodily expressions is associated with sustained neural activity, potentially reflecting the increased cognitive demands required to process bodily cues, which are often perceived as more ambiguous than facial expressions (Stekelenburg & de Gelder, 2004). This finding supports the notion that focusing on bodily expressions requires greater attentional engagement and cognitive resources, as individuals are less accustomed to relying exclusively on body cues for emotional recognition. In contrast, focusing on facial expressions leads to more efficient and automatic processing, occurring earlier in the timeline of emotional processing (Stekelenburg & de Gelder, 2004).

Finally, the greater influence of bodily expressions on facial emotion recognition compared to the reverse highlights the potential modulation of the contextual effect by attentional focus. Our results suggest that attentional focus may influence the processing of emotional cues from facial and bodily expressions, as previous behavioral (Lecker et al., 2020) and neural studies (Chen et al., 2022; Gu et al., 2013). Meta-analyses also support this idea, showing that attention and cognitive load modulate both early and late stages of facial emotion processing, potentially influencing the automaticity of these processes (Schindler & Bublatzky, 2020). However, no previous study has yet examined the influence of cognitive load on the integration of emotional signals when the attentional focus is on bodily expressions rather than facial expressions. Therefore, future research should investigate these contextual effects in greater depth to clarify the neural mechanisms underlying the bidirectional integration of facial and bodily expressions and how attentional focus and cognitive load interact with this process.

The Role of Cognitive Load on the Facial and Bodily Expressions Integration

Our behavioral results suggest that the contextual effect occurs automatically, satisfying the efficiency criterion for automaticity (Moors & De Houwer, 2006). Specifically, we found no interaction between congruence and cognitive load for either the accuracy or response time). This

indicates that the integration of facial and bodily expressions is similarly efficient under both high and low cognitive load, operating without significant cognitive effort, even during a concurrent task (Aviezer et al., 2011). These findings support the hypothesis of automatic contextual integration, consistent with prior behavioral studies (Aviezer et al., 2011).

At neural level, although no interaction between congruence and cognitive load was observed at either early or later stages, suggesting that the contextual effect may operate automatically, satisfying the efficiency criterion for automaticity (Moors & De Houwer, 2006). However, this interpretation should be considered with caution due to the absence of a clear contextual effect in the EEG data. Indeed, while the level of cognitive load (low vs. high) does not seem to affect the integration of facial and bodily expressions, it remains possible that the mere presence of a concurrent task, regardless of its cognitive demands, may influence this integration. Moreover, the lack of a direct correspondence between the behavioral data and the EEG components studies, although not entirely uncommon, is somewhat surprising. Typically, ERP effects are detected even when behavioral measures fail to capture effects, due to the higher sensitivity of neural recordings. In the present study, this dissociation may partly stem from the use of conservative analytical techniques, which could have masked subtle neural effects. Again, a no-task baseline condition could, in the future, help to clarify how dual-task demands impact the integration of facial and bodily emotional cues at the neural level.

Limitations

A limitation concerns the difference in perceived intensity among the emotions (anger, happiness, and sadness) and expression (facial vs. bodily expressions), as well as the differences in intensity and arousal between facial and bodily expressions in the incongruent condition. Specifically, happiness was rated as more intense than anger, and anger more intense than sadness. Bodily expressions were rated as more intense than facial expressions. Moreover, in

incongruent conditions, bodily expressions were rated as more intense and arousing than facial expressions. Given that higher emotional intensity and arousal can facilitate emotion recognition (Morningstar et al., 2021), such differences may have influenced the emotion recognition task.

Conclusion

Emotions are an integral part of daily life, making it essential to understand how we perceive and integrate them to respond effectively. Our study aimed to explore the integration of facial and bodily expressions using behavioral and neural measures while manipulating cognitive load as a criterion for automaticity. Our findings strongly suggest an automatic integration of facial and bodily emotional signals at the behavioral level, although this effect appears more nuanced at the neural level. We also observed a bidirectional asymmetry in the influence of bodily expressions on facial expression recognition and vice versa, suggesting the role of attentional focus in modulating the interaction between facial and bodily cues. Furthermore, temporal differences were noted in emotional processing depending on attentional focus. Facial expressions were processed more quickly and efficiently, whereas bodily expressions required extended attentional engagement, likely due to their greater ambiguity and dependence on contextual interpretation. Our research contributes to enhancing the understanding of emotional signal integration by being the first study to use both behavioral and neural measures, alongside high-quality stimuli that closely resemble real-life perception, to characterize the automaticity of this integration. Future research should further investigate the neural mechanisms underlying these contextual effects, including a no-load control condition which would help determine whether cognitive demands directly influence the integration process or if the contextual effect functions independently of such demands. Additionally, exploring how attention and cognitive load interact with the processing of congruent and incongruent expressions could deepen our understanding of the automaticity of this integration and its role in emotional recognition.

Author contributions

A-S.P: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing-original draft, Writing-reviewing, and editing.

S.R: Conceptualization, Funding acquisition, Resources, Supervision, Writing-reviewing and editing, Investigation, Methodology, Project administration.

Funding

This work was supported by “Natural Sciences and Engineering Research Council of Canada” (NSERC) Discovery Grant (Grant numbers: RGPIN-2020-06706).

Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

Data availability statement

Supplementary material available at: <https://osf.io/5at6g/>

References

- Abo Foul, Y., Eitan, R., & Aviezer, H. (2018). Perceiving emotionally incongruent cues from faces and bodies: Older adults get the whole picture. *Psychology and Aging, 33*(4), 660-666. <https://doi.org/10.1037/pag0000255>
- Aviezer, H., Bentin, S., Dudarev, V., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion, 11*(6), 1406-1414. <https://doi.org/10.1037/a0023578>
- 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. <https://doi.org/10.1111/j.1467-9280.2008.02148.x>
- Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225-1229. <https://doi.org/10.1126/science.1224313>
- Bago, B., Frey, D., Vidal, J., Houde, O., Borst, G., & De Neys, W. (2018). Fast and slow thinking: Electrophysiological evidence for early conflict sensitivity. *Neuropsychologia, 117*, 483-490. <https://doi.org/10.1016/j.neuropsychologia.2018.07.017>
- Belin, P., Campanella, S., & Ethofer, T. E. (2012). *Integrating face and voice in person perception*. Springer Science & Business Media.
- Cao, F., Zeng, K., Li, W., Liu, S., Zhang, L., Katembu, S., & Xu, Q. (2022). Influence of scene-based expectation on facial expression perception: The moderating effect of cognitive load. *Biological Psychology, 168*, Article 108247. <https://doi.org/10.1016/j.biopsycho.2021.108247>

- Chen, T., Sun, Y., Feng, C., & Feng, W. (2022). In identifying the source of the incongruent effect. *Journal of Psychophysiology*, 36(3), 167-176. <https://doi.org/10.1027/0269-8803/a000290>
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249. <https://doi.org/10.1038/nrn1872>
- de Gelder, B., de Borst, A. W., & Watson, R. (2015). The perception of emotion in body expressions. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(2), 149-158. <https://doi.org/10.1002/wcs.1335>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9-21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Durston, A. J., & Itier, R. J. (2021). The early processing of fearful and happy facial expressions is independent of task demands – Support from mass univariate analyses. *Brain research*, 1765, Article 147505. <https://doi.org/10.1016/j.brainres.2021.147505>
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99(3), 550-553. <https://doi.org/10.1037/0033-295X.99.3.550>
- Fields, E. C., & Kuperberg, G. R. (2020). Having your cake and eating it too: Flexibility and power with mass univariate statistics for ERP data. *Psychophysiology*, 57(2), Article e13468. <https://doi.org/10.1111/psyp.13468>
- Focker, J., Gondan, M., & Roder, B. (2011). Preattentive processing of audio-visual emotional signals. *Acta Psychologica*, 137(1), 36-47. <https://doi.org/10.1016/j.actpsy.2011.02.004>
- Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: A validation study. *Cognition & Emotion*, 22(6), 1094-1118. <https://doi.org/10.1080/02699930701626582>

- Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711-1725. <https://doi.org/10.1111/j.1469-8986.2011.01273.x>
- Gu, Y., Mai, X., & Luo, Y. J. (2013). Do bodily expressions compete with facial expressions? Time course of integration of emotional signals from the face and the body. *PLoS One*, *8*(7), Article e66762. <https://doi.org/10.1371/journal.pone.0066762>
- Hudson, A., Durston, A. J., McCrackin, S. D., & Itier, R. J. (2021). Emotion, gender and gaze discrimination tasks do not differentially impact the neural processing of angry or happy facial expressions: A mass univariate ERP analysis. *Brain Topography*, *34*(6), 813-833. <https://doi.org/10.1007/s10548-021-00873-x>
- Jaspers-Fayer, F., Maffei, A., Goertzen, J., Kleffner, K., Coccaro, A., Sessa, P., & Liotti, M. (2022). Spatiotemporal dynamics of covert vs. overt emotional face processing in dysphoria. *Frontiers in Behavioral Neuroscience*, *16*, Article 920989. <https://doi.org/10.3389/fnbeh.2022.920989>
- Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage*, *58*(2), 665-674. <https://doi.org/10.1016/j.neuroimage.2011.06.035>
- Karaaslan, A., Durmuş, B., & Amado, S. (2020). Does body context affect facial emotion perception and eliminate emotional ambiguity without visual awareness? *Visual Cognition*, *28*(10), 605-620. <https://doi.org/10.1080/13506285.2020.1846649>
- Kret, M. E., & de Gelder, B. (2010). Social context influences recognition of bodily expressions. *Experimental Brain Research*, *203*(1), 169-180. <https://doi.org/10.1007/s00221-010-2220-8>

- Kret, M. E., & de Gelder, B. (2013). When a smile becomes a fist: The perception of facial and bodily expressions of emotion in violent offenders. *Experimental Brain Research*, 228(4), 399-410. <https://doi.org/10.1007/s00221-013-3557-6>
- Kret, M. E., Stekelenburg, J. J., Roelofs, K., & de Gelder, B. (2013). Perception of face and body expressions using electromyography, pupillometry and gaze measures. *Frontiers in Psychology*, 4, Article 28. <https://doi.org/10.3389/fpsyg.2013.00028>
- Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, 4(1). <https://doi.org/10.1177/2515245920951503>
- Lecker, M., Dotsch, R., Bijlstra, G., & Aviezer, H. (2020). Bidirectional contextual influence between faces and bodies in emotion perception. *Emotion*, 20(7), 1154-1164. <https://doi.org/10.1037/emo0000619>
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Leys, C., Delacre, M., Mora, Y. L., Lakens, D., & Ley, C. (2019). How to classify, detect, and manage univariate and multivariate outliers, with emphasis on pre-registration. *International Review of Social Psychology*, 32(1). <https://doi.org/10.5334/irsp.289>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764-766. <https://doi.org/10.1016/j.jesp.2013.03.013>
- Li, X. (2021). Recognition characteristics of facial and bodily expressions: Evidence from ERPs. *Frontiers in Psychology*, 12, Article 680959. <https://doi.org/10.3389/fpsyg.2021.680959>

- Lima, C. F., Anikin, A., Monteiro, A. C., Scott, S. K., & Castro, S. L. (2019). Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion, 19*(2), 219-233.
<https://doi.org/10.1037/emo0000429>
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*, Article 213.
<https://doi.org/10.3389/fnhum.2014.00213>
- Luo, W., Feng, W., He, W., Wang, N. Y., & Luo, Y. J. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *Neuroimage, 49*(2), 1857-1867. <https://doi.org/10.1016/j.neuroimage.2009.09.018>
- McCrackin, S. D., & Itier, R. J. (2021). Feeling through another's eyes: Perceived gaze direction impacts ERP and behavioural measures of positive and negative affective empathy. *Neuroimage, 226*, Article 117605. <https://doi.org/10.1016/j.neuroimage.2020.117605>
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences, 102*(45), 16518-16523. <https://doi.org/10.1073/pnas.0507650102>
- Mondloch, C. J., Nelson, N. L., & Horner, M. (2013). Asymmetries of influence: Differential effects of body postures on perceptions of emotional facial expressions. *PLoS One, 8*(9), Article e73605. <https://doi.org/10.1371/journal.pone.0073605>
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297-326. <https://doi.org/10.1037/0033-2909.132.2.297>
- Morningstar, M., Gilbert, A. C., Burdo, J., Leis, M., & Dirks, M. A. (2021). Recognition of vocal socioemotional expressions at varying levels of emotional intensity. *Emotion, 21*(7), 1570-1575. <https://doi.org/10.1037/emo0001024>

- Palmer, J. A., Kreutz-Delgado, K., & Makeig, S. (2012). AMICA: An adaptive mixture of independent component analyzers with shared components.
<https://www.bing.com/ck/a?!&&p=8840a577b0cf5960efad0914b7629377c6b1446468e5ef5128083a778e73b064JmltdHM9MTc1NzI4OTYwMA&ptn=3&ver=2&hsh=4&fclid=01b0ca0b-33c3-629b-3260-de7a32f863a5&u=a1aHR0cHM6Ly9zY2NuLnVjc2QuZWR1L35qYXNvbi9hbWljYV9hLnBkZg&ntb=1>
- Paulmann, S., Jessen, S., & Kotz, S. A. (2009). Investigating the multimodal nature of human communication. *Journal of Psychophysiology*, 23(2), 63-76. <https://doi.org/10.1027/0269-8803.23.2.63>
- Peschard, V., Maurage, P., & Philippot, P. (2014). Towards a cross-modal perspective of emotional perception in social anxiety: Review and future directions. *Frontiers in Human Neuroscience*, 8, Article 322. <https://doi.org/10.3389/fnhum.2014.00322>
- Polich, J. (1990). P300, probability, and interstimulus interval. *Psychophysiology*, 27(4), 396-403. <https://doi.org/10.1111/j.1469-8986.1990.tb02333.x>
- Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: Time course and topographic evoked-potentials mapping. *Human Brain Mapping*, 26(1), 65-79. <https://doi.org/10.1002/hbm.20130>
- Puce, A., McNeely, M. E., Berrebi, M. E., Thompson, J. C., Hardee, J., & Brefczynski-Lewis, J. (2013). Multiple faces elicit augmented neural activity. *Frontiers in Human Neuroscience*, 7, Article 282. <https://doi.org/10.3389/fnhum.2013.00282>
- Puffet, A. S., & Rigoulot, S. (2025). Validation of the Emotionally Congruent and Incongruent Face-Body Static Set (ECIFBSS). *Behavior Research Methods*, 57(1), Article 41. <https://doi.org/10.3758/s13428-024-02550-w>

- Reschke, P. J., & Walle, E. A. (2021). The unique and interactive effects of faces, postures, and scenes on emotion categorization. *Affective Science*, 2(4), 468-483.
<https://doi.org/10.1007/s42761-021-00061-x>
- Righart, R., & de Gelder, B. (2006). Context influences early perceptual analysis of faces: An electrophysiological study. *Cerebral Cortex*, 16(9), 1249-1257.
<https://doi.org/10.1093/cercor/bhj066>
- Russell, A., Buerkner, P., Herve, M., Love, J., & Singmann, H. (2021). *Package 'emmeans'. R package version.*
- Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex*, 130, 362-386.
<https://doi.org/10.1016/j.cortex.2020.06.010>
- Seubert, J., Kellermann, T., Loughhead, J., Boers, F., Brensinger, C., Schneider, F., & Habel, U. (2010). Processing of disgusted faces is facilitated by odor primes: A functional MRI study. *Neuroimage*, 53(2), 746-756. <https://doi.org/10.1016/j.neuroimage.2010.07.012>
- Smith, N. K., Cacioppo, J. T., Larsen, J. T., & Chartrand, T. L. (2003). May I have your attention, please: Electro cortical responses to positive and negative stimuli. *Neuropsychologia*, 41(2), 171-183. [https://doi.org/10.1016/S0028-3932\(02\)00147-1](https://doi.org/10.1016/S0028-3932(02)00147-1)
- Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: An ERP study on the body-inversion effect. *Neuroreport*, 15(5), 777-780.
<https://doi.org/10.1097/01.wnr.0000119730.93564.e8>
- Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion*, 8(1), 81-95. <https://doi.org/10.1037/1528-3542.8.1.81>

- van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E.-J., Derks, K., Dablander, F., Gronau, Q. F., Kucharský, Š., Gupta, A. R. K. N., Sarafoglou, A., Voelkel, J. G., Stefan, A., Ly, A., Hinne, M., Matzke, D., & Wagenmakers, E.-J. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *L'Année psychologique*, *120*(1), 73-96. <https://doi.org/10.3917/anpsy1.201.0073>
- van den Bergh, D., Wagenmakers, E.-J., & Aust, F. (2023). Bayesian repeated-measures analysis of variance: An updated methodology implemented in JASP. *Advances in Methods and Practices in Psychological Science*, *6*(2). <https://doi.org/10.1177/25152459231168024>
- van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*(3), 487-494. <https://doi.org/10.1037/1528-3542.7.3.487>
- van den Stock, J., vandenbulcke, M., Sinke, C. B., Goebel, R., & de Gelder, B. (2014). How affective information from faces and scenes interacts in the brain. *Social Cognitive and Affective Neuroscience*, *9*(10), 1481-1488. <https://doi.org/10.1093/scan/nst138>
- van Dillen, L. F., & Derks, B. (2012). Working memory load reduces facilitated processing of threatening faces: An ERP study. *Emotion*, *12*(6), 1340-1349. <https://doi.org/10.1037/a0028624>
- van Doorn, J., van den Bergh, D., Bohm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharsky, S., Ly, A., Marsman, M., Matzke, D., Gupta, A., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E. J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*(3), 813-826. <https://doi.org/10.3758/s13423-020-01798-5>

- Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, *17*, 3-28. <https://doi.org/10.1007/BF00987006>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. <https://ggplot2.tidyverse.org>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), Article 1686. <https://doi.org/10.21105/joss.01686>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). dplyr: A grammar of data manipulation. <https://github.com/tidyverse/dplyr>
- Willis, M. L., Palermo, R., & Burke, D. (2011). Judging approachability on the face of it: The influence of face and body expressions on the perception of approachability. *Emotion*, *11*(3), 514-523. <https://doi.org/10.1037/a0022571>
- Yang, Y. F., Brunet-Gouet, E., Burca, M., Kalunga, E. K., & Amorim, M. A. (2020). Brain processes while struggling with evidence accumulation during facial emotion recognition: An ERP study. *Frontiers in Human Neuroscience*, *14*, Article 340. <https://doi.org/10.3389/fnhum.2020.00340>
- Zhang, D., Zhao, T., Liu, Y., & Chen, Y. (2015). Comparison of facial expressions and body expressions: An event-related potential study. *Acta Psychologica Sinica*, *47*(8), 963-970. <https://doi.org/10.3724/sp.J.1041.2015.00963>
- Zhang, M., Liu, T., Jin, Y., He, W., Huang, Y., & Luo, W. (2019). The asynchronous influence of facial expressions on bodily expressions. *Acta Psychologica*, *200*, Article 102941. <https://doi.org/10.1016/j.actpsy.2019.102941>

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67(6), 361-370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>

Transition vers le Chapitre 4

Le second article explore l'automaticité de l'effet contextuel à travers la manipulation du critère d'efficacité. Les résultats suggèrent que le niveau de charge cognitive n'influence pas l'effet contextuel, ni sur le plan comportemental ni sur le plan neural. En revanche, des différences apparaissent selon le focus attentionnel. L'expression corporelle semble exercer une influence plus marquée sur la reconnaissance de l'expression faciale que l'inverse. De plus, les données d'activité cérébrale indiquent qu'une plus grande allocation de ressources est observée pour la reconnaissance des expressions faciales à un stade précoce du traitement émotionnel, tandis que l'inverse se produit à un stade plus tardif. Ces résultats suggèrent que l'influence bidirectionnelle entre les expressions faciales et corporelles est modulée par l'orientation de l'attention sur l'un ou l'autre de ces indices émotionnels.

La prochaine étude examinera l'automaticité de l'effet contextuel entre expressions faciales et corporelles en manipulant la rapidité de traitement, celle-ci étant un indice de l'automaticité, les processus automatiques se caractérisant par une survenue rapide et précoce. Pour cela, trente participants ont effectué une tâche de reconnaissance d'émotions faciales ou corporelles, en utilisant des stimuli combinant des expressions faciales et corporelles congruents ou incongruents, présentés dans des conditions avec ou sans contrainte de temps. Pendant la tâche, nous avons enregistré l'activité cérébrale des participants avec un EEG, ainsi que les taux de reconnaissance et leurs temps de réponse.

Chapitre 4

Article 3 – How Speed of Processing Influences the Integration of Facial and Bodily
Emotional Expressions: An Event-Related Potential Study

How Speed of Processing Influences the Integration of Facial and Bodily Emotional Expressions: An Event-Related Potential Study

Anne-Sophie Puffet^{1,2}, and Simon Rigoulot^{1,2}

¹Department of Psychology, University of Quebec at Trois-Rivières.

²Research group CogNAC (Cognition, Neurosciences, Affect and Behaviour).

Author Note

Correspondence concerning this article should be addressed to Anne-Sophie Puffet, anne-sophie.puffet@uqtr.ca, and Simon Rigoulot, simon.rigoulot@uqtr.ca

This work was supported by “Natural Sciences and Engineering Research Council of Canada” (NSERC) Discovery Grant (Grant numbers: RGPIN-2020-06706).

The authors report there are no competing interests to declare.

Author’s contribution:

Anne-Sophie Puffet: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing-original draft, Writing-reviewing, and editing. **Simon Rigoulot:** Conceptualization, Funding acquisition, Resources, Supervision, Writing-reviewing and editing, Investigation, Methodology, Project administration, Validation.

Abstract

Emotions play a crucial role in everyday life by ensuring effective social interactions and guiding appropriate behavioral responses. Emotions are perceived through multiple emotional cues, in particular facial and bodily expressions. Notably, their simultaneous perception impacts emotional recognition, reflecting integration between channels. However, whether this integration occurs through automatic or controlled processing remains unclear, a distinction crucial for understanding how emotional signals are processed during everyday social interactions. To address this, we manipulated speed of emotional processing as a criterion for automaticity, testing whether face-body integration meets conditions for automatic processing. Thirty participants performed an emotion recognition task using emotionally congruent and incongruent face-body stimuli under both time-constrained and unconstrained conditions. Brain activity (EEG), response accuracy and reaction times were measured. Bayesian ANOVAs revealed that congruent expressions improved emotional recognition. Bodily expressions also had a stronger influence on facial expressions than vice versa. At the neural level, congruent stimuli elicited greater activity at early (between 110ms and 310ms) and late stages of emotional processing (beyond 300ms post-stimulus), indicating that coherent emotional information is prioritized and processed more efficiently. Across all analyses, the integration of facial and bodily expressions remained consistent in time-constrained and unconstrained conditions, at behavioral and neural levels, suggesting that this process operates automatically. The study provides new insights and complements previous findings into the perceptual mechanisms underlying emotion recognition by suggesting that the integration of facial and bodily emotional cues operates automatically, and by highlighting the bidirectional yet asymmetrical influences between facial and bodily expressions.

Keywords: processing speed; contextual effect; facial emotion; bodily expression; automaticity

Introduction

Social interactions are fundamental to human existence, and their harmonious development depends significantly on our ability to accurately decode emotions of others. Traditionally, research has primarily focused on the processing of isolated facial expressions (de Gelder, 2009). While important, facial expressions represent only a partial channel of emotional communication. Indeed, emotions are conveyed through multiple visual cues, like facial emotion, bodily expressions (Gu et al., 2013) or contextual background scenes (Kret & de Gelder, 2010), as well as other sensory modalities, including voice and odors (Focker et al., 2011; Seubert et al., 2010). This predominant focus on isolated facial expressions has thus overlooked the complexity of real-world affective communication, limiting a comprehensive understanding of emotions perception. Moreover, it remains unclear whether the simultaneous perception of facial and bodily expressions relies on automatic or controlled processing. Understanding this distinction is crucial, as automatic processing enables individuals to respond quickly during social interactions and supports behaviors that promote effective adaptation to their environment. This study aims to investigate the automaticity of simultaneous perception of facial and bodily expressions in conditions that closely resemble real-life situations, by manipulating processing speed, a criterion to assess automatic processing.

Studies have shown that the simultaneous perception of facial and bodily cues is essential for accurate emotion recognition (Abo Foul et al., 2018; Lecker et al., 2020; Mondloch et al., 2013; Willis et al., 2011). Specially, recognizing facial emotions is facilitated when facial and bodily expressions are congruent, meaning they convey the same emotion. Conversely, recognition becomes more challenging when these cues are incongruent, with bodily expressions conveying a different emotion than the face (Abo Foul et al., 2018; Lecker et al., 2020; Mondloch et al., 2013; Willis et al., 2011). This integration, commonly referred to in the literature as the

contextual effect, supports the notion that emotional facial and bodily expressions are perceived holistically (Aviezer et al., 2012a; Puffet et al., 2025).

Although previous research has explored the contextual effect between facial and bodily expressions, it remains unclear whether this integration occurs automatically or requires voluntary, effortful processing. This question is essential because automatic emotional perception is linked to evolutionary benefits. For example, it allows individuals to respond rapidly during social interactions and supports behaviors that enhance adaptation to the environment. Indeed, detecting and reacting quickly to someone expressing fear is adaptive, as it enables the individual to engage in context-appropriate behaviors, such as fleeing or defending oneself, to ensure survival. Moors and De Houwer (2006) proposed a model that defines three criteria for identifying automatic process: fast, efficient and unconscious. Several studies, applying these criteria, suggest that the recognition of isolated emotional signals, such as facial or vocal expressions, can take place automatically (Lima et al., 2019; Tracy & Robins, 2008). For instance, Tracy and Robins (2008) showed that the accuracy of facial emotion recognition remained stable even when participants had limited time to respond or were simultaneously performing a cognitively demanding task. What remains to be determined is whether this automaticity applies only to single emotional cues or whether it also extends to the integration of multiple emotional signals, such as those conveyed simultaneously by the face and body.

To address this question, several studies investigated whether the contextual effect observed between facial and bodily expressions reflects automatic integration. One study showed that this effect between facial and bodily expressions can occur *without conscious awareness*, as it was observed even when bodily expressions were presented very briefly (33ms) (Karaaslan et al., 2020). In addition, this effect appears to be *efficient*, as it is not influenced by cognitive load (Aviezer et al., 2011; Puffet & Rigoulot, 2025a). In one experiment, participants were asked to

recognize facial or bodily expressions while simultaneously performing a memory task under either high cognitive load (e.g., memorizing seven different digits) or a low cognitive load (e.g., memorizing seven same digits). Regardless of the level of cognitive load, facial expression recognition was consistently influenced by bodily expressions in the same way (Puffet & Rigoulot, 2025a).

A final criterion manipulated to investigate the automaticity of facial and bodily cue integration is processing speed. A previous study explored the speed of emotional processing by testing facial emotion recognition in the presence of congruent and incongruent bodily expressions (Meeren et al., 2005). Stimuli were presented for 200ms, and participants were instructed to respond as quickly as possible by identifying whether the facial expressions showed anger or fear. The results indicated that the contextual effect persisted despite the brief exposure and speeded responses, providing supportive evidence for automatic processing (Meeren et al., 2005).

This study also examined the automaticity of facial-body integration using electroencephalography (EEG), a technique that allows precise characterization of the timing of brain activity following stimulus perception. Results showed that the integration of facial and bodily cues occurs rapidly at the neural level (Meeren et al., 2005). More broadly, ERP studies have identified three stages in the processing of emotional facial and bodily expressions: an early (~100ms, P100/N100), an intermediate (~200ms, P200/N200), and a late stage (~300ms, P300/LPP) (Gu et al., 2013). Early neural responses suggest automatic processing and a negativity bias, with larger P100 amplitudes for threatening bodily expressions like fear, regardless of attention is directed toward facial or bodily emotions (Chen et al., 2022; Gu et al., 2013). At 115ms post-stimulus, increased brain activity has been observed in response to incongruent compared to congruent facial and bodily expressions, even when presented briefly.

This early conflict detection supports the hypothesis that facial-body integration operates automatically, as it aligns with the speed criterion (Meeren et al., 2005). At the intermediate stage, incongruence between facial and bodily cues is still processed, as shown by increased N200 (Chen et al., 2022; Gu et al., 2013) amplitudes for mismatched expressions, supporting rapid integration. At later stages, attention influences emotional processing. Specifically, when attention is directed toward the face, P300 amplitudes are greater for congruent than for incongruent expressions, suggesting sustained attention and more extensive elaboration of congruent information compared to incongruent ones (Gu et al., 2013; Li, 2021). These findings highlight an early prioritization of bodily information, rapid conflict detection, and provide evidence regarding the automatic nature of facial and bodily cue integration for the speed criterion (Meeren et al., 2005).

Although previous studies have shown that the integration of facial and bodily expressions occurs rapidly (Meeren et al., 2005), it remains unclear whether this rapid processing meets the criteria for automaticity. To the best of our knowledge, only one ERP study has directly examined the neural mechanisms underlying this integration by manipulating processing speed, which limits our understanding of its temporal dynamics (Meeren et al., 2005). One important consideration is that previously the speed criterion for facial and bodily expression integration has been assessed by briefly presenting congruent and incongruent stimuli (Meeren et al., 2005). However, brief stimulus presentation alone may not fully capture the rapidity or automaticity of processing, as emotional information can continue to be processed after the stimulus offset (Moors & De Houwer, 2006). One possibility to address this would be to prevent the processing of emotional expressions after stimulus presentation to ensure that processing does not persist beyond the exposure to the cues. To do this, a more robust approach involves combining short presentation times with speeded response demands, minimizing the influence of controlled

processing (Moors & De Houwer, 2006). Moreover, this previous study (Meeren et al., 2005) relied on frequentist analyses, which do not allow for showing the absence of an effect, as suggested in the study, thereby limiting the validity of the results. A more appropriate approach relies on Bayesian analyses, which provide evidence in favor of one hypothesis over another (Lee & Wagenmakers, 2013; van den Bergh et al., 2020, 2023; van Doorn et al., 2021). Finally, most research on facial and bodily expression integration has focused on how bodily expressions influence facial emotion recognition, overlooking the potential impact of facial expressions on the perception of bodily emotions. Given that attentional focus play a crucial role in emotional processing (Gu et al., 2013; Lecker et al., 2020; Zhang et al., 2019), it is essential to investigate its influence on the automaticity of facial and bodily expression integration.

Consequently, this study aims to investigate the automaticity of facial and bodily expression integration, focusing on the speed criterion (Moors & De Houwer, 2006). To achieve this, participants will perform an emotion recognition task briefly presented congruent and incongruent facial and bodily expressions. They will be instructed to respond either quickly or without time pressure, allowing us to manipulated processing speed. ERP recordings will be used to examine the temporal dynamics of integration, and attention will be directed toward either facial or bodily cues. As the hypothesis proposes automatic integration, independent of response speed, we will use Bayesian analyses to assess the strength of evidence for the absence of a speed effect (Lee & Wagenmakers, 2013; van den Bergh et al., 2020, 2023; van Doorn et al., 2021).

Three hypotheses will be tested using both behavioral and EEG measures:

- *Hypothesis 1*: we anticipate observing the contextual effect between facial and bodily cues, characterized by improved emotion recognition accuracy and faster responses in the congruent compared to incongruent conditions (Lecker et al., 2020). On the brain level, we expect that incongruent expressions will evoke higher amplitudes in early

ERP components (P100, N100, P200, N200), reflecting rapid conflict detection (Gu et al., 2013).

- *Hypothesis 2*: we predict that task instructions directing attention on face or bodily cues will influence the contextual effect (Lecker et al., 2020). Specifically, the influence of facial expressions on recognizing bodily emotions (when focusing on the body) should be weaker than the reverse effect of bodily cues on facial emotion recognition (when focusing on the face). Regarding ERP components, we expect that focusing attention on the face will enhance P300 amplitudes for congruent pairs rather than incongruent ones, indicating facilitated emotional processing and more efficient sustained attention allocation (Gu et al., 2013; Li, 2021).
- *Hypothesis 3*: we propose that processing speed will not affect these contextual effects (Meeren et al., 2005), suggesting the integration occurs automatically. This would be reflected by a strong level of confidence in the alternative hypothesis, i.e., the absence of differences between fast and slow conditions in both behavioral (accuracy of recognition and reaction times) and brain measures (amplitudes of ERPs of interest, P100, N100, P200, N200).

Methods

Participants

The sample size for this study was estimated using the *SuperPower* package (version 0.2.0; power = 0.80, standard deviation = 0.05) in R studio (version 4.3.0) (Lakens & Caldwell, 2021), resulting in a minimum required total of 25 participants. A total of 34 right-handed participants (20 females; $M_{\text{age}} = 24 \pm 3.66$ years; range 18-35 years) were recruited between January and April 2025 via social networks and a specific University platform for advertising experiments. Each participant received CAD 30 as compensation for their participation in a

three-hour session. Inclusion criteria required participants to be between 18 and 35 years old. Exclusion criteria included a history of neurological disorders (e.g., traumatic brain injury), diagnosed psychological conditions (e.g., psychosis), uncorrected visual or auditory impairments, and the use of medications or substances (e.g., drugs) affecting the nervous system. To assess mood, participants completed the French version of *The Hospital Anxiety and Depression Scale* (HADS; Means of Anxiety score = 5.97 +/- 2.38; Means of depression score = 3.18 +/- 2.34) (Zigmond & Snaith, 1983). Four participants were excluded from the analysis due to an issue with the data collection software. The final sample included 30 participants (17 females; $M_{\text{age}} = 24 \pm 3.50$ years; range 18-35 years) with mean HADS score of 6 (± 2.41) for anxiety and 3.43 (± 2.23) for depression.

Sociodemographic information, including date of birth, gender, dominant hand, native language, and education level, was also collected (see <https://osf.io/stf56/> for details). All participants provided their informed consent, and the study was approved by the UQTR ethics committee (certificate number: CERPPE-24-36-07.05). The protocol was not registered.

Materials

The stimuli were selected from the “Validation of the Emotionally Congruent and Incongruent Face-Body Static Set” (ECIFBSS) (Puffet & Rigoulot, 2025b), which consists of 1 952 images depicting facial and bodily expressions (49 images per actor) in both congruent and incongruent conditions. For this study, we selected 120 stimuli in total, including 60 congruent images (20 for each emotion: happiness, anger, and sadness) and 60 incongruent images (ten for each combination of differing emotional expressions; e.g., 10 happy faces with angry body and 10 happy faces with sad body). The mean recognition accuracy was 86% for facial expressions and 87% for bodily expressions.

The ECIFBSS provided the evaluation of intensity, valence and arousal (see Table 1; <https://osf.io/stf56/>). Multiple ANOVAs were performed to explore differences in these emotional dimensions (valence, arousal, and intensity) as a function of emotion (happiness, sadness, and anger), congruence (congruent vs. incongruent), and mode of expression (facial vs. bodily expressions).

In terms of valence, the ANOVA revealed a significant main effect of emotion ($F[2, 228] = 996.178, p < 0.001, \eta^2 = 0.884$), an interaction between emotion and mode of expression ($F[2, 228] = 6.617, p = 0.002, \eta^2 = 0.006$), an interaction between mode of expression and congruence ($F[1, 228] = 6.136, p = 0.014, \eta^2 = 0.003$), and a three-way interaction between emotion, mode of expression and congruence ($F[2, 228] = 4.594, p = 0.011, \eta^2 = 0.004$). Bonferroni-adjusted paired *t*-tests showed that happiness was perceived significantly as more positive than both anger ($t[228] = -39.943, p < 0.001, d = -6.316, CI [-2.998, -2.663]$), and sadness ($t[228] = 37.225, p < 0.001, d = 5.886, CI [2.471, 2.805]$). Simple effects analysis revealed a significant effect of mode of expression for sadness ($F[1, 228] = 8.777, p = 0.004$) and a significant effect of congruence during face condition ($F[1, 228] = 6.915, p = 0.010$). These findings suggest that for sadness, facial expressions were rated as more negative than bodily expressions. Additionally, during the face condition, congruent stimuli were rated as more positive than incongruent ones. No other significant effects were found ($ps > 0.05$).

Table 1. Mean and Standard Deviation of Valence, Arousal and Intensity for Facial and Bodily Emotions

	Valence	Arousal	Intensity
Facial emotion			
Happiness	6.73 (+/- 0.68)	5.52 (+/- 0.58)	5.90 (+/- 1.06)
Anger	3.55 (+/- 0.34)	5.12 (+/- 0.45)	4.95 (+/- 1.08)
Sadness	3.77 (+/- 0.24)	4.44 (+/- 0.32)	4.18 (+/- 0.63)
Happiness	6.53 (+/- 0.42)	5.58 (+/- 0.52)	6.45 (+/- 0.90)
Bodily emotion			
Anger	3.57 (+/- 0.31)	5.38 (+/- 0.37)	5.62 (+/- 0.72)
Sadness	3.93 (+/- 0.34)	4.38 (+/- 0.31)	4.14 (+/- 0.76)

In terms of arousal, the ANOVA revealed a significant main effect of emotion ($F[2, 228] = 98.607, p < 0.001, \eta^2 = 0.428$) and congruence ($F[1, 228] = 5.716, p = 0.018, \eta^2 = 0.012$), and an interaction between mode of expression and emotion ($F[2, 228] = 6.822, p = 0.001, \eta^2 = 0.03$). Bonferroni-adjusted paired t -tests showed that happiness was perceived as more arousing than anger ($t[228] = -5.542, p < 0.001, d = -0.876, CI [-0.594, -0.239]$) and anger was rated as more arousing than sadness ($t[228] = 8.404, p < 0.001, d = 1.329, CI [0.455, 0.810]$). Moreover, congruent conditions were rated as more arousing than incongruent conditions ($t[228] = 2.391, p = 0.018, d = 0.309, CI [0.026, 0.268]$). Simple effects analysis revealed a significant effect of mode of expression during anger emotion ($F[1, 228] = 11.634, p < 0.001$) suggesting that for anger, bodily expressions were rated as more arousing than facial expressions. No other significant effects were found ($ps > 0.05$).

In terms of intensity, the ANOVA revealed a significant main effect of emotion ($F[2, 228] = 73.680, p < 0.001, \eta^2 = 0.356$) and an interaction between mode of expression and emotion ($F[2, 228] = 10.226, p < 0.001, \eta^2 = 0.049$). Bonferroni-adjusted paired t -tests showed that happiness was perceived as more intense than anger ($t[228] = -6.788, p < 0.001, d = -1.073, CI [-1.405, -0.680]$) and anger was rated as more intense than sadness ($t[228] = 5.321, p < 0.001,$

$d = 0.841$, CI [0.455, 1.180]). Simple effects analysis revealed a significant effect of mode of expression during anger ($F[1, 228] = 18.424$, $p < 0.001$) and sadness ($F[1, 228] = 4.374$, $p < 0.001$) suggesting that for anger, bodily expressions were rated as more intense than facial expressions and for sadness, bodily expressions were rated as less intense than facial expressions. No other significant effects were found ($ps > 0.05$).

Since brightness and contrast affect the amplitude of ERP components (Puce et al., 2013), these visual properties were standardized across all stimuli (OSF; <https://osf.io/stf56/>). Brightness and contrast values were measured using ImageJ software (version 1.53). The mean brightness was 232.75 (+/- 2.84) for congruent conditions and 232.90 (+/- 2.70) for incongruent conditions, while the mean of contrast was 61.56 (+/- 4.69) and 61.54 (+/- 4.86) respectively. Two one-way ANOVAs were conducted to examine potential differences in brightness and contrast between congruent and incongruent conditions. Results could not reveal significant effect of congruency on either brightness or contrast ($ps > 0.767$).

Experimental Design

The experimental employed a within-subject design, manipulating four independent variables: attentional focus (on face emotion or bodily expression), processing speed (fast or slow), congruency (congruent or incongruent between facial and bodily expressions), and type of emotional expression (happiness, sadness, and anger). We measured three dependent variables: reaction time, accuracy rate, and brain activity through the amplitude of relevant ERP components. A total of 480 trials were conducted, evenly distributed across the two attentional focus conditions (240 trials each for facial and bodily expressions). Within each attentional focus condition, participants completed 120 trials for each congruency condition (congruent and incongruent), and 120 trials per speed condition (fast and slow). For congruent emotion pairing

(happiness, sadness, or anger), each emotion appeared in 40 trials. For incongruent emotion pairing (e.g., angry face with sad body), each specific combination appeared in 10 trials.

General Procedure

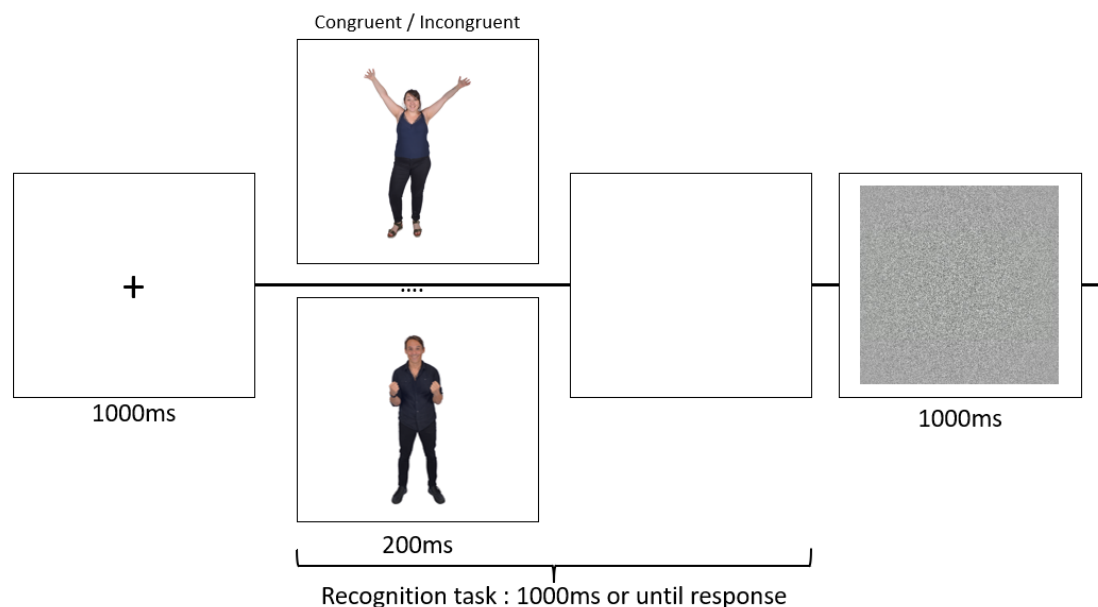
Before the experiment, participants completed three online questionnaires via LimeSurvey software: a sociodemographic questionnaire, the HADS (Zigmond & Snaith, 1983), and a consent form.

The experiment lasted approximately three hours, including one hour for the task and two hours for encephalography (EEG) equipment setup. It took place in a dedicated EEG laboratory at the University of Quebec at Trois-Rivières (UQTR). Participants were seated comfortably in a Faraday cage, 60cm from an HP EliteDisplay E232 computer screen (59cm diagonal, 29cm height, 51cm width). The computer used for stimulus presentation was an HP ProDesk 600 G1 TWR. The task was programmed using E-Prime 2.0 software.

The experiment consisted of two identical blocks presented in a counterbalanced order across participants. In the first block, participants were instructed to focus on recognizing facial expressions. In the second, they focused on recognizing bodily expressions. Each of these blocks included two speed conditions: a fast condition, in which participants were instructed to respond as quickly as possible, and a slow condition, in which they could take as much time as needed. The order of the speed conditions was also counterbalanced across participants. Each part began with a training session (five trials), following by an experimental session with 240 randomized trials. Each trial started with a fixation cross displayed for 1000ms, followed by an image presented for 200ms, depicting a person expressing an emotion through both face and body posture in either a congruent or incongruent condition. During image presentation, participants were instructed to recognize the emotion expressed by either the face or the body, depending on the block. They were also required to respond either as quickly as possible or at their own pace,

depending on the speed condition (fast or slow). Responses were recorded via a keyboard using three directional arrows mapped to anger, sadness, or happiness, counterbalanced across participants. After each response, a visual mask appeared for 1000ms before the next trial began (see Figure 1).

Figure 1. *Shema of Experimental Procedure*



Statistics Analysis

The behavioral and EEG data were analyzed using MATLAB (R2021a), JASP (version 0.18.1; RRID:SCR_015823), and R (version 4.3.0). The package ggplot2 (version 3.4.2; RRID:SCR_014601) (Wickham, 2016), tidyverse (version 1.2.0; RRID:SCR_019186) (Wickham et al., 2019), dplyr (version 1.1.2; RRID :SCR_017102) (Wickham et al., 2023), and emmeans (version 1.8.7; RRID :SCR_018734) (Russell et al., 2021) were used. The analyses conducted are available on the OSF platform (<https://osf.io/stf56/>).

Behavioral Analysis

We analyzed the emotion recognition task by assessing two behavioral metrics: the accuracy rate of correct categorization and response time. For accuracy, we calculated the

Unbiased hit rates (Hu score) (Wagner, 1993) to adjust for potential responses biases that could artificially inflate the accuracy score. For instance, a participant who categorizes all facial expressions as “anger” may obtain a high accuracy score for anger, but this would not reflect true recognition. The Hu score addresses this by considering the squared frequency of correct responses for a target emotion, divided by the product of the total number of stimuli representing that emotion and the overall frequency of selecting that emotion category (Wagner, 1993). Outliers in both Hu scores and response times were identified and excluded using the *Mean Absolute Deviation* (MAD) function in R (Leys et al., 2013, 2019).

We then conducted classical repeated measures ANOVAs on the arcsine-transformed Hu score, as suggested in the literature (Goeleven et al., 2008; Wagner, 1993) and on response times, incorporating three factors: congruency, attentional focus, and speed. Since traditional repeated measures ANOVAs do not directly support the null hypothesis (i.e., no difference between conditions), we complemented them with repeated measures Bayesian ANOVAs using the same factors. For Bayesian analyses, we applied default prior options for fixed effects ($r = 0.5$ for the fixed effects) as recommended (van den Bergh et al., 2020; van Doorn et al., 2021). The Bayesian ANOVAs generated two types of models: the null model and alternative models, which included one or more factors and their interactions. We calculated Bayes factors (BF10) for each model using JASP software to assess the strength of evidence in favor of each alternative model relative to the null model or the best-performing model (van den Bergh et al., 2023). Specifically, a BF10 ratio quantifies how much more likely the observed data are under the alternative model (i.e., the most likely model) compared to the null model. Evidence for the null hypothesis is indicated by BF10 values between 1 and 1/3 (weak evidence), between 1/3 and 1/10 (moderate evidence), and below 1/10 (strong evidence). Conversely, B10 values between 1 and 3 (weak evidence), 3 and 10 (moderate evidence), and above 10 (strong evidence) support the alternative hypothesis (Lee

& Wagenmakers, 2013; van Doorn et al., 2021). For example, a BF₁₀ of 30 suggests that the data are thirty times more likely to occur under the alternative model than the null model (van den Bergh et al., 2020). Additionally, we calculated *inclusion Bayes factor* (BF_{inclusion}) for each factor (congruency, focus, and speed) to evaluate the evidence for including these factors and their interactions within the model set. This approach enabled us to assess the predictive value of each factor in the dataset (van den Bergh et al., 2020).

ERP Recording, Preprocessing, and Analysis

EEG data were recorded using BrainVision software with 64 scalp electrodes, and Fpz served as the ground electrode. The data were sampled at 500Hz per channel for offline analysis. Vertical (VEOG) and horizontal (HEOG) electrooculograms were recorded using three additional electrodes: one placed below the left eye (Fp2), and two on the left (FT9) and right (FT10) sides of the eyes. Reference electrodes (TP9 and TP10) were placed on the earlobes. All electrode impedances were kept below 25 k Ω .

Data pre-processing was performed using EEGLab (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes in MATLAB, following the pipeline recommended by Lopez-Calderon and Luck (Lopez-Calderon & Luck, 2014). The data were re-referenced offline to the average, followed by filtering with a 0.1 Hz high-pass filter. Independent Component Analysis (ICA) was conducted by the first author (ASP) using AMICA (Palmer et al., 2012). Signals were segmented into epochs ([-200ms; 700ms]) around the target stimuli. A low-pass filter with a 30 Hz cutoff was applied, and two rounds of artifact detection were carried out. The first round used the “moving window peak-to-peak threshold” method from ERPLAB, with an 85uV threshold, a 200ms window size, and a 50ms window step. The second-round targeted eye movements at electrodes Fp2, FT9, and FT10, using the “step-like artifacts” detection method, with a 30uV threshold, a 200ms window size, and a 50ms window step. After

artifact rejection, an average of 97.56% of data was retained for further analysis (Face focus: 97.57% ($M = 7025$ trials); body focus: 97.54% ($M = 7023$ trials); Congruency: 97.47% ($M = 3509$ trials); Incongruency: 97.64% ($M = 3515$ trials); Fast speed: 97.63% ($M = 3514.5$ trials); Slow speed: 97.49% ($M = 3509.5$ trials)).

The EEG data were analyzed using the Factorial Mass Univariate Toolbox (FMUT) (Fields & Kuperberg, 2020). This approach involves conducting thousands of inferential statistical tests iteratively across all electrodes within a temporal window of interest, applying multiple comparison corrections. (Durston & Itier, 2021; Fields & Kuperberg, 2020; Hudson et al., 2021; Jaspers-Fayer et al., 2022; McCrackin & Itier, 2021). Mass univariate analyses are exploratory and do not require *a priori* assumptions, offering greater statistical power compared to traditional spatiotemporal averaging methods (Fields & Kuperberg, 2020; Groppe et al., 2011).

We first conducted an initial exploratory mass univariate analysis across all electrodes and time points from -100ms pre-stimulus to the end of the epoch (700ms). Based on this analysis and previous research (Gu et al., 2013; Meeren et al., 2005; van Dillen & Derks, 2012; Zhang et al., 2015), we identified one early time windows (110 – 310ms) of interest at fronto-central (AF4, AF7, AF8, F5, F6, F7, F8, FC1, FCz, C1, C2, C3, and Cz) and at parieto-central sites (Pz, P2, P7, CPz, CP1, CP2, and CP4). We also identified one later time window (300 – 700ms) in the fronto-central (AF7, AF8, AF4, F2, F5, F6, F7, F8, Fz, FC1, FC5, and FCz), centro-parietal (CPz, CP1, CP2, CP4, Pz, P2, P7, Cz, C1, C2, and C3) and occipital sites (O1, O2, and OZ).

The mass univariate analysis used an alpha level of 0.05 and incorporated three within-subjects factors: speed, focus, and congruence. In FMUT, ANOVAs were conducted using 100,000 permutations for each data point, with correction for multiple comparisons via permutation test, as recommended (Durston & Itier, 2021; Fields & Kuperberg, 2020; Hudson et al., 2021). In addition, repeated-measures Bayesian ANOVAs were performed on the amplitude

of components averaged over the same time windows and electrodes sites as the FMUT analysis, using the three factors: speed, focus, and congruence. To address the automaticity hypothesis, we expect no interaction between processing speed and emotional congruence (see hypothesis 3).

Results

Preliminary Analyses: Behavioral and EEG Data

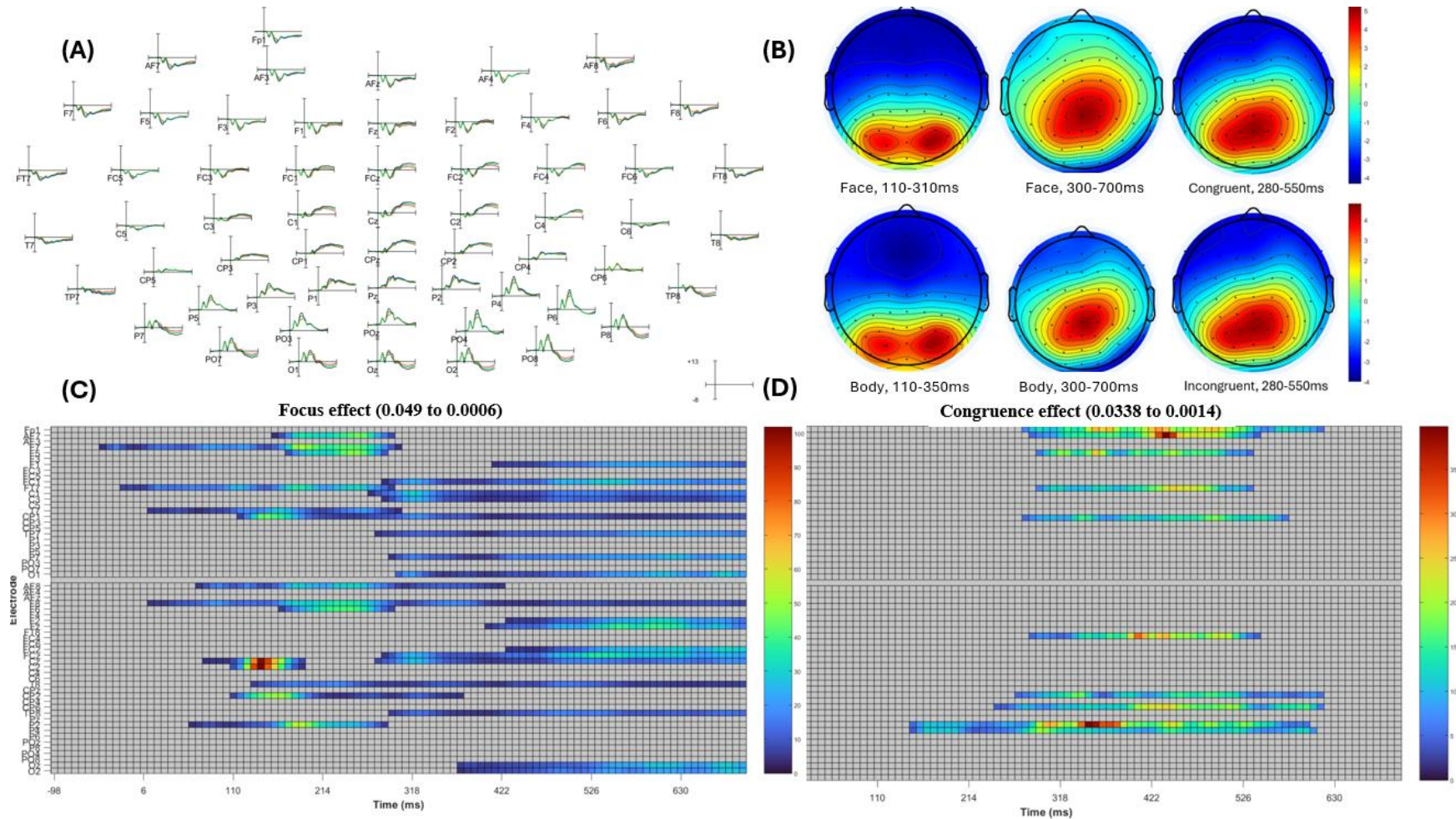
We used the MAD function (Leys et al., 2013, 2019) to identify outlier trials and participants. For Hu scores, no participants or trials were excluded due to Hu score outliers. For reaction times, no participants and 9.8% of the trials were excluded due to response time outliers.

For ERP data, in order to identify the spatiotemporal dynamics of the effects of focus, congruence and speed on ERP responses, we conducted an exploratory analysis including all electrodes and the entire -100 to 700ms post-stimulus time window.

This exploratory analysis revealed a significant effect of Focus, primarily localized over fronto-central sites (AF7, AF8, F7, F8, F6, F5, Fz, FC1, FCz, Cz, C2, and C1) and centro-parietal sites (CP1, CP2, P2, P7, C1, Cz, and C2) between 110ms to 310ms, with a peak around 280ms post-stimulus. A later effect of Focus was also observed between 300 and 700ms mainly over occipital (O1, O2, and Oz) and fronto-central sites (FC2, FC1, FCz, Cz, C1, C3, F8, F2, F1, and Fz; p-values for these effects ranged from 0.0497 to 0.0002). Additionally, a significant main effect of Congruence was found at centro-parietal sites (CPz, CP1, CP4, Pz, P1, and P2) between 280ms and 550ms, peaking around 400ms post-stimulus (p-values ranged from 0.0326 to 0.0015).

Finally, a significant effect of speed was observed at fronto-central sites (F6, and FC5) between 380 and 530ms (p-values ranged from 0.0409 to 0.0202), along with a significant interaction between Congruence x Speed at frontal sites (AF4, and F6) between 120 and 350ms (p-values ranged from 0.0302 to 0.0041). No other effects were significant ($p > 0.2221$; see Figure 2).

Figure 2. Plot ERP Waveforms of All Electrodes (A), Scalp Mapping for Focus and Congruence Conditions (B), and FMUT map of Focus and Congruence Effects (C and D)



Note. For grand averaged ERPs for all electrodes (A) elicited the cerebral activity between -100 and 700ms post-stimulus. For C and D graphs: coloured sections correspond to the significant F-values as indicated by the colour bar.

These results indicate that the main effects of Focus and Congruence on neural processing are predominantly localized to specific scalp sites and time windows, whereas the effects of speed and its interaction with congruence are present but more spatially restricted and temporally limited. This pattern of findings guided our subsequent, more focused analyses on specific electrodes and time windows.

Specifically, FMUT analyses focused on two time-windows: (1) early time-window from 110 to 310ms post-stimulus, with a peak around 280ms, and (2) late time-window from 300 to 700ms, peaking around 400ms.

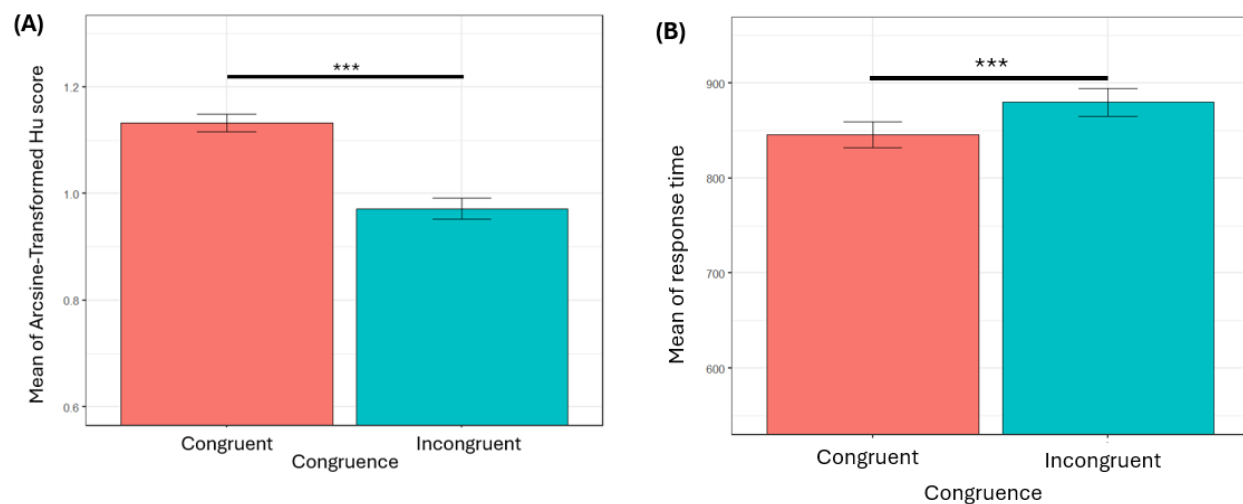
Hypothesis 1: Contextual Effect

A 2 (Focus) x 2 (Speed) x 2 (Congruency) repeated measures ANOVA on the arcsine-transformed Hu scores revealed a significant main effect of congruence ($F[1, 29] = 84.292$, $p < 0.001$, $\eta^2 = 0.173$). Bonferroni-adjusted paired t -tests showed that emotions were better recognized during congruent conditions rather than incongruent conditions ($t[29] = 9.181$, $p < 0.001$, $d = 0.988$, CI [0.125, 0.197]). Bayesian repeated measures ANOVA indicated strong evidence that the best model, i.e., the one with the largest Bayes Factor, includes the main effect of congruence, focus, and focus * congruence interaction ($BF_{10} = 1.208e+20$). Analysis showed strong evidence for including of congruence factor ($BF_{inclusion} = 1.865e+7$; see Figure 3).

The same repeated measures ANOVA on reaction times also revealed a significant main effect of congruence ($F[1, 29] = 36.794$, $p < 0.001$, $\eta^2 = 0.016$). Bonferroni-adjusted paired t -tests showed that response times were significantly longer for incongruent compared to congruent stimuli ($t[29] = -6.066$, $p < 0.001$, $d = -0.309$, CI [-45.651, -22.629]). Bayesian repeated measures ANOVA indicated strong evidence that the best model, i.e., the one with the largest Bayes Factor, includes the main effect of congruence, focus, speed and the focus * congruence

interaction ($BF_{10} = 2.897e+16$). Analyses provided strong evidence for the including of congruence factor ($BF_{inclusion} = 21845.251$; see Figure 3).

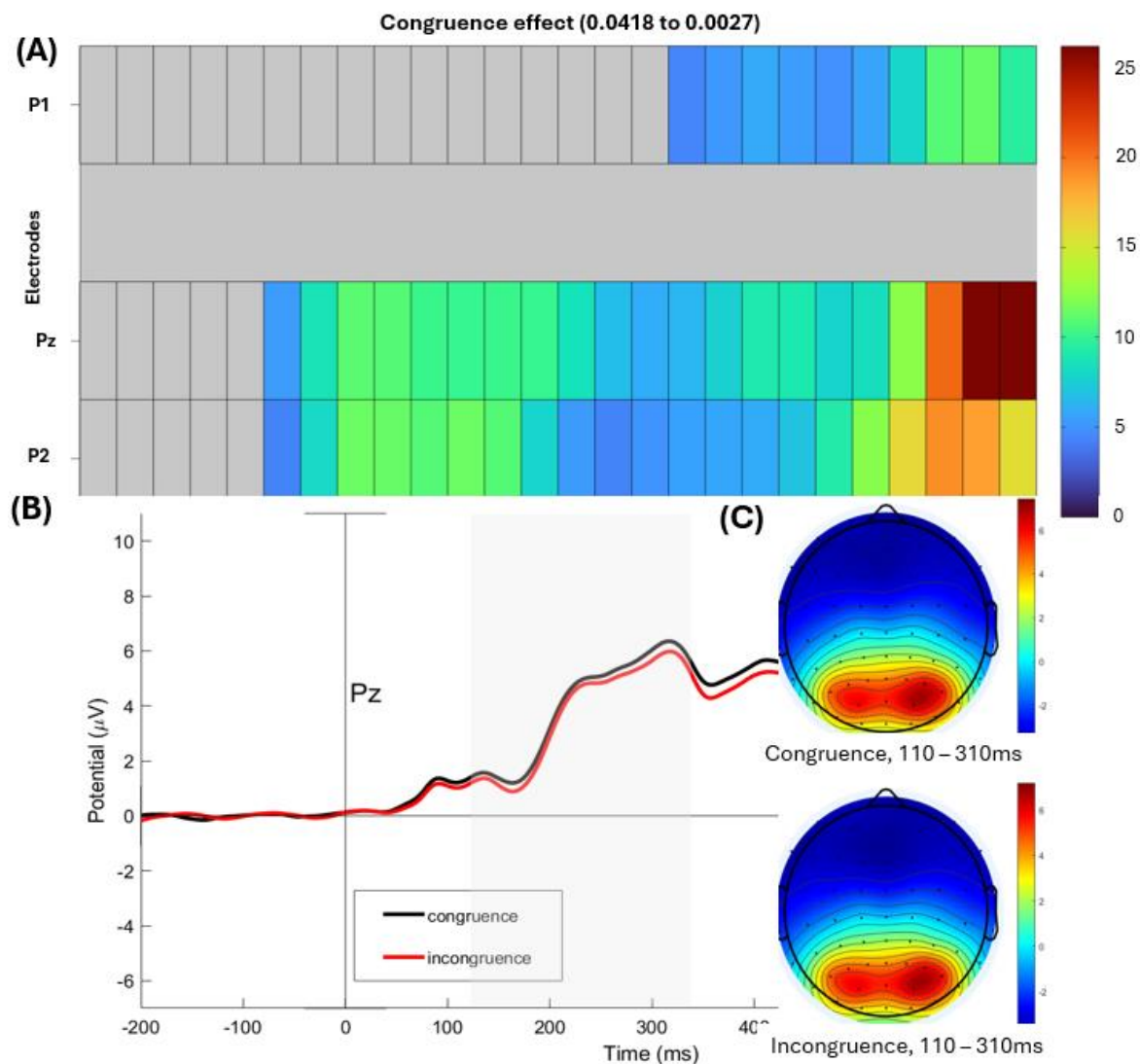
Figure 3. Histogram of Means of Arcsine Transformation (A) and Response Time (B) for the Congruence



Note. *** $p < 0.001$.

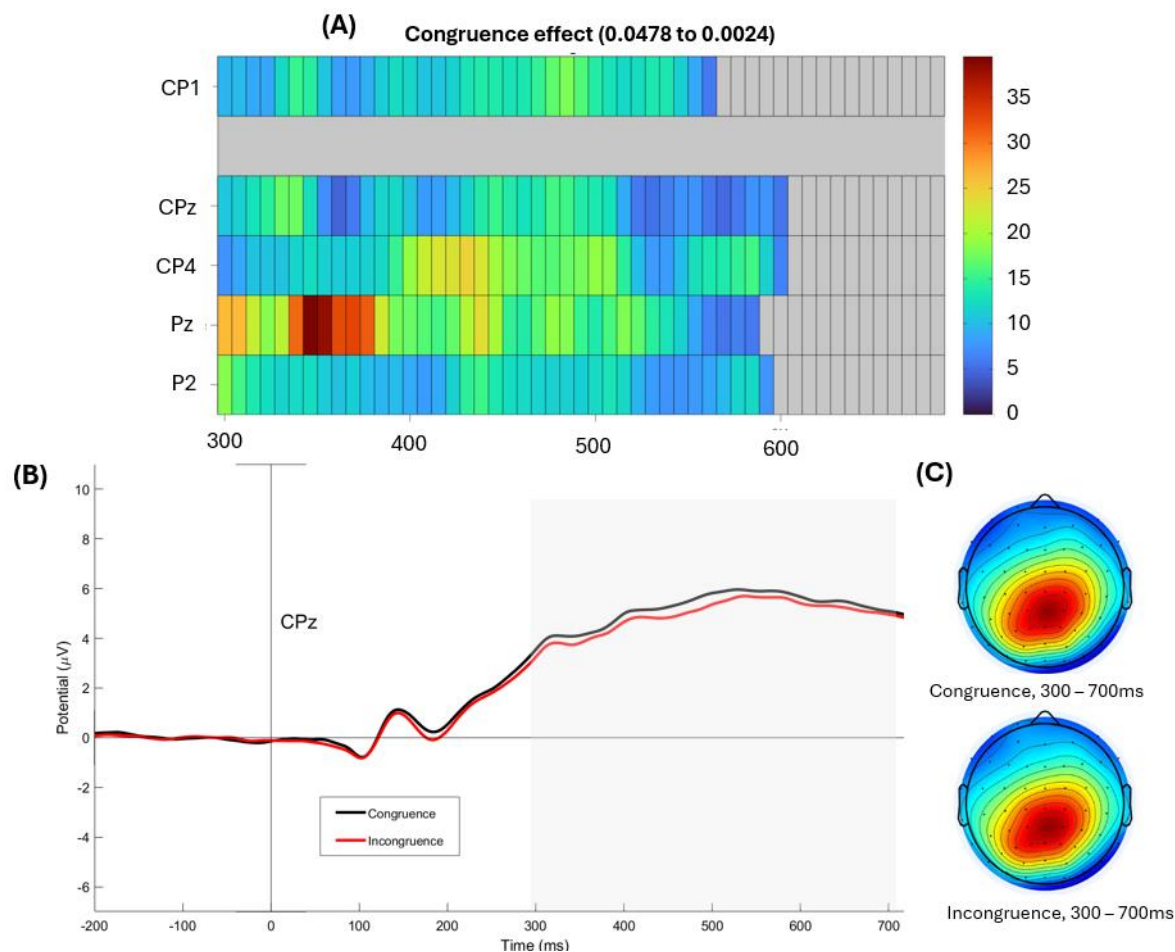
Finally, on the time-window 110-310ms, analyses FMUT on ERPs showed a significant main effect of congruence (p-values ranged from 0.0418 to 0.0027), primarily over parietal sites (P1, Pz and P2). More specifically, amplitudes were also higher in the congruent condition (Mean = 3.93uV +/- 2.2) than in the incongruent condition (Mean = 3.69uV +/- 2.25). Bayesian repeated measures ANOVA provided strong evidence in favor of a best model including main effects of Focus, Speed, Congruence, and Focus * Congruence and Speed * Congruence interactions ($BF_{10} = 82315.923$). The analysis provided strong evidence for including the congruence factor ($BF_{inclusion} = 38.873$; see Figure 4).

Figure 4. FMUT map of Congruence Effect (A), Grand Averaged ERP on the Parietal Sites (B), Scalp map (C)



On late time-window (300-700ms), analyses FMUT on ERPs showed at centro-parietal sites (CP1, CP4, CPz, P2, and Pz) revealed a significant main effect of congruence (p-values ranged from 0.0478 to 0.0024). More specifically, amplitudes were also higher in the congruent condition (Mean = 4.398 μV +/- 1.903) than in the incongruent condition (Mean = 4.083 μV +/- 1.90). Bayesian repeated measures ANOVA showed strong evidence that the best model includes a main effect of Focus and Congruence ($BF_{10} = 6201.146$). The analysis provided strong evidence for the inclusion of the congruence factor ($BF_{inclusion} = 1070.670$; see Figure 5).

Figure 5. FMUT map of Congruence Effect (A), Grand Averaged ERP on the Centro-Parietal Sites(B), Scalp map (C)

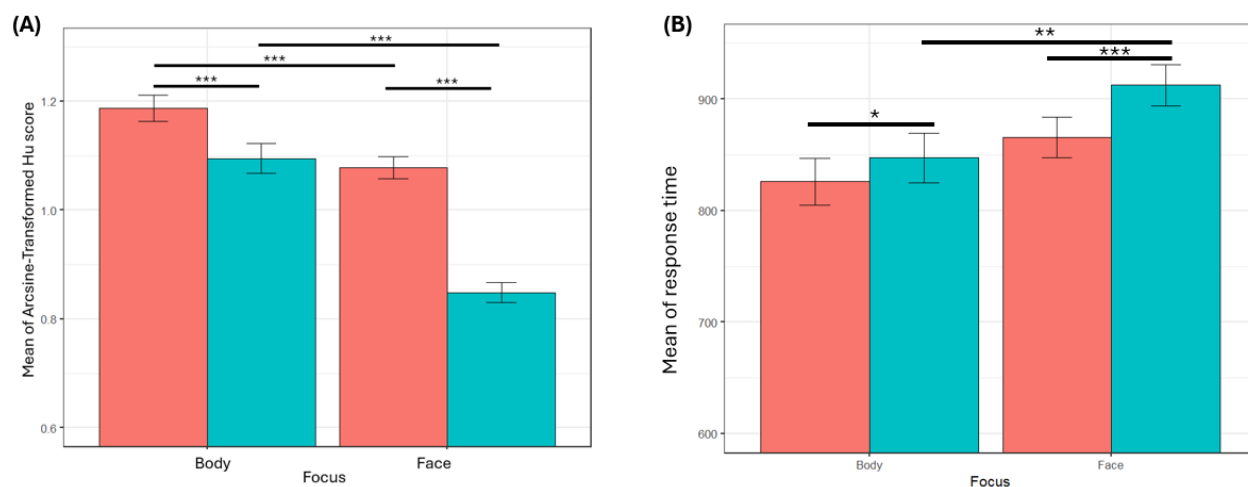


Hypothesis 2: Attentional Focus Effect

The repeated measures ANOVA on arcsine transformation to the Hu Score revealed a significant main effect of focus ($F[1, 29] = 68.854, p < 0.001, \eta^2 = 0.212$) and an interaction effect between focus and congruence ($F[1, 29] = 26.449, p < 0.001, \eta^2 = 0.032$). Bonferroni-adjusted paired t -tests showed that emotions were better recognized when attention was focused on bodily expressions rather than facial ones ($t[29] = 8.298, p < 0.001, d = 1.094, CI [0.134, 0.222]$). The simple effects analysis also showed a significant effect of congruence when participants focused on the face ($F[1, 29] = 256.663, p < 0.001$) and on the body posture ($F[1, 29] = 11.101, p = 0.002$), suggesting that emotional expressions were more accurately

recognized in congruent compared to incongruent conditions, with a larger difference observed when attention was directed to the facial rather than the bodily expression. Bayesian analyses indicated strong evidence for including focus factor ($BF_{inclusion} = 3.335e+6$) and the congruence * focus interaction ($BF_{inclusion} = 661.778$; see Figure 6).

Figure 6. Histogram of Means of Arcsine Transformation (A) and Response Time (B) for the Congruence and Focus Interaction



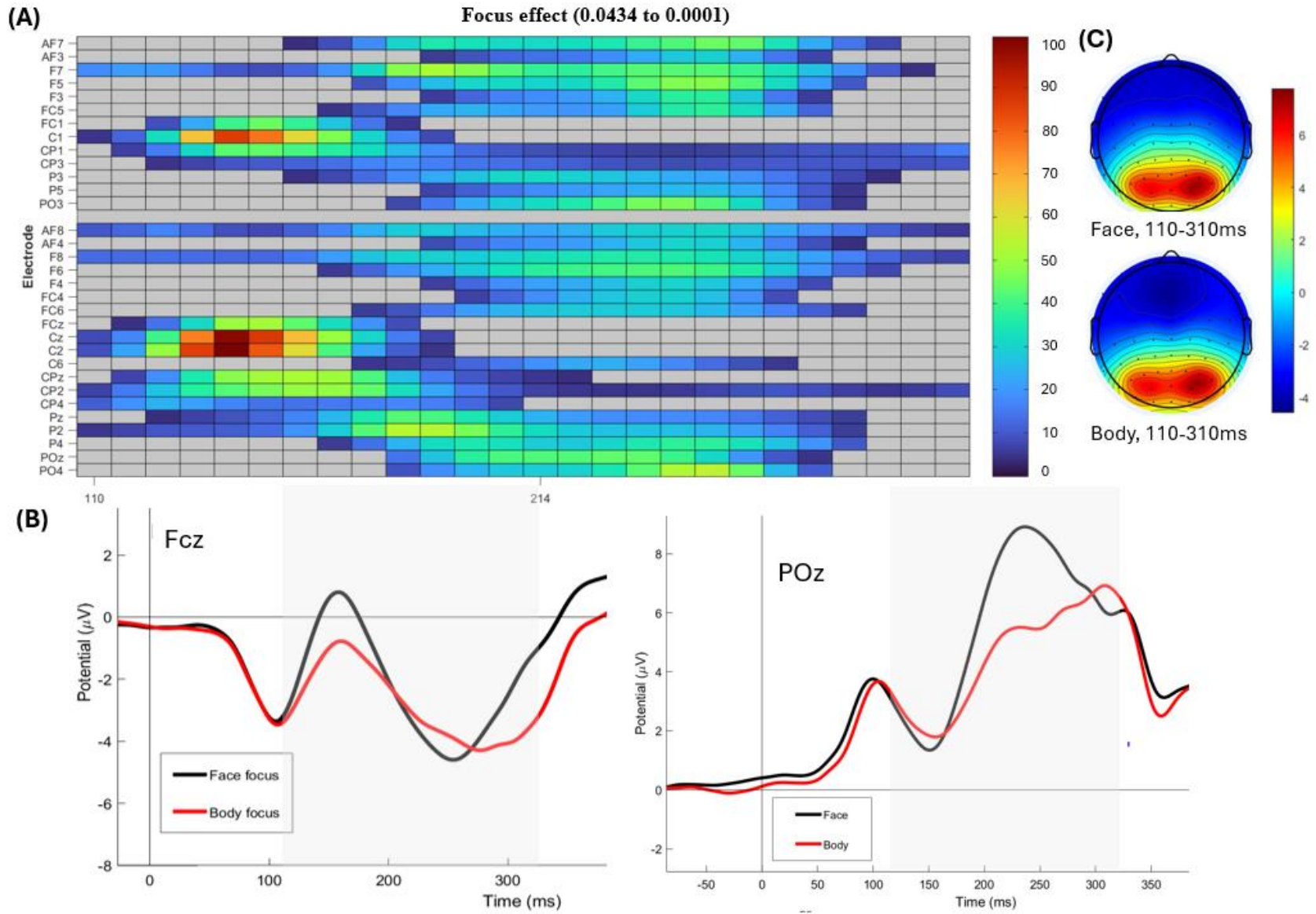
Note. * $p < 0.01$. ** $p < 0.001$. *** $p < 0.0001$.

For reaction times, the repeated measures ANOVA revealed a significant main effect of focus ($F[1, 29] = 11.153, p = 0.002, \eta^2 = 0.037$), and an interaction effect between focus and congruence ($F[1, 29] = 11.676, p = 0.002, \eta^2 = 0.002$). Bonferroni-adjusted paired t -tests showed that response times were longer during attentional focus on the face rather than body ($t[29] = -3.340, p = 0.002, d = -0.475, CI [-84.488, -20.309]$). The simple effects analysis also showed a significant effect of congruence when participants focused on the face ($F[1, 29] = 40.457, p < 0.001$) and on the body posture ($F[1, 29] = 12.560, p = 0.001$), suggesting that the response time were longer in incongruent compared to congruent conditions, with a larger difference observed when attention was directed to the bodily rather than the facial expression. Bayesian repeated measures ANOVA indicated strong evidence for including focus

factor ($BF_{inclusion} = 13.613$) and moderate evidence for including of the congruence * focus interaction ($BF_{inclusion} = 661.778$; see Figure 6).

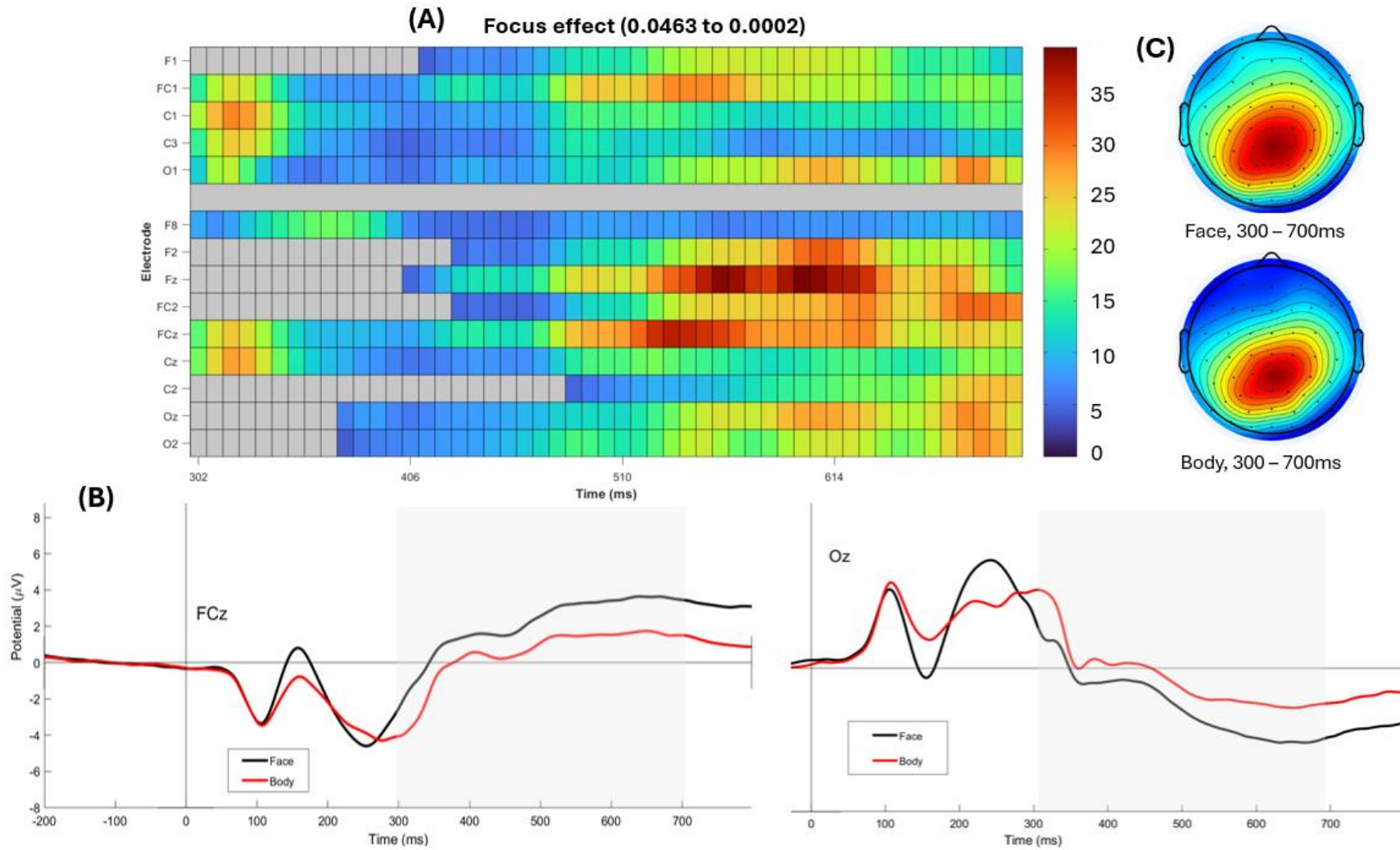
Finally, on the time-window 110-310ms, analyses FMUT on ERPs showed a main effect of Focus (p-values ranged from 0.0434 to 0.0001), predominantly over fronto-central sites (AF3, AF4, AF7, AF8, F3, F4, F5, F6, F7, F8, FC1, FC4, FC5, FC6, FCz, C1, C2, C6, and Cz) and centro-parieto-occipital sites (CP1, CP2, CP3, CP4, CPz, Pz, P2, P3, P4, P5, PO3, PO4, and POz). At fronto-central and centro-parieto-occipital sites, amplitudes were greater in the face-focus condition (*Fronto-central*: Mean = $-2.684\mu V \pm 1.654$; *centro-parieto-occipital*: Mean = $4.224\mu V \pm 1.796$) compared to the body-focus condition (*Fronto-central*: Mean = $-2.059\mu V \pm 1.504$; *centro-parieto-occipital*: Mean = $3.135\mu V \pm 1.40$). Bayesian repeated measures ANOVA indicated on the fronto-central ($BF_{10} = 3615.645$) and centro-parieto-occipital ($BF_{10} = 231545.138$) sites, strong evidence in favor of a best model including main effects of Focus, Speed, Congruence, and Focus x Congruence and Speed * Congruence interactions. In fronto-central ($BF_{inclusion} = 16.327$) and centro-parieto-occipital ($BF_{inclusion} = 12.187$) sites, analyses indicated strong evidence for the inclusion of the Focus (see Figure 7).

Figure 7. FMUT map of Focus Effect (A), Grand Averaged ERP on the Fronto-Central and Parieto-Occipital sites(B), Scalp map (C)



On the time-window 300-700ms, analyses FMUT on ERPs showed a significant main effect of Focus (p-values ranged from 0.0463 to 0.0002) across the fronto-central site (Fz, F1, F2, F8, FCz, FC1, FC2, C1, C2, C3, and Cz) and occipital site (O1, O2, and Oz). More specifically, amplitudes were greater in the face-focus condition (*Fronto-central*: Mean = 0.179uV +/- 1.620; *occipital*: Mean = -1.963uV +/- 2.814) compared to the body-focus condition (*Fronto-central*: Mean = -0.368uV +/- 1.576; *occipital*: Mean = -0.419uV +/- 2.443). Bayesian repeated measures ANOVA revealed strong evidence that the best model at the fronto-central (BF10 = 10.06) and occipital sites (BF10 = 157.044) includes a main effect of Focus. At the centro-parietal sites, strong evidence for a model including both focus and congruence (BF10 = 6201.146). Inclusion Bayes factors indicated strong evidence for the inclusion of the Focus at the fronto-central (BF_{inclusion} = 12.155) and occipital (BF_{inclusion} = 165.907) sites, while at the centro-parieto sites, the evidence for including focus factor was weaker but still anecdotal to moderate (BF_{inclusion} = 2.528; see Figure 8).

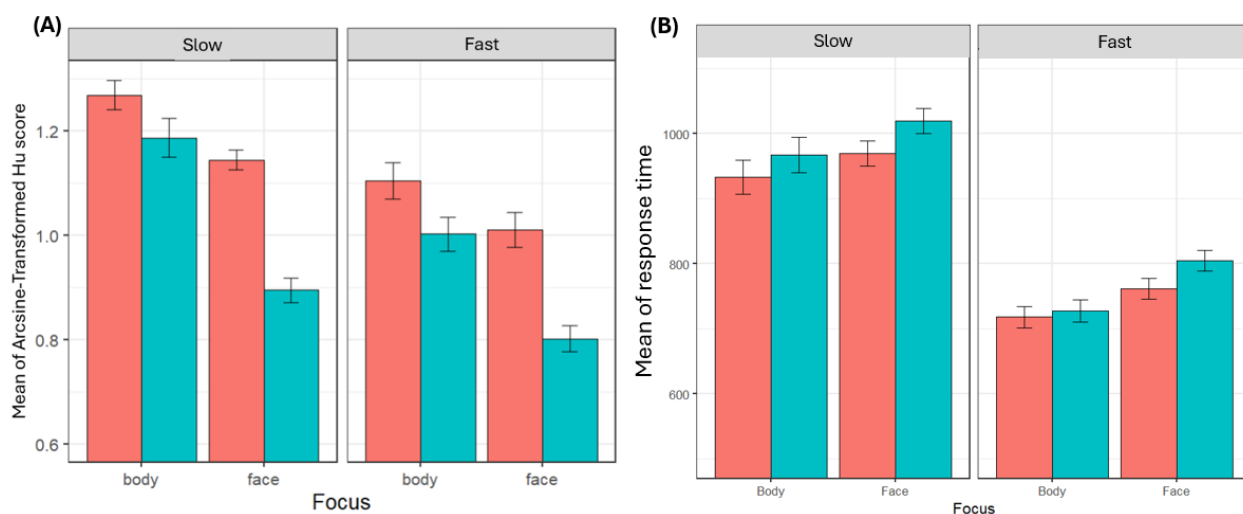
Figure 8. FMUT map of Focus Effect (A), Grand Averaged ERP on the Fronto-Central and Occipital Sites(B), Scalp map (C)



Hypothesis 3: Speed Effect

The repeated measures ANOVA on the arcsine-transformed Hu scores revealed a significant main effect of speed ($F[1, 29] = 29.816, p < 0.001, \eta^2 = 0.139$). Emotions were better recognized when the task was performed slowly rather than quickly ($t[29] = 5.460, p < 0.001, d = 0.887, CI [0.090, 0.198]$). Bayesian repeated measures ANOVA indicated strong evidence for including the speed factor ($BF_{inclusion} = 2916.472$). Additionally, there was weak evidence of the inclusion of congruence X speed interaction ($BF_{inclusion} = 0.315$). Bayesian analysis provided weak evidence against the inclusion of models with speed ($BF_{inclusion} < 0.997$; see Figure 9).

Figure 9. Histogram of Means of Arcsine Transformation (A) and Response Time (B) for the Speed *Congruence*Focus Interaction



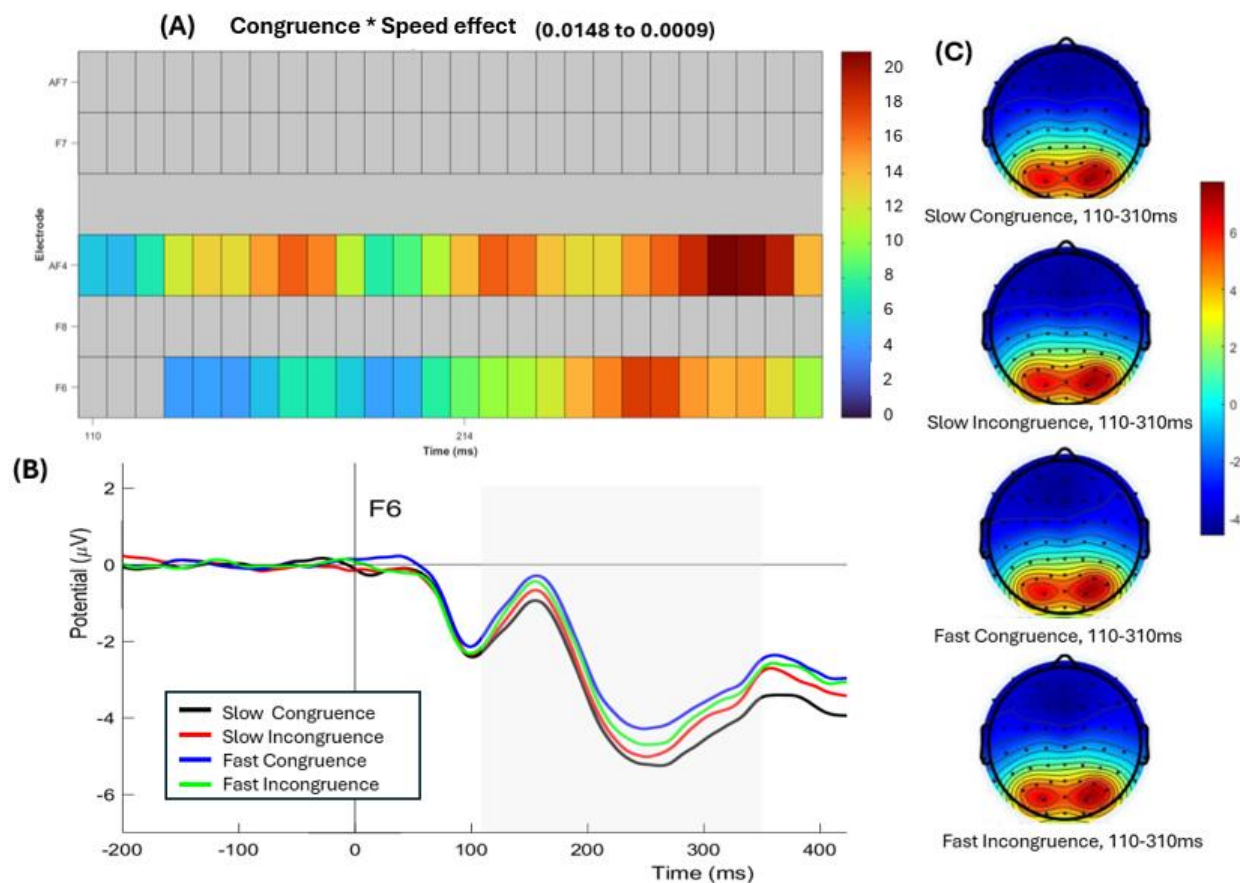
The same repeated measures ANOVA on reaction times revealed a significant main effect of speed ($F[1, 29] = 160.680, p < 0.001, \eta^2 = 0.655$). Bonferroni-adjusted paired t -tests showed that response times were longer during slow condition rather than fast condition ($t[29] = 12.676, p < 0.001, d = 1.986, CI [183.957, 254.740]$). Bayesian repeated measures ANOVA indicated strong evidence in favor of speed factor ($BF_{inclusion} = 2.554e+10$). Additionally, there was weak evidence of the inclusion of congruence * speed interaction ($BF_{inclusion} = 0.760$). Bayesian

analysis provided strong evidence against the inclusion of models with speed ($BF_{inclusion} < 0.084$; see Figure 9).

Finally, on the time-window 110-310ms, analyses FMUT on ERPs showed, at the frontal sites (F6, F7, F8, and AF7), a significant main effect of speed (p-values ranged from 0.0473 to 0.0104) and an interaction between congruence and speed (p-values ranged from 0.0148 to 0.0009) at frontal sites (F6 and AF4). More specifically, in the slow condition, amplitudes were higher for the congruent condition (Mean = $-3.54\mu V \pm 2.35$) compared to incongruent condition (Mean = $-3.18\mu V \pm 2.40$). No difference between the congruent and incongruent conditions was observed in the fast condition. Bayesian repeated measures ANOVA revealed strong evidence that the best model includes a main effect of Focus, Speed, Congruence, Focus*Speed, Focus*Congruence, Speed *Congruence, Focus*Speed *Congruence ($BF_{10} = 6.157 \times 10^6$). Analysis provided strong evidence for including the speed factor ($BF_{inclusion} = 141.218$) and Speed *Congruence ($BF_{inclusion} = 10.737$; see Figure 10).

On the time-window 300-700ms, analyses FMUT on ERPs show no significant effect with speed condition ($ps > 0.4$).

Figure 10. FMUT map of Congruence * Speed Effect (A), Grand Averaged ERP on the Frontal Sites(B), Scalp map (C)



Discussion

The present study investigated the automatic nature of the contextual effect between facial and bodily expressions by manipulating one criterion of automaticity discussed in the literature, processing speed (Moors & De Houwer, 2006). As expected from our *first hypothesis*, results showed that emotions were recognized more accurately and more quickly when facial and bodily expressions were congruent compared to when they were incongruent. However, contrary to our hypothesis, early brain activity (110-310ms) was greater for emotionally congruent than incongruent face-body pairings. Consistent with our *second hypothesis*, bodily expressions were recognized more accurately and elicited faster responses than facial expressions, suggesting a stronger influence of body postures on the recognition of facial emotions than vice versa. In

terms of ERP components, as expected, greater later brain responses (300-700ms) were observed for congruent versus incongruent stimuli when attention was directed toward facial expressions. This effect was also present when attention was focused on bodily expressions. Finally, in line with our *third hypothesis*, no behavioral interaction was found between congruence and processing speed, suggesting that the contextual effect is not modulated by speed at the behavioral level. However, at the neural level, greater ERP amplitudes (110-310ms) for congruent versus incongruent stimuli were found only in the slow condition, not in the fast one, thereby challenging the assumption of automaticity. We discuss these effects in the following sections.

The Contextual Effect

In line with earlier behavioral studies, our results indicate that the recognition of emotional expressions is more effective when facial and bodily cues convey the same emotion, and tends to be disrupted when these cues are incongruent (Lecker et al., 2020; Mondloch et al., 2013; Puffet & Rigoulot, 2025c; Willis et al., 2011). Specifically, participants were more accurate and responded more quickly in congruent conditions than in incongruent ones. These findings reinforce the notion that emotional recognition benefits from the joint processing of facial and bodily signals, which appear to be integrated (Kret, Stekelenburg et al., 2013; Meeren et al., 2005). Such contextual effect supports the existence of a shared and unified mechanism for recognition emotional information across different expressive channels (de Gelder et al., 2015). Moreover, these findings are also consistent with previous research showing that bodily expressions help reduce the ambiguity of facial expressions, thereby facilitating emotional recognition (Karaaslan et al., 2020; Meeren et al., 2005). Conversely, inadequate integration of emotional context, such as body language, can negatively impact the recognition of facial expressions. This has been observed in certain clinical conditions, such as alcohol use disorder,

where emotional recognition is impaired due to a diminished ability to incorporate contextual bodily cues into the recognition of facial expressions (Brion et al., 2017; Creupelandt et al., 2020). These observations underscore the critical role of emotional context in the accurate recognition of emotions.

Our neural findings during the early stages of emotional processing differ from those reported in several previous studies on face-body emotional integration (Chen et al., 2022; Gu et al., 2013; Li, 2021; Wang et al., 2017). While prior research often found enhanced early brain responses, particularly in the N2 component, to incongruent face-body pairs, interpreted as early conflict detection (Gu et al., 2013), our study instead revealed greater neural activity for emotionally congruent stimuli within the 110-310ms time window. This pattern suggests that congruent face-body emotional cues, which provide coherent and redundant information, may be processed more efficiently at early perceptual stages. Supporting this interpretation, similar findings have been reported in multimodal emotion integration studies involving facial and vocal expressions, where enhanced P2 amplitudes for congruent stimuli were observed (Balconi & Carrera, 2011; Pourtois et al., 2002), reflecting a preferential processing of emotionally consistent cues. Furthermore, our neurophysiological results align with the behavioral data, which also showed faster and more accurate emotion recognition for congruent compared to incongruent face-body pairs. Together, these findings suggest that rather than being driven by early conflict detection, emotional face-body integration may benefit from an early facilitation mechanism when they are congruent, leading to more effective recognition of emotional content. However, this result needs to be interpreted in the context of the specific experimental design used in the present study. Specifically, instructing participants to respond as quickly as possible in the fast condition, combined with the brief presentation of the stimuli (200ms), may have introduced a response bias that could confound the interpretation of the results. Nevertheless, the differences

observed with previous studies using facial and bodily expressions cannot be fully explained by these procedural aspects, as the prior studies also employed a short stimulus presentation and required fast responses from participants (Gu et al., 2013; Li, 2021).

The finding of enhanced early processing for congruent expressions (Balconi & Carrera, 2011) contrasts with the early conflict detection reported in other studies on face-body emotional incongruence (Chen et al., 2022; Gu et al., 2013). This difference may partly be explained by variations in the quality of the stimuli used. In the present study, we employed a high-quality set of stimuli (Puffet & Rigoulot, 2025b) specifically designed to reduce perceptual biases commonly introduced by image manipulations. Many previous studies have created congruent and incongruent face-body combinations by merging facial and bodily expressions from different image databases, often in grayscale. This process can lead to image degradation and the introduction of perceptual inconsistencies that may confound the interpretation of emotional incongruence. For example, mismatched lighting, disproportionate scaling between face and body or poorly cropped facial images can create perceptual artifacts that exaggerate the incongruence due to physical, rather than emotional, mismatches. This is particularly relevant given that early ERP components like the P100 are known to be sensitive to low-level visual features (e.g., color, brightness, contrasts) (Puce et al., 2013; Yang et al., 2020), and the P200 has been shown to respond to implausible or mismatched face-body configurations (Hu et al., 2020). Therefore, using ecologically more valid, high-quality stimuli is essential to minimize perceptual biases and ensure that observed neural effects reflect genuine emotional integration rather than artifacts of stimulus construction. It is thus possible that the use of higher-quality stimuli in the present study may account for the differences observed compared to previous research.

The Role of Attentional Focus on the Facial and Bodily Expressions Integration

Attentional focus of participants was also manipulated. Specifically, participants were instructed to direct their attention and recognize emotions either from facial or bodily expressions. This design allowed for the examination of the bidirectionality of the contextual effect (Lecker et al., 2020), and whether this directional influence occurs automatically, regardless of the attentional focus. The findings of this study indicate that emotions are recognized more quickly and accurately when attention is directed toward bodily expressions rather than facial expressions. The greater ease in recognizing bodily expressions may be explained by several factors. First, body expressions tend to be more homogenous than facial expression. For instance, expressions of anger through body posture typically involve clenched fists and muscle tension, with limited variability in how this emotion is physically conveyed. In contrast, facial expressions are ambiguous and exhibit variability across individuals and contexts, which may complicate their recognition (de Gelder et al., 2015). Second, bodily expressions occupied a larger visual area on the screen compared to facial expressions, which may have further facilitated their recognition (Zhang et al., 2015), especially considering the brief exposure time used in this experimental design.

Regarding the contextual effect, our results support the idea of bidirectionality: emotional recognition was consistently faster and more accurate in congruent compared to incongruent conditions, regardless of whether attention was directed to the face or the body. This suggests that bodily expressions can serve as context for interpreting facial emotion, but also that the reverse is possible (Kret, Roelofs et al., 2013; Lecker et al., 2020). However, this bidirectional influence appears to be asymmetrical. Specifically, facial expressions were more susceptible to influence from bodily cues than vice versa (Lecker et al., 2020). One possible explanation for this asymmetry lies in the natural attentional bias toward faces in human social perception.

Individuals tend to rely more heavily on facial cues when interpreting emotions and often underestimate the importance of other modalities, such as body language, which is frequently treated as a contextual source of information. This bias may explain why facial expressions exert weaker contextual effects on bodily expressions because in everyday life, faces are rarely used as contextual cues for interpreting body posture. As a result, using facial expressions as a contextual source for body emotion recognition may not be intuitive and could reduce the integration of contextual cues when attention is focused on the body (Aviezer et al., 2012b; Kret, Stekelenburg et al., 2013; Lecker et al., 2020).

At the neural level, both early (around 230ms post-stimulus) and late (from 300ms post-stimulus) brain responses show greater amplitude when attention is directed toward facial expressions compared to body expressions. These findings are consistent with previous research suggesting that facial expressions are processed rapidly and with high priority in the brain (Meeren et al., 2005; Stekelenburg & de Gelder, 2004; Zhang et al., 2015). Neural activity thus appears to be primarily engaged in processing facial cues, reflecting lower efficiency in the processing of body expressions (Zhang et al., 2015). This pattern indicates that attention is preferentially allocated to facial expressions and remains engaged during later stages of processing. These neural results align with behavioral data showing better recognition accuracy for facial expressions and a weaker influence of facial cues on the recognition of body expressions. Altogether, this suggests that facial expressions are prioritized as the main emotional target, rather than being interpreted as contextual information (Lecker et al., 2020). Finally, from approximately 300ms post-stimulus, we also observed sustained neural activity that was greater for emotionally congruent compared to incongruent stimuli, regardless of whether attention was directed toward facial or bodily expressions, suggesting no attentional focus-related difference in the processing of incongruence.

The Role of Speed on the Facial and Bodily Expressions Integration

In the present study, the processing speed of emotional expressions was also manipulated. Stimuli were briefly displayed for 200ms, and participants were required to perform the emotion recognition task under either fast or unrestricted response conditions. This manipulation of response time was important to consider, as a short stimulus presentation does not necessarily imply that emotional processing is also brief, it may continue beyond the initial perceptual stage (Moors & De Houwer, 2006).

At the behavioral level, the results showed no interaction between speed and congruence, neither for accuracy nor for response time. This suggests that the integration of facial and bodily expressions occurs similarly whether participants have a short or long time to perform the emotion recognition task. Such findings indicate that emotional integration happens rapidly, consistent with the characteristics of automatic processing and in line with previous research (Meeren et al., 2005). These results further support the hypothesis that the integration of facial and bodily emotional expressions operates automatically (Aviezer et al., 2011; Karaaslan et al., 2020; Meeren et al., 2005).

At the neural level, the results are difficult to interpret due to discrepancies between the findings from FMUT and Bayesian analyses. Moreover, no clear interaction between speed and congruence was found, apart from a weak early effect (between 110 and 310ms) at fronto-central and centro-parieto-occipital sites. This effect indicates a contextual influence in the slow condition, but not during the fast condition, raising questions about the automaticity of facial-body emotional integration. Indeed, the fact that the contextual effect between facial and bodily expressions occurs only in the slow condition suggests that when participants have limited time to respond, the interaction between facial and bodily cues may not occur, implying that a minimum processing time is required for this integration to take place. However, we believe that

this result needs to be considered carefully. Indeed, while the Bayesian ANOVA provided moderate to strong evidence supporting the inclusion of the interaction term in the model, the FMUT detected a statistically significant interaction between processing speed and congruence at only two electrodes (F6 and AF4) during early processing (110-310ms). This discrepancy, where the FMUT shows a weak effect limited to two electrodes, yet the Bayesian ANOVA strongly supports the interaction, highlights a fundamental difference between frequentist and Bayesian approaches. The FMUT relies on fixed significance thresholds ($p < 0.05$) (Fields & Kuperberg, 2020). In contrast, Bayesian ANOVA evaluates the relative evidence for different model (van den Bergh et al., 2023; van Doorn et al., 2021), and can highlight potentially relevant effect even when the evidence is not strong enough to draw firm conclusions, especially when an effect slightly improves overall model fit. In this case, the $BF_{inclusion}$ for the interaction terms suggests that the strong evidence support the interaction between Speed and Congruence (van den Bergh et al., 2023; van Doorn et al., 2021). Furthermore, the FMUT showed that only a small and spatially diffuse number of electrodes exhibited a significant interaction between Speed and Congruence. These findings suggest that, even at the neural level, the contextual effect between facial and bodily cues unfolds in an automatic manner according to the speed criterion. However, the weak interaction effect observed between speed and congruence further supports the need for cautious interpretation and underscores the value of combining both statistical approaches for a more comprehensive understanding of the data.

Limitation

Although the present study supports previous findings (Meeren et al., 2005) and contributes new insights into the processing speed of facial and bodily emotional cues, it also has certain limitations that temper the conclusions. Specifically, while response of participants speed was manipulated (fast vs. slow conditions), the presentation duration of facial and bodily

expressions remained fixed at 200ms across all conditions. Although his design is methodologically sound, it does not allow for examination of whether longer presentation times would modulate the contextual effect between facial and bodily expressions. An experimental design varying both stimulus duration and response time would provide more robust interpretation. Future studies should therefore continue investigating the processing speed of facial and bodily emotional cues to strengthen and refine our understanding of these mechanisms.

Conclusion

These emotions are conveyed through both facial and bodily expressions, two distinct but complementary emotional channels that jointly contribute to the accurate recognition of emotions. The present study investigates this contextual effect, with a particular focus on its automaticity by manipulating the time available for emotional processing. Our findings, based on both behavioral and neural measures, provide converging evidence that the integration of facial and bodily expressions occurs automatically. Specifically, the time constraints imposed on the recognition task did not modulate the contextual influence of one channel on the other, indicating that this integration does not rely on deliberate or controlled processing. Moreover, the contextual effect appears to be bidirectional yet asymmetrical, while facial expression recognition is more strongly influenced by body posture than the reverse. To our knowledge, this is the first study to investigate the automatic, bidirectional contextual integration between facial and bodily expressions using both behavioral and neural measures while varying processing speed. These results advance our understanding of the underlying mechanisms of emotional perception and highlight the importance of considering multimodal integration in emotion recognition.

Declaration of interest statement

The authors report there are no competing interests to declare.

References

- Abo Foul, Y., Eitan, R., & Aviezer, H. (2018). Perceiving emotionally incongruent cues from faces and bodies: Older adults get the whole picture. *Psychology and Aging, 33*(4), 660-666. <https://doi.org/10.1037/pag0000255>
- Aviezer, H., Bentin, S., Dudarev, V., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion, 11*(6), 1406-1414. <https://doi.org/10.1037/a0023578>
- Aviezer, H., Trope, Y., & Todorov, A. (2012a). Holistic person processing: Faces with bodies tell the whole story. *Journal of Personality and Social Psychology, 103*(1), 20-37. <https://doi.org/10.1037/a0027411>
- Aviezer, H., Trope, Y., & Todorov, A. (2012b). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225-1229. <https://doi.org/10.1126/science.1224313>
- Balconi, M., & Carrera, A. (2011). Cross-modal integration of emotional face and voice in congruous and incongruous pairs: The P2 ERP effect. *Journal of Cognitive Psychology, 23*(1), 132-139. <https://doi.org/10.1080/20445911.2011.473560>
- Brion, M., D'Hondt, F., Lannoy, S., Pitel, A. L., Davidoff, D. A., & Maurage, P. (2017). Crossmodal processing of emotions in alcohol-dependence and Korsakoff syndrome. *Cognitive Neuropsychiatry, 22*(5), 436-451. <https://doi.org/10.1080/13546805.2017.1373639>
- Chen, T., Sun, Y., Feng, C., & Feng, W. (2022). In identifying the source of the incongruent effect. *Journal of Psychophysiology, 36*(3), 167-176. <https://doi.org/10.1027/0269-8803/a000290>

- Creupelandt, C., D'Hondt, F., de Timary, P., Falagiarda, F., Collignon, O., & Maurage, P. (2020). Selective visual and crossmodal impairment in the discrimination of anger and fear expressions in severe alcohol use disorder. *Drug and Alcohol Dependence*, 213, Article 108079. <https://doi.org/10.1016/j.drugalcdep.2020.108079>
- de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3475-3484. <https://doi.org/10.1098/rstb.2009.0190>
- de Gelder, B., de Borst, A. W., & Watson, R. (2015). The perception of emotion in body expressions. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(2), 149-158. <https://doi.org/10.1002/wcs.1335>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9-21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Durston, A. J., & Itier, R. J. (2021). The early processing of fearful and happy facial expressions is independent of task demands – Support from mass univariate analyses. *Brain research*, 1765, Article 147505. <https://doi.org/10.1016/j.brainres.2021.147505>
- Fields, E. C., & Kuperberg, G. R. (2020). Having your cake and eating it too: Flexibility and power with mass univariate statistics for ERP data. *Psychophysiology*, 57(2), Article e13468. <https://doi.org/10.1111/psyp.13468>
- Focker, J., Gondan, M., & Roder, B. (2011). Preattentive processing of audio-visual emotional signals. *Acta Psychologica*, 137(1), 36-47. <https://doi.org/10.1016/j.actpsy.2011.02.004>
- Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: A validation study. *Cognition & Emotion*, 22(6), 1094-1118. <https://doi.org/10.1080/02699930701626582>

- Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711-1725. <https://doi.org/10.1111/j.1469-8986.2011.01273.x>
- Gu, Y., Mai, X., & Luo, Y. J. (2013). Do bodily expressions compete with facial expressions? Time course of integration of emotional signals from the face and the body. *PLoS One*, *8*(7), Article e66762. <https://doi.org/10.1371/journal.pone.0066762>
- Hu, Y., Baragchizadeh, A., & O'Toole, A. J. (2020). Integrating faces and bodies: Psychological and neural perspectives on whole person perception. *Neuroscience & Biobehavioral Reviews*, *112*, 472-486. <https://doi.org/10.1016/j.neubiorev.2020.02.021>
- Hudson, A., Durston, A. J., McCrackin, S. D., & Itier, R. J. (2021). Emotion, gender and gaze discrimination tasks do not differentially impact the neural processing of angry or happy facial expressions: A mass univariate ERP analysis. *Brain Topography*, *34*(6), 813-833. <https://doi.org/10.1007/s10548-021-00873-x>
- Jaspers-Fayer, F., Maffei, A., Goertzen, J., Kleffner, K., Coccaro, A., Sessa, P., & Liotti, M. (2022). Spatiotemporal dynamics of covert vs. overt emotional face processing in dysphoria. *Frontiers in Behavioral Neuroscience*, *16*, Article 920989. <https://doi.org/10.3389/fnbeh.2022.920989>
- Karaaslan, A., Durmuş, B., & Amado, S. (2020). Does body context affect facial emotion perception and eliminate emotional ambiguity without visual awareness? *Visual Cognition*, *28*(10), 605-620. <https://doi.org/10.1080/13506285.2020.1846649>
- Kret, M. E., & de Gelder, B. (2010). Social context influences recognition of bodily expressions. *Experimental Brain Research*, *203*(1), 169-180. <https://doi.org/10.1007/s00221-010-2220-8>

- Kret, M. E., Roelofs, K., Stekelenburg, J. J., & de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience*, 7, Article 810.
<https://doi.org/10.3389/fnhum.2013.00810>
- Kret, M. E., Stekelenburg, J. J., Roelofs, K., & de Gelder, B. (2013). Perception of face and body expressions using electromyography, pupillometry and gaze measures. *Frontiers in Psychology*, 4, Article 28. <https://doi.org/10.3389/fpsyg.2013.00028>
- Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, 4(1).
<https://doi.org/10.1177/2515245920951503>
- Lecker, M., Dotsch, R., Bijlstra, G., & Aviezer, H. (2020). Bidirectional contextual influence between faces and bodies in emotion perception. *Emotion*, 20(7), 1154-1164.
<https://doi.org/10.1037/emo0000619>
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Leys, C., Delacre, M., Mora, Y. L., Lakens, D., & Ley, C. (2019). How to classify, detect, and manage univariate and multivariate outliers, with emphasis on pre-registration. *International Review of Social Psychology*, 32(1). <https://doi.org/10.5334/irsp.289>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764-766.
<https://doi.org/10.1016/j.jesp.2013.03.013>
- Li, X. (2021). Recognition characteristics of facial and bodily expressions: Evidence from ERPs. *Frontiers in Psychology*, 12, Article 680959. <https://doi.org/10.3389/fpsyg.2021.680959>

- Lima, C. F., Anikin, A., Monteiro, A. C., Scott, S. K., & Castro, S. L. (2019). Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion, 19*(2), 219-233.
<https://doi.org/10.1037/emo0000429>
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*, Article 213.
<https://doi.org/10.3389/fnhum.2014.00213>
- McCrackin, S. D., & Itier, R. J. (2021). Feeling through another's eyes: Perceived gaze direction impacts ERP and behavioural measures of positive and negative affective empathy. *Neuroimage, 226*, Article 117605. <https://doi.org/10.1016/j.neuroimage.2020.117605>
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences, 102*(45), 16518-16523. <https://doi.org/10.1073/pnas.0507650102>
- Mondloch, C. J., Nelson, N. L., & Horner, M. (2013). Asymmetries of influence: Differential effects of body postures on perceptions of emotional facial expressions. *PLoS One, 8*(9), Article e73605. <https://doi.org/10.1371/journal.pone.0073605>
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297-326. <https://doi.org/10.1037/0033-2909.132.2.297>
- Palmer, J. A., Kreutz-Delgado, K., & Makeig, S. (2012). AMICA: An adaptive mixture of independent component analyzers with shared components.
<https://www.bing.com/ck/a?!&&p=8840a577b0cf5960efad0914b7629377c6b1446468e5ef5128083a778e73b064JmltdHM9MTc1NzI4OTYwMA&ptn=3&ver=2&hsh=4&fclid=01b0ca0b-33c3-629b-3260-de7a32f863a5&u=a1aHR0cHM6Ly9zY2NuLnVjc2QuZWR1L35qYXNvbi9hbWljYV9hLnBkZg&ntb=1>

- Pourtois, G., Debatisse, D., Despland, P. A., & de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Cognitive Brain Research, 14*(1), 99-105. [https://doi.org/10.1016/S0926-6410\(02\)00064-2](https://doi.org/10.1016/S0926-6410(02)00064-2)
- Puce, A., McNeely, M. E., Berrebi, M. E., Thompson, J. C., Hardee, J., & Brefczynski-Lewis, J. (2013). Multiple faces elicit augmented neural activity. *Frontiers in Human Neuroscience, 7*, Article 282. <https://doi.org/10.3389/fnhum.2013.00282>
- Puffet, A. S., & Rigoulot, S. (2025a). The role of cognitive load in automatic integration of emotional information from face and body. *Scientific Reports, 15*(1), Article 28184. <https://doi.org/10.1038/s41598-025-12511-8>
- Puffet, A. S., & Rigoulot, S. (2025b). Validation of the Emotionally Congruent and Incongruent Face-Body Static Set (ECIFBSS). *Behavior Research Methods, 57*(1), Article 41. <https://doi.org/10.3758/s13428-024-02550-w>
- Puffet, A.-S., Saleh, G., & Rigoulot, S. (2025c). Exploring the impact of facial-body emotion incongruence: Insight from a systematic review. *Journal of Nonverbal Behavior*. <https://doi.org/10.1007/s10919-025-00488-x>
- Russell, A., Buerkner, P., Herve, M., Love, J., & Singmann, H. (2021). *Package 'emmeans'. R package version*.
- Seubert, J., Kellermann, T., Loughhead, J., Boers, F., Brensinger, C., Schneider, F., & Habel, U. (2010). Processing of disgusted faces is facilitated by odor primes: A functional MRI study. *Neuroimage, 53*(2), 746-756. <https://doi.org/10.1016/j.neuroimage.2010.07.012>
- Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: An ERP study on the body-inversion effect. *Neuroreport, 15*(5), 777-780. <https://doi.org/10.1097/01.wnr.0000119730.93564.e8>

- Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion, 8*(1), 81-95. <https://doi.org/10.1037/1528-3542.8.1.81>
- van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E.-J., Derks, K., Dablander, F., Gronau, Q. F., Kucharský, Š., Gupta, A. R. K. N., Sarafoglou, A., Voelkel, J. G., Stefan, A., Ly, A., Hinne, M., Matzke, D., & Wagenmakers, E.-J. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *L'Année psychologique, 120*(1), 73-96. <https://doi.org/10.3917/anpsy1.201.0073>
- van den Bergh, D., Wagenmakers, E.-J., & Aust, F. (2023). Bayesian repeated-measures analysis of variance: An updated methodology implemented in JASP. *Advances in Methods and Practices in Psychological Science, 6*(2). <https://doi.org/10.1177/25152459231168024>
- van Dillen, L. F., & Derks, B. (2012). Working memory load reduces facilitated processing of threatening faces: An ERP study. *Emotion, 12*(6), 1340-1349. <https://doi.org/10.1037/a0028624>
- van Doorn, J., van den Bergh, D., Bohm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharsky, S., Ly, A., Marsman, M., Matzke, D., Gupta, A., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E. J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review, 28*(3), 813-826. <https://doi.org/10.3758/s13423-020-01798-5>
- Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior, 17*, 3-28. <https://doi.org/10.1007/BF00987006>
- Wang, L., Xia, L., & Zhang, D. (2017). Face-body integration of intense emotional expressions of victory and defeat. *PLoS One, 12*(2), Article e0171656. <https://doi.org/10.1371/journal.pone.0171656>

- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. <https://ggplot2.tidyverse.org>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), Article 1686. <https://doi.org/10.21105/joss.01686>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). dplyr: A grammar of data manipulation. <https://github.com/tidyverse/dplyr>
- Willis, M. L., Palermo, R., & Burke, D. (2011). Judging approachability on the face of it: The influence of face and body expressions on the perception of approachability. *Emotion*, 11(3), 514-523. <https://doi.org/10.1037/a0022571>
- Yang, Y. F., Brunet-Gouet, E., Burca, M., Kalunga, E. K., & Amorim, M. A. (2020). Brain processes while struggling with evidence accumulation during facial emotion recognition: An ERP study. *Frontiers in Human Neuroscience*, 14, Article 340. <https://doi.org/10.3389/fnhum.2020.00340>
- Zhang, D., Zhao, T., Liu, Y., & Chen, Y. (2015). Comparison of facial expressions and body expressions: An event-related potential study. *Acta Psychologica Sinica*, 47(8), 963-970. <https://doi.org/10.3724/sp.J.1041.2015.00963>
- Zhang, M., Liu, T., Jin, Y., He, W., Huang, Y., & Luo, W. (2019). The asynchronous influence of facial expressions on bodily expressions. *Acta Psychologica*, 200, Article 102941. <https://doi.org/10.1016/j.actpsy.2019.102941>
- Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67(6), 361-370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>

Discussion générale

Les émotions occupent une place centrale dans notre quotidien, intervenant notamment dans l'établissement et le maintien de relations sociales de qualité, indispensables à notre bien-être. Pour mieux comprendre comment nous percevons et traitons les informations émotionnelles, la recherche a exploré la manière dont les expressions faciales sont intégrées dans un ensemble d'indices émotionnels, telles que les postures corporelles (Kret, Stekelenburg et al., 2013; Lecker et al., 2020; Meeren et al., 2005; Mondloch et al., 2013). Autrement dit, la perception des émotions ne repose pas uniquement sur un signal isolé, mais résulte de l'intégration dynamique de multiples indices disponibles dans l'environnement. Toutefois, ce champ reste confronté à plusieurs limites importantes : (1) la majorité des études utilisent des stimuli peu écologiques, souvent issus de banques d'images distinctes pour les visages et les corps, ce qui nuit à la validité écologique des résultats; (2) il existe encore peu de connaissances sur le caractère automatique ou volontaire du traitement de cet effet contextuel entre les expressions faciales et corporelles; (3) enfin, la plupart des recherches sur l'automatisme de l'effet contextuel adoptent une approche uniquement comportementale, offrant peu d'éléments sur les mécanismes neurocognitifs impliqués. Face à ces constats, ce travail de thèse avait pour objectif principal de mieux caractériser l'effet contextuel entre les expressions faciales et corporelles, en se focalisant sur l'étude de son automatisme. Pour cela, nous avons combiné une approche comportementale et électrophysiologique (composantes

PE), et avons développé un ensemble de stimuli de haute qualité permettant de se rapprocher d'une perception émotionnelle en contexte réel.

Dans un premier temps, l'article 1 valide la création d'un matériel composé de stimuli visuels émotionnels congruents et incongruents combinant des expressions faciales et corporelles, comblant ainsi les lacunes méthodologiques des études antérieures. Ensuite, deux études expérimentales (articles 2 et 3) ont permis de tester empiriquement l'effet contextuel en manipulant deux critères d'automaticité : la charge cognitive (article 2) et la rapidité de traitement (article 3). Les résultats de ces deux dernières études convergent en montrant un effet contextuel avec une reconnaissance émotionnelle plus rapide et plus précise en condition congruente qu'incongruente. De plus, les postures corporelles semblent exercer une influence plus marquée sur la reconnaissance des expressions faciales qu'inversement. Enfin, sur les plans comportemental et neural, ni la charge cognitive (article 2), ni la rapidité de traitement (article 3) n'interagissent significativement avec le facteur de congruence, suggérant un traitement automatique de cette intégration contextuelle selon ces deux critères. Néanmoins, certaines divergences entre les deux études émergent. Dans l'article 2, aucun effet de la congruence n'est observé tant précocement que tardivement au niveau de l'activité cérébrale. En revanche, deux stades de traitement émotionnel distincts émergent selon le focus attentionnel : une activité cérébrale précoce (autour de 100ms) plus importante lorsque l'attention est dirigée vers le visage, tandis qu'une activité soutenue plus tardive (> 400 ms) est observée lorsque l'attention est dirigée vers le corps. A l'inverse, dans l'article 3, l'activité cérébrale est

significativement plus importante pour les stimuli congruents plutôt qu'incongruents, aussi bien aux étapes précoces que tardives du traitement, et ce, quel que soit le focus attentionnel. De plus, les expressions faciales induisent une activité plus importante que les expressions corporelles tant précocement que tardivement dans le traitement émotionnel. Ces résultats seront discutés dans la suite de la discussion.

Création d'une banque de stimuli pour une approche plus écologiquement valide

Le premier apport de ce travail de thèse est la création et la validation d'un ensemble de stimuli d'expressions faciales et corporelles en situation de congruence et d'incongruence (Puffet & Rigoulot, 2025a). Cette banque de stimuli, nommée ECIFBSS, regroupe 1 952 images d'émotions faciales et corporelles représentées à la fois en condition congruente et incongruente. Pour ce faire, 40 acteurs ont exprimé sept émotions (colère, dégoût, peur, surprise, tristesse, joie et neutre) avec leur visage et leur corps à la fois en situation de congruence et d'incongruence, générant des combinaisons faciales-corporelles à partir des mêmes individus, garantissant une homogénéité. Cette approche constitue une avancée méthodologique importante, en réponse aux limites des études antérieures, dans lesquelles les expressions faciales et corporelles étaient souvent issues de banques d'images distinctes, puis fusionnées artificiellement pour créer des combinaisons congruentes ou incongruentes (Karaaslan et al., 2020; Kret, Stekelenburg et al., 2013; Lecker et al., 2020; Mondloch et al., 2013). Ce type de procédure introduit des biais perceptifs importants, en raison de différences de contraste, de luminosité ou encore de qualité visuelle entre les deux sources d'images. Or, plusieurs études ont montré

que ces paramètres influencent directement la reconnaissance émotionnelle au niveau comportemental (Wallbott, 1992), ainsi que les réponses cérébrales précoces, notamment la composante P100, particulièrement sensible aux caractéristiques visuelles des stimuli (couleur, position du champ visuel, contraste, fréquences spatiales, etc.) (Gu et al., 2013). Dans cette perspective, l'utilisation de stimuli uniformes, standardisés et de haute qualité apparaît indispensable pour garantir la validité des résultats, en limitant les biais perceptifs susceptibles de fausser l'interprétation des effets observés.

Par ailleurs, la validation de la banque ECIFBSS a été réalisée par deux études complémentaires permettant de recueillir des données portant sur les taux de reconnaissance émotionnelle, mais également sur des dimensions affectives telles que l'activation, la valence, l'intensité et l'authenticité des expressions faciales et corporelles. Les résultats issus de ces études de validation sont en adéquation avec les travaux précédents de création de stimuli (de Gelder & van den Stock, 2011; Goeleven et al., 2008; Thoma et al., 2013; Tottenham et al., 2009; van der Schalk et al., 2011). Ils montrent notamment que les émotions faciales sont globalement mieux reconnues que les expressions corporelles. Plus précisément, l'émotion faciale de joie apparaît comme l'émotion la mieux reconnue, tandis que la peur est celle la moins bien reconnue. Du côté des expressions corporelles, la colère est la mieux reconnue, alors que le dégoût est identifié comme ayant le plus faible taux de reconnaissance (de Gelder & van den Stock, 2011; Goeleven et al., 2008; Thoma et al., 2013; Tottenham et al., 2009; van der Schalk et al., 2011). Ces résultats peuvent s'expliquer par la fréquence d'exposition et la

variabilité dans l'expression des émotions concernées. En effet, la joie est une émotion fréquemment rencontrée dans la vie quotidienne, facilitant sa reconnaissance (Calvo et al., 2014), tandis que la colère, notamment lorsqu'elle est exprimée par le corps, est souvent représentée par une configuration stéréotypée (par exemple, poings serrés et posture tendue), facilitant sa reconnaissance. A l'inverse, la peur ou le dégoût sont exprimés de manière plus variable et sont moins fréquemment rencontrés dans le quotidien, rendant leur identification plus difficile (van den Stock et al., 2007). En outre, à l'instar d'autres banques de stimuli émotionnels (Dawel et al., 2017; Ong et al., 2023), l'évaluation de l'authenticité suggère que les expressions faciales et corporelles sont perçues comme modérément authentiques. L'authenticité est un facteur très peu pris en considération dans les banques de stimuli émotionnels. En effet, aucune des banques majoritairement utilisées dans les études sur les effets contextuels ne fournit ce type de données (de Gelder & van den Stock, 2011; Lopez et al., 2017; Thoma et al., 2013). Pourtant, l'authenticité d'une expression émotionnelle a une implication sur la reconnaissance de l'émotion exprimée (Dawel et al., 2017). C'est pourquoi, il semble important que ce critère soit plus systématiquement inclus dans les banques de stimuli émotionnels. Finalement, les évaluations de la valence et de l'activation s'alignent sur le modèle dimensionnel des émotions (Russell, 1980), renforçant ainsi la validité théorique des stimuli proposés. Ces données offrent à la communauté scientifique un outil expérimental rigoureux et standardisé, permettant un contrôle de variables émotionnelles susceptibles d'influencer la reconnaissance des émotions. Par conséquent, la banque ECIFBSS représente une

ressource innovante et essentielle pour la conception d'études futures plus écologiques et méthodologiquement robustes.

Automaticité de l'intégration des expressions faciales et corporelles

Les articles 2 et 3 ont porté spécifiquement sur l'exploration de deux critères d'automaticité tels que définis par Moors et De Houwer (2006) : la charge cognitive (article 2) et la rapidité de traitement (article 3) (Moors & De Houwer, 2006). Ces critères avaient déjà été manipulés dans des travaux antérieurs. Concernant la charge cognitive, Aviezer et ses collègues (2011) ont mené une étude comportementale dans laquelle chaque essai débutait par une tâche de mémorisation d'une séquence complexe de cinq chiffres et une lettre (p. ex., 183K65) dans une condition de haute charge cognitive ou uniquement une lettre dans une condition de faible charge cognitive. Après avoir perçu cette série de chiffres, les participants devaient identifier l'émotion exprimée par le visage dans des images combinant expressions faciales et postures corporelles congruentes ou incongruentes. Finalement, les participants retranscrivaient la série de chiffres et lettres qu'ils avaient perçus au début de l'essai (Aviezer et al., 2011). Les résultats montraient que le niveau de charge cognitive n'altérait pas l'influence des expressions corporelles sur la reconnaissance des émotions faciales, suggérant ainsi un traitement automatique au regard de ce critère. De manière similaire, le critère de rapidité a été investigué dans une étude combinant des mesures comportementales et EEG (Meeren et al., 2005). Les participants percevaient rapidement (200ms) des expressions faciales et corporelles congruentes ou incongruentes et identifiaient l'émotion faciale. Les résultats indiquaient

que la rapidité de présentation n'atténuait pas l'effet contextuel, et que des composantes précoces comme la P100 montraient une amplitude plus élevée pour les stimuli émotionnellement incongruents plutôt que congruents, soulignant une détection rapide et précoce du conflit émotionnel (Meeren et al., 2005). A notre connaissance, il n'existe aucune autre étude ayant manipulé ces deux critères dans l'effet contextuel entre les expressions faciales et corporelles n'offrant donc qu'une vision partielle de l'automatisme de ce traitement.

Charge cognitive

Les résultats de l'article 2 renforcent les observations comportementales précédentes (Aviezer et al., 2011), en soutenant l'hypothèse selon laquelle l'effet contextuel entre les expressions faciales et corporelles opère de manière automatique au regard du critère de charge cognitive. En effet, aucune différence significative n'a été observée entre les conditions de haute et de basse charge cognitive en fonction de la congruence, que ce soit en termes de taux de reconnaissance ou de temps de réponse. Ces résultats suggèrent que l'intégration des informations faciales et corporelles ne requiert pas de ressources cognitives particulières. L'article 2 apporte également un éclairage nouveau à travers des données de composantes PE. Au niveau neurophysiologique, les données vont dans le même sens que les données comportementales, c'est-à-dire que le niveau de charge cognitive n'a pas modifié les effets d'intégration entre les indices faciaux et corporels. Toutefois, cette absence d'effet doit être interprétée avec prudence. En effet, aucune différence significative n'a été détectée entre les conditions congruentes et incongruentes

au niveau de l'activité cérébrale, ce qui pourrait indiquer un traitement similaire de ces stimuli. Une explication possible réside dans la conception du protocole expérimental, qui ne comportait pas de condition sans charge cognitive. Il est donc envisageable que la seule présence d'une tâche de mémorisation, qu'elle soit exigeante ou non, ait altéré le traitement des signaux émotionnels. Cette hypothèse est appuyée par des travaux antérieurs montrant, dans une étude combinant expressions faciales et scènes émotionnelles, que des effets de congruence au niveau de la composante N170 n'émergeaient que dans la condition sans charge cognitive, et disparaissaient dès qu'une charge cognitive, même minimale, était introduite (Cao et al., 2022). Ainsi, il est possible que ce ne soit pas l'intensité de la charge cognitive qui module l'effet contextuel, mais le fait même d'exécuter une double tâche. Dans cette perspective, l'absence de condition contrôle (sans charge cognitive) dans notre étude constitue une limite méthodologique. Des recherches futures devraient inclure une telle condition afin de déterminer si un effet contextuel est observé uniquement en l'absence de toute charge cognitive. Si tel est le cas, cela remettrait en question l'idée que l'effet contextuel est entièrement automatique, en suggérant qu'il dépend au contraire de la disponibilité de ressources attentionnelles minimales.

Rapidité de traitement

Dans l'étude 3, le critère de rapidité de traitement a été manipulé par une présentation brève des stimuli (200ms), comme dans les travaux de Meeren et ses collègues (2005), ainsi que par la variation du délai de réponse (rapide ou lent). Les participants devaient,

selon la condition, répondre rapidement ou non, ce qui permettait de contrôler le temps de traitement des informations émotionnelles après la disparition du stimulus. Les résultats comportementaux ne montrent aucune interaction significative entre la rapidité de réponse et la congruence, ni sur les taux de reconnaissance ni sur les temps de réponse. Ces observations suggèrent que l'intégration des expressions faciales et corporelles se déroule de façon similaire, que les participants disposent d'un temps de réponse limité ou étendu, renforçant ainsi les résultats de la littérature antérieure (Meeren et al., 2005). Sur le plan neurophysiologique, aucune interaction claire entre la rapidité de traitement et la congruence n'a été observée. Un effet limité a toutefois été relevé entre 110 et 310ms dans certains sites (fronto-centraux et centro-parieto-occipitaux), indiquant une activité cérébrale accrue pour les stimuli congruents plutôt qu'incongruents uniquement lors de la condition lente et non lors de la condition rapide. Ces résultats suggèrent que l'intégration contextuelle entre les expressions faciales et corporelles est plus marquée lorsque les participants disposent de davantage de temps, ce qui nuance l'hypothèse d'une automaticité du traitement. En effet, si l'effet contextuel n'apparaît que lorsque le temps n'est pas limité, cela pourrait indiquer que l'influence des indices faciaux et corporels nécessite un laps de temps minimum pour se manifester lors de la tâche de reconnaissance émotionnelle. Dans ce cas, cet effet ne répondrait pas au critère de rapidité de traitement défini par Moors et De Houwer (2006) comme caractéristique de l'automaticité d'un traitement émotionnel. Toutefois, cet effet reste difficile à interpréter, car il se limite à deux électrodes (F6 et AF4), témoignant d'un effet faible et spatialement restreint qui doit être considéré avec prudence. Par ailleurs, l'étude 3 présente une limite méthodologique

susceptible d'avoir influencé les résultats. En effet, la rapidité de présentation des stimuli n'a pas été manipulée. Ceux-ci ont été présentés uniquement de manière brève (200ms), ce qui ne permet pas de déterminer si des différences seraient apparues avec un temps d'exposition plus long, comme dans l'étude 2 (1500ms). Cette contrainte réduit également la possibilité de comparer directement les résultats des études 2 et 3.

Intégration des deux articles expérimentaux

Les deux dernières études de cette thèse (articles 2 et 3) relèvent certaines divergences dans les résultats obtenus, en particulier au niveau de l'activité neurale. Le premier point de divergence concerne l'effet contextuel. Alors que l'effet contextuel entre les expressions faciales et corporelles est clairement observé dans les mesures comportementales des deux études, il n'apparaît pas au niveau de l'activité cérébrale dans l'article 2, contrairement à l'article 3. Plus précisément, alors qu'aucune différence significative n'est trouvée entre les stimuli congruents et incongruents au niveau de l'activité cérébrale dans l'article 2, les résultats de l'article 3 mettent en évidence une activité cérébrale accrue pour les stimuli congruents, tant à des activités cérébrales précoces que tardives, suggérant une priorisation du traitement des informations émotionnelles consistantes par rapport à des informations incongruentes (Balconi & Carrera, 2011; Pourtois et al., 2002). Cette divergence pourrait être expliquée par le fait que dans l'article 3, aucune manipulation de la charge cognitive n'a été effectuée, renforçant ainsi l'idée que les deux niveaux de charge cognitive testés dans l'article 2 ne suffisent pas à eux seuls à établir le caractère automatique de l'intégration émotionnelle.

Afin d'étayer de manière robuste l'hypothèse d'un traitement automatique de l'effet contextuel, il serait pertinent de conduire une nouvelle étude explorant le dernier critère d'automaticité, la conscience. Cela pourrait être réalisé à travers un protocole comparant des conditions de perception consciente et non consciente (Karaaslan et al., 2020). Une telle recherche permettrait de déterminer si l'intégration contextuelle des expressions faciales et corporelles dépend d'un traitement conscient des stimuli, et de tester indirectement l'hypothèse selon laquelle l'absence de manipulation de la charge cognitive, même faible, pourrait favoriser l'émergence de cet effet au niveau cérébral.

Les deux études montrent également des résultats similaires concernant le focus attentionnel. Dans les deux études expérimentales (articles 2 et 3), nous avons manipulé le focus attentionnel afin d'examiner l'influence des instructions sur l'effet contextuel. Concrètement, les participants devaient, selon les essais, focaliser leur attention sur le visage ou le corps, et identifier l'émotion exprimée par le canal ciblé. Ce protocole permettrait d'évaluer la bidirectionnalité de l'effet contextuel, c'est-à-dire la capacité des expressions corporelles à influencer la reconnaissance des émotions faciales, mais aussi l'effet inverse (Kret, Stekelenburg et al., 2013; Lecker et al., 2020). Les résultats des deux études révèlent des tendances similaires. D'une part, l'effet contextuel semble être bidirectionnel. En effet, les indices faciaux et corporels influencent réciproquement la reconnaissance émotionnelle, ce qui concorde avec la littérature actuelle (Kret, Stekelenburg et al., 2013; Lecker et al., 2020). D'autre part, une asymétrie dans cette influence émerge de manière systématique : les postures corporelles semblent exercer un

effet plus fort sur la reconnaissance des expressions faciales qu'inversement. Ce déséquilibre renforce l'idée d'une dominance des postures corporelles dans le traitement émotionnel (Lecker et al., 2020). Cette bidirectionnalité asymétrique, observée dans nos études et introduite par Lecker et ses collègues (2020), renforce l'hypothèse selon laquelle les émotions perçues à travers le corps et le visage ne sont pas traitées de manière séparée et équivalente, mais plutôt de manière holistique (Aviezer et al., 2012a) et hiérarchisée. Ainsi, les deux canaux véhiculent une information émotionnelle pertinente, mais les expressions corporelles semblent être associées à un traitement émotionnel facilité influençant ainsi plus fortement la reconnaissance des émotions faciales. Plusieurs hypothèses peuvent être avancées pour expliquer cette asymétrie. D'une part, les postures corporelles sont plus saillantes visuellement et offrent des signaux plus facilement détectables, favorisant ainsi leur traitement précoce et leur impact sur les visages. D'autre part, cette asymétrie pourrait s'expliquer par des différences d'intensité ou d'activation émotionnelle entre les deux indices. En effet, dans les deux études, les expressions corporelles ont été évaluées comme plus intenses et plus activatrices que les expressions faciales. Or, il a été démontré que ces deux dimensions affectives facilitent la reconnaissance émotionnelle dans un contexte multimodal (Morningstar et al., 2021), ce qui pourrait expliquer la prédominance des postures corporelles dans l'intégration émotionnelle.

Cependant, au niveau de l'activité cérébrale, les données sur le focus attentionnel des études 2 et 3 divergent. Notamment, dans l'étude 2, l'activité cérébrale précoce (P100,

N100, P250, N250) était plus importante lorsque l'attention était focalisée sur les expressions faciales plutôt que corporelles (Meeren et al., 2005; Stekelenburg & de Gelder, 2004). Plus tardivement, à partir de 400ms, une attention plus soutenue était associée à la focalisation de l'attention sur la posture corporelle. Cela montrait un traitement préférentiel précoce des expressions faciales et un besoin plus important de ressources attentionnelles tardivement sur les postures corporelles. Dans l'étude 3, aussi bien précocement que tardivement, l'activité cérébrale était plus importante lorsque l'attention était focalisée sur les expressions faciales plutôt que corporelles (Meeren et al., 2005; Stekelenburg & de Gelder, 2004; Zhang et al., 2015), suggérant également un traitement précoce préférentiel des expressions faciales, mais qui semble se prolonger tardivement dans le traitement. Cette différence au niveau du traitement tardif entre les études 2 et 3 pourrait probablement s'expliquer par le devis expérimental et particulièrement la présentation rapide des stimuli émotionnels dans l'étude 3. En effet, les stimuli émotionnels étaient présentés durant 200ms (vs. 1500ms dans l'étude 2) limitant ainsi le temps de traitement disponible pour réaliser la tâche de reconnaissance émotionnelle et potentiellement diminuant la possibilité de focaliser de manière soutenue l'attention sur les postures corporelles plus tardivement. Par conséquent, si une présentation courte des stimuli émotionnels influence la focalisation de l'attention tardivement sur les indices corporels, cela pourrait indiquer une modulation du temps de présentation sur le traitement des indices faciaux et corporels. Afin de déterminer si le temps de présentation des stimuli influence le traitement des indices faciaux et corporels,

une nouvelle étude pourrait inclure l'ensemble des conditions expérimentales, en manipulant à la fois le temps de réponse et le temps de présentation des stimuli.

Proposition d'un modèle intégratif

Les deux études menées, portant respectivement sur la manipulation de la charge cognitive et la rapidité de traitement, font état d'un traitement similaire des indices faciaux et corporels au niveau comportemental. Sur la base de ces résultats, un modèle intégratif peut être proposé, suggérant que la perception simultanée des expressions faciales et corporelles repose sur un traitement automatique selon deux critères de Moors et De Houwer (2006; la charge cognitive et la rapidité de traitement) (Aviezer et al., 2011; Meeren et al., 2005). Plus spécifiquement, l'influence croisée entre les expressions faciales et corporelles, c'est-à-dire l'impact des expressions corporelles sur la reconnaissance des expressions faciales et inversement, se produit automatiquement, sans être modulée, ni par le niveau de charge cognitive, ni par le temps imparti à la tâche de reconnaissance. Ainsi, les résultats comportementaux des deux études convergent pour soutenir l'idée d'un processus d'intégration émotionnelle automatique entre les visages et les corps (Aviezer et al., 2011; Meeren et al., 2005). Un tel fonctionnement automatique aurait un intérêt adaptatif majeur dans la vie quotidienne. En effet, le traitement rapide et efficace des indices émotionnels issus de sources multiples permet une lecture contextuelle fine, favorisant des réactions adaptées face aux situations sociales ou potentiellement menaçantes. Par exemple, percevoir immédiatement l'intention hostile d'un individu s'approchant dans une ruelle sombre grâce à l'intégration automatique de

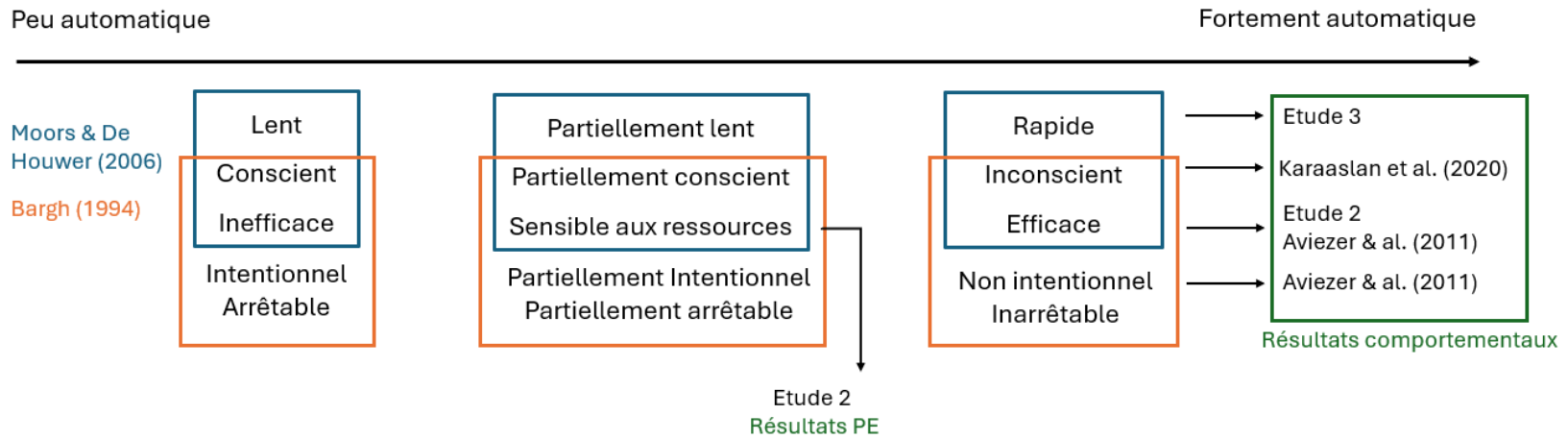
ses expressions faciales et corporelles permettrait une réaction rapide et efficace basée sur une analyse intégrée et donc plus précise de la situation. En somme, ce travail de thèse apporte des observations issues des études 2 et 3 qui soutiennent l'existence d'un modèle théorique d'intégration automatique des indices faciaux et corporels, répondant aux critères de rapidité et d'efficacité proposés par Moors et De Houwer (2006), et traduisant une exigence fondamentale de traitement efficace et rapide des signaux émotionnels de l'environnement. Cependant, ce modèle serait enrichi par la considération d'autres critères notamment le critère de conscience également proposé par Moors et De Houwer ou encore en prenant en compte d'autres modèles comme celui de Bargh et ses collègues (1996). Le critère de conscience n'a pas été manipulé lors de ce travail de thèse, mais une précédente étude comportementale a déjà exploré ce critère d'automatisme. Dans cette étude, les participants percevaient d'abord une combinaison d'expressions faciales et corporelles présentée durant 33ms, puis uniquement l'expression faciale pendant 150ms, avant d'effectuer une tâche de reconnaissance de cette expression faciale (Karaaslan et al., 2020). Les résultats ont montré que l'intégration des indices faciaux et corporels se réalisait même lorsque les participants n'étaient pas conscients d'avoir perçu simultanément les expressions faciales et corporelles (présentation durant 33ms), supportant ainsi le modèle d'automatisme suggéré par ce travail de thèse (Karaaslan et al., 2020).

Il est important de souligner que, selon le modèle de Moors et De Houwer (2006), l'automatisme n'est pas nécessairement une caractéristique dichotomique d'un traitement

cognitif. Chaque critère est indépendant, et un traitement peut être automatique avec des degrés variables selon différents axes (Moors & De Houwer, 2006). Autrement dit, l'automatisme se place sur un continuum plutôt qu'en catégories strictes. Ce positionnement s'aligne également avec le modèle de Bargh et ses collègues (1996), qui introduit quant à lui d'autres critères, notamment l'absence d'intention préalable et l'inarrêtabilité du traitement (Bargh, 2014; Bargh & Tota, 1988; Bargh et al., 1996). Le critère de non-intention, proposé par Bargh et ses collègues, a été précédemment exploré dans le cas de l'effet contextuel entre des indices faciaux et corporels (Aviezer et al., 2011). Dans cette étude, les participants étaient répartis en trois groupes : le premier avait pour consigne explicite d'ignorer le corps, avec information sur l'influence potentielle des postures corporelles; le second groupe croyait que les postures corporelles étaient aléatoires et sans importance pour la tâche; le dernier groupe réalisait la tâche de reconnaissance sans instruction particulière. Malgré ces consignes, les résultats montrent que les participants n'étaient pas capables d'ignorer intentionnellement le corps, et aucune différence significative entre les groupes n'a été observée (Aviezer et al., 2011). Ainsi, à la lumière des études précédentes et des résultats présentés dans ce travail de thèse, l'effet contextuel entre les expressions faciales et corporelles correspond, au niveau comportemental, à plusieurs critères d'automatisme, positionnant ce traitement cognitif comme fortement automatique sur le continuum (voir Figure 12).

Figure 12

Schéma représentant le continuum d'automatisme dans lequel s'inscrivent nos résultats comportementaux et électrophysiologiques



Cependant, au niveau de l'activité cérébrale, l'effet contextuel ne semble pas être positionné de manière similaire sur le continuum. En effet, l'absence d'effet contextuel dans l'étude 2 (manipulant la charge cognitive), contrastant avec sa présence lors du traitement émotionnel précoce dans l'étude 3 (manipulant la rapidité de traitement), suggère que l'intégration ne serait pas entièrement automatique au niveau des mécanismes cérébraux sous-jacents. Il est ainsi possible que certains processus neuronaux impliqués dans cette intégration requièrent des ressources cognitives pour s'activer de manière efficace, suggérant un traitement non pas entièrement automatique, mais plutôt partiellement automatique (voir Figure 12). Autrement dit, l'intégration des expressions faciales et corporelles semblerait se dérouler de manière automatique sur le plan comportemental, mais impliquerait, au niveau cérébral, des mécanismes plus contrôlés pouvant être influencés par la charge cognitive (Cao et al., 2022; Puffet & Rigoulot, 2025b).

Nos travaux suggèrent donc que la charge cognitive pourrait potentiellement interférer avec l'intégration des expressions faciales et corporelles rendant le traitement émotionnel plus difficile au niveau cérébral (Cao et al., 2022; Puffet & Rigoulot, 2025b). Cette hypothèse coïncide avec de nombreux modèles de traitement émotionnel suggérant l'implication d'un réseau dynamique constitué d'une voie automatique et une seconde inférentielle (Pessoa, Japee et al., 2005; Pessoa, Padmala et al., 2005; Vuilleumier, 2002). C'est notamment le cas du modèle des émotions en tant qu'information sociale (*Emotions as Social Information Model, EASI*) postulant de l'existence de deux voies distinctes par

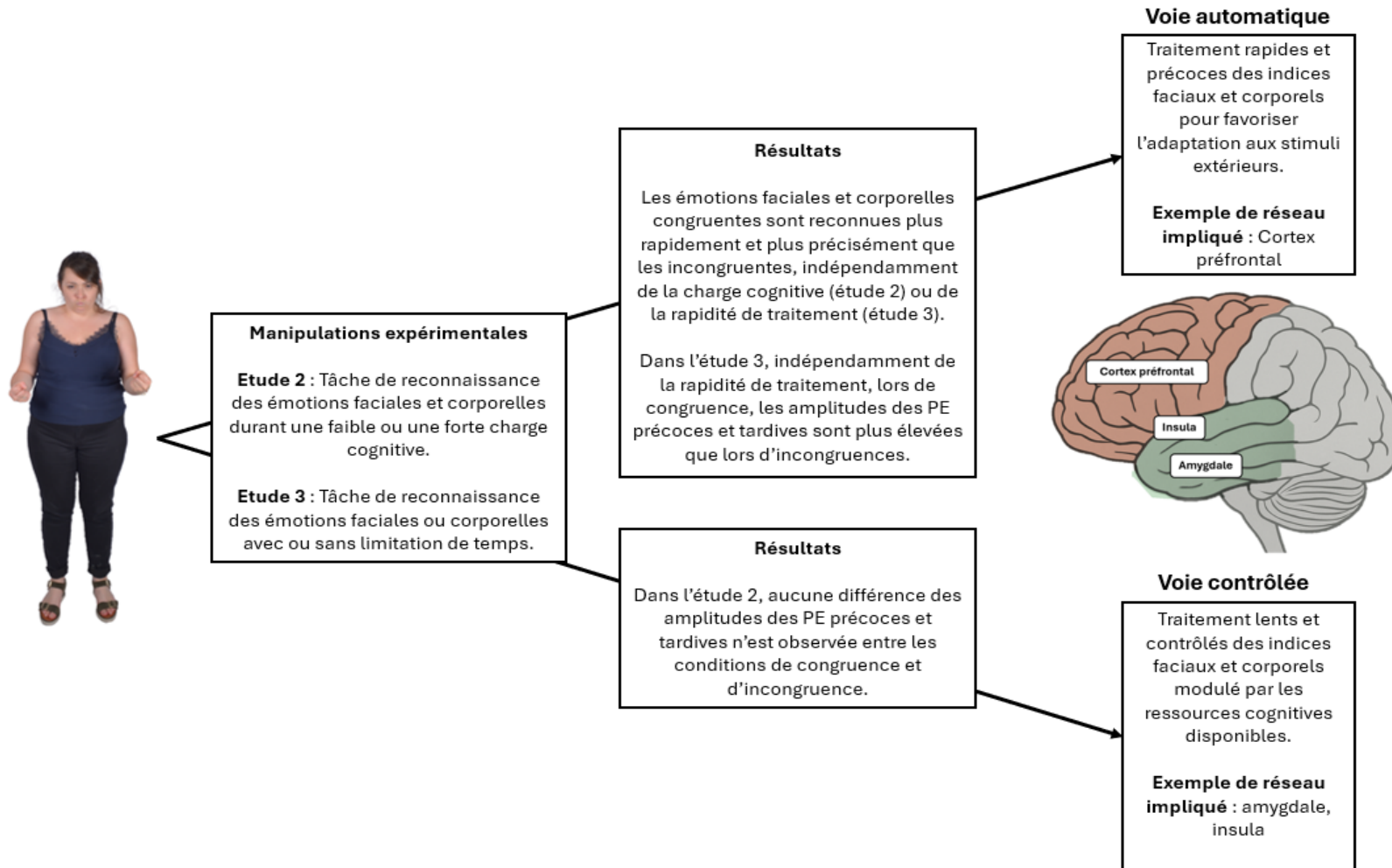
lesquelles les émotions d'autrui influencent notre comportement : une voie affective, rapide et automatique, et une voie inférentielle, plus lente et contrôlée (van Kleef, 2009). La voie affective est mobilisée lors de la perception d'expressions émotionnelles déclenchant une réponse émotionnelle immédiate, comme dans les phénomènes de contagion émotionnelle, c'est-à-dire le traitement par lequel l'émotion perçue chez autrui induit une émotion similaire chez soi. A l'inverse, la voie inférentielle intervient lorsque nous cherchons à interpréter les états internes d'autrui, ses intentions ou ses motivations, par exemple, en inférant les causes d'une frustration exprimée par un proche. Cette seconde voie mobilise davantage de ressources cognitives, et son activation dépend du contexte attentionnel et motivationnel (van Kleef, 2009). L'intégration des indices faciaux et corporels pourrait donc suivre une dynamique similaire à celle proposée par le modèle EASI. Ainsi, certains aspects de l'intégration, notamment ceux observés au niveau comportemental (études 2 et 3) ou cérébral lorsque le traitement est rapide (étude 3), suggèrent un traitement automatique, analogue à la voie affective. En revanche, lorsque les ressources cognitives disponibles sont plus faibles (étude 2), l'intégration des indices émotionnels faciaux et corporels, au niveau cérébral, n'est pas purement automatique, mais modulée par les ressources cognitives disponibles, comme le propose le modèle EASI pour le traitement émotionnel en contexte social (van Kleef, 2009). Nous pouvons également supposer, sur la base des travaux proposant des modèles à deux voies (Pessoa Japee et al., 2005; Pessoa, Padmala et al., 2005; Vuilleumier, 2002), que ces voies reposent sur des réseaux neuronaux distincts. La voie inférentielle impliquerait notamment le cortex préfrontal, en lien avec des fonctions cognitives comme l'attention, tandis que la

voie affective, plus automatique, mobiliserait préférentiellement des structures telles que l'amygdale ou l'insula (voir Figure 13). Etant donné que les études 2 et 3 portent sur les potentiels évoqués, il n'est pas possible d'identifier précisément les structures cérébrales impliquées. C'est pourquoi des recherches complémentaires seraient pertinentes dans ce cadre.

Pour conclure, l'intégration des expressions faciales et corporelles semble relever d'un traitement partiellement automatique, dont le degré d'automaticité varie en fonction de la charge cognitive disponible. Nos résultats suggèrent que les ressources cognitives ne sont pas toujours indispensables à cette intégration, mais qu'elles deviennent nécessaires dans certains contextes spécifiques, notamment lorsque la tâche requiert un traitement plus élaboré. La charge cognitive modulerait le déroulement temporel de l'activité cérébrale au niveau précoce (à partir de 110ms) influençant ainsi l'intégration des indices faciaux et corporels. Ce travail de thèse pose donc les bases d'un nouveau modèle d'intégration des expressions faciales et corporelles reposant sur un modèle de traitement partiellement automatique modulé par la charge cognitive (voir Figure 13). Il serait pertinent que de nouvelles études étoffent ce modèle afin de mieux comprendre les mécanismes sous-jacents à cette intégration émotionnelle ainsi que les réseaux neuronaux impliqués.

Figure 13

Représentation graphique du modèle proposé sur base du travail de thèse



Limites et perspectives

Ce travail de thèse constitue une avancée significative dans le champ des neurosciences affectives, en particulier dans l'étude de l'effet contextuel entre les expressions faciales et corporelles. Toutefois, certaines limites doivent être prises en compte afin d'ouvrir la voie à des perspectives de recherche complémentaires.

La première limite concerne la nature statique des stimuli créés et validés dans cette thèse. En effet, bien que la banque ECIFBSS offre une qualité supérieure aux stimuli utilisés précédemment en proposant des combinaisons congruentes et incongruentes d'expressions faciales et corporelles issues des mêmes acteurs, elle ne reflète qu'une représentation figée de l'expression émotionnelle. Or, dans la vie réelle, les émotions sont majoritairement perçues sous une forme dynamique. Des travaux ont montré que les émotions dynamiques sont généralement mieux reconnues en raison de l'apport du mouvement dans l'expression émotionnelle (Cunningham & Wallraven, 2009), suggérant ainsi que l'utilisation de stimuli statiques compromet la reconnaissance émotionnelle. Toutefois, il convient de noter qu'une étude antérieure ayant étudié les effets contextuels à partir de stimuli dynamiques (Nelson & Mondloch, 2017) n'a observé aucune différence majeure avec une étude utilisant des stimuli statiques (Lecker et al., 2020), indiquant que les mécanismes d'intégration émotionnelle restent comparables, quel que soit le format. Finalement, la mise à disposition de stimuli émotionnels statiques, à la communauté scientifique, reste essentielle notamment dans le cadre de certains protocoles de recherche comme lors de l'utilisation d'EEG, pour lesquels l'utilisation de vidéos introduit des

variables temporelles et perceptives plus complexes à contrôler. Dès lors, la banque ECIFBSS, malgré son format statique, constitue une avancée significative pour l'étude de la perception émotionnelle dans des conditions plus proches de la réalité que les paradigmes existants et représente la première étape vers la création d'outils méthodologiques plus écologiques. Une seconde limite liée à la banque ECIFBSS est qu'elle est constituée principalement d'émotions dites de base (joie, colère, tristesse, dégoût et neutre), limitant ainsi la représentativité des émotions plus complexes (par exemple l'émerveillement, la jalousie, l'agacement, l'amour), pourtant fréquentes dans les interactions sociales. Une perspective consisterait donc à développer une banque de stimuli dynamiques incluant ces émotions plus complexes afin de mieux rendre compte de la richesse de l'expérience émotionnelle.

Une deuxième limite concerne la sélection restreinte d'émotions dans les études expérimentales (joie, tristesse, colère). Cette limitation soulève la question de la généralisation des résultats à d'autres émotions, y compris celles peu étudiées ou plus ambiguës (par exemple la surprise). De plus, le facteur émotionnel n'a pas été intégré dans les analyses statistiques, notamment en raison du faible nombre d'essais par condition, ce qui aurait limité la puissance statistique et complexifié l'interprétation des analyses bayésiennes. Cette absence empêche d'évaluer si les effets observés sur l'automatisme s'appliquent de manière équivalente aux trois émotions testées. Il serait donc pertinent, dans de futurs travaux, d'inclure une gamme plus large d'émotions (de base et complexes) et de tester leurs effets en interaction avec la charge cognitive ou la rapidité de traitement.

Troisièmement, comme cela a été explicité dans l'introduction de la thèse, des différences interindividuelles modulent l'effet contextuel. En effet, certaines personnes ont tendance à s'appuyer davantage sur le contexte (par exemple, les expressions corporelles), tandis que d'autres se focalisent davantage sur le visage pour reconnaître l'émotion exprimée par autrui (Ensenberg-Diamant et al., 2025). Ces différences interindividuelles entraînent des variations dans l'effet contextuel, notamment les individus sensibles au contexte adoptent un traitement plus holistique des stimuli émotionnels, les rendant plus réceptifs à l'effet contextuel (Ensenberg-Diamant et al., 2025). Cependant, dans les études de cette thèse, nous n'avons pas contrôlé ces différences interindividuelles. Il est donc possible que celles-ci aient eu un impact sur les observations résultant des études 2 et 3. Il serait donc pertinent dans de futures études de prendre en compte ces différences et explorer si celles-ci ont un impact sur l'effet contextuel et notamment son traitement automatique.

Cette thèse de doctorat souligne également le rôle central des postures corporelles dans la perception émotionnelle. Explorer l'influence des postures corporelles dans la reconnaissance émotionnelle est essentielle, notamment car dans de nombreuses situations sociales, les expressions faciales ne constituent pas toujours une source d'informations émotionnelles fiables. Par exemple, elles peuvent être partiellement visibles, peu informatives ou volontairement contrôlées. Dans ces cas, le corps fournit des indices émotionnels essentiels à une reconnaissance efficace des émotions et compense le manque de clarté des expressions faciales. Comprendre comment ces indices corporels sont

reconnus et intégrés contribue à affiner les modèles théoriques sur la reconnaissance émotionnelle et ouvre des perspectives de recherche pertinente notamment dans des contextes cliniques, par exemple où les signaux émotionnels peuvent être altérés (p.ex. dans les troubles anxieux, de stress post-traumatique ou d'usage d'alcool). Enfin, les résultats de cette thèse apportent des éléments fondamentaux sur l'intégration des visages et des corps en montrant que l'effet contextuel n'est pas totalement automatique, notamment en raison de l'influence bidirectionnelle et asymétrique. Des travaux futurs sont nécessaires pour approfondir ces mécanismes et préciser comment les indices corporels contribuent, à différents stades de traitement, à la reconnaissance émotionnelle.

Par ailleurs, cette thèse s'est focalisée sur l'exploration de deux des trois critères constituant le traitement automatique de l'effet contextuel entre les expressions faciales et corporelles, c'est-à-dire la rapidité de traitement et la charge cognitive (Moors & De Houwer, 2006), excluant ainsi le critère de conscience. Les deux études menées apportent des éléments nouveaux sur le traitement de l'effet contextuel, mais restent cependant limitées par leurs limites procédurales citées précédemment et l'absence de manipulation du troisième critère de conscience. Ainsi la mise en place d'un protocole expérimental manipulant le niveau de conscience, par exemple en comparant des conditions de présentation consciente et inconsciente (Karaaslan et al., 2020), constituerait une perspective de recherche pertinente. Une telle étude permettrait non seulement de renforcer les conclusions tirées des travaux présentés dans cette thèse, mais aussi de mieux cerner les mécanismes sous-jacents à l'effet contextuel. Par ailleurs, il serait également

pertinent de réaliser une étude utilisant un paradigme de reconnaissance implicite des émotions. En effet, nos deux expériences portant sur l'automatisme reposent exclusivement sur des tâches de reconnaissance explicite des émotions. Ce type de paradigme engage inévitablement des processus conscients, attentionnels et décisionnels, ce qui limite la possibilité d'isoler les composantes automatiques du traitement émotionnel. Il pourrait, par exemple, être demandé aux participants de réaliser une tâche parallèle non émotionnelle comme demander le genre ou l'âge des individus sur l'image (Monroe et al., 2013) et d'observer si cette tâche a un impact sur l'effet contextuel des indices faciaux et corporels. L'inclusion d'un tel paradigme implicite dans une tâche de reconnaissance des indices faciaux et corporels permettrait donc de dissocier les traitements intentionnels des réponses réellement automatiques, offrant ainsi une compréhension plus fine de l'automatisme du processus d'intégration des indices émotionnels. Enfin, d'autres modèles théoriques de l'automatisme incluent des critères complémentaires, tels que l'intentionnalité du traitement et son caractère non-arrêtable (Bargh, 2014; Bargh & Tota, 1988; Bargh et al., 1996). Ces dimensions ont déjà été explorées dans une étude comportementale antérieure (Aviezer et al., 2011), qui a mis en évidence une automatisme de l'effet contextuel. Cependant, à ce jour, aucune donnée ne permet de déterminer si cette automatisme se manifeste également au niveau de l'activité cérébrale selon ces modèles (Bargh, 2014; Bargh & Tota, 1988; Bargh et al., 1996).

En outre, cette thèse s'est focalisée sur l'intégration de signaux émotionnels visuels, à savoir les expressions faciales et corporelles. Or, dans les situations naturelles, les

émotions sont perçues à travers de multiples canaux sensoriels, notamment les indices auditifs (voix, musique) ou contextuels (environnement visuel comme un lieu de fête ou de conflit). Ne prendre en compte que les indices faciaux et corporels réduit donc la portée écologique des résultats. Il serait essentiel que les recherches futures explorent l'effet contextuel dans des contextes multimodaux afin de mieux comprendre l'intégration des émotions telle qu'elle se produit dans la vie quotidienne. Par exemple en étudiant les interactions entre des expressions faciales et corporelles perçues simultanément avec des stimuli sonores tels que des voix ou de la musique. Cela permettrait de recréer des situations de la vie quotidienne comme lors d'une rencontre avec un ami, un entretien d'embauche, etc.

Finalement, ce travail de thèse s'est centré sur une population majoritairement composée d'étudiants universitaires. Cette homogénéité de l'échantillon limite la possibilité d'examiner l'influence de facteurs interindividuels tels que l'âge ou les différences dans la prise en compte du contexte émotionnel. Or, de récentes études montrent que la sensibilité au contexte n'est pas uniforme entre les individus. Autrement dit, il existe de fortes variations interindividuelles dans le poids relatif accordé au visage ou au contexte, comme la posture corporelle, dans la reconnaissance émotionnelle. En effet, certaines personnes s'appuient davantage sur les informations contextuelles, tandis que d'autres privilégient principalement les indices faciaux (Ensenberg-Diamant et al., 2025). De même, l'âge semble moduler l'utilisation du contexte émotionnel. Par exemple, les adultes plus âgés accordent un poids plus important au corps que les jeunes adultes

lorsqu'ils reconnaissent une émotion, ce qui suggère une approche plus holistique potentiellement liée à une expertise sociale accrue (Abo Foul et al., 2018). Ces facteurs interindividuels pourraient donc influencer la manière dont les indices faciaux et corporels interagissent lors de reconnaissance émotionnelle. Dans nos deux études portant sur l'automatisme de l'effet contextuel, nous n'avons pas intégré ces variables interindividuelles, ni l'âge, ni les différences de sensibilité au contexte, ce qui a pu introduire un biais dans nos résultats. Par ailleurs, notre travail s'est limité à une population tout-venant, sans prise en compte de troubles susceptibles de modifier la perception émotionnelle, tels que le trouble de l'usage d'alcool ou le trouble de stress post-traumatique. Certaines études montrent que les personnes présentant un trouble d'usage d'alcool présentent un déficit cross-modal, c'est-à-dire qu'elles utilisent moins systématiquement les émotions contextuelles comme les voix ou le contexte pour reconnaître des émotions faciales entraînant des déficits dans la reconnaissance émotionnelle (Maurage & Campanella, 2014; Maurage & Campanella, 2013; Maurage et al., 2007; Maurage, Joassin et al., 2013; Maurage et al., 2008; Maurage, Rombaux et al., 2013). Concernant le trouble de stress post-traumatique, ce domaine est encore très peu exploré, il existe à notre connaissance une seule étude montrant que les émotions menaçantes multimodales (visage et voix) telles que la colère sont mieux reconnues que les émotions positives telles que la joie, suggérant une sensibilité accrue aux menaces et une incapacité à percevoir les signaux de sécurité (Williams et al., 2018). Cependant, que cela soit pour le trouble d'usage d'alcool ou pour le stress post-traumatique, certaines limites sont présentes dans la littérature : (1) nous avons très peu de connaissance sur les

mécanismes sous-jacents aux effets contextuels limitant nos connaissances des activités cérébrales associées à cet effet; (2) les études actuelles utilisent des stimuli de très faibles qualités remettant en question la validité écologique des résultats obtenus; (3) aucune étude actuelle n'a permis d'explorer le caractère automatique ou volontaire des effets contextuels lors de la présence d'un trouble clinique tel que l'usage d'alcool ou le stress post-traumatique. Une perspective de ce travail de thèse serait par conséquent d'étendre les études réalisées présentement chez une population clinique afin de mieux comprendre comment les émotions provenant de multiples indices émotionnels sont traitées et influencées par des troubles psychopathologiques.

Conclusion générale

Ce travail de thèse apporte à la communauté scientifique des outils innovants et des avancées théoriques majeures sur le traitement simultané des expressions faciales et corporelles. D'un point de vue méthodologique, la création et la validation de la banque ECIFBSS constituent une contribution importante en mettant à disposition des chercheurs un ensemble d'images d'expressions faciales et corporelles en situations congruentes et incongruentes, à la fois plus proches des conditions de perception réelle et rigoureusement contrôlée sur des dimensions telles que l'activation, la valence, l'intensité et l'authenticité. Sur le plan théorique, ce travail soutient et prolonge les modèles comportementaux antérieurs suggérant une automaticité de l'effet contextuel entre les indices faciaux et corporels, en mobilisant deux critères : la charge cognitive et la rapidité de traitement (Aviezer et al., 2011; Meeren et al., 2005). Par ailleurs, les données électrophysiologiques issues de ce travail apportent un éclairage nouveau, en montrant que cette intégration émotionnelle semble se réaliser de manière automatique au niveau cérébral. En effet, aucune interaction entre la congruence émotionnelle et les critères d'automaticité n'a été observée, ce qui suggère un traitement rapide et ne nécessitant pas de ressource cognitive. Ces résultats devront toutefois être consolidés par des études futures, afin de mieux caractériser les mécanismes cérébraux sous-jacents et de vérifier l'absence de biais méthodologiques pouvant affecter cette interprétation. En définitive, cette thèse pose les fondations d'une recherche intégrant données comportementales et cérébrales sur l'effet

contextuel entre expressions faciales et corporelles, en proposant des stimuli validés, contrôlés et écologiques, dans une perspective d'exploration plus proche de la perception émotionnelle réelle.

Références générales

- Abbenhuis, M. A., Raaijmakers, W. G. M., Raaijmakers, J. G. W., & van Woerden, G. J. M. (1990). Episodic memory in dementia of the Alzheimer type and in normal ageing: Similar impairment in automatic processing. *The Quarterly Journal of Experimental Psychology Section A*, *42*(3), 569-583. <https://doi.org/10.1080/14640749008401237>
- Abo Foul, Y., Eitan, R., & Aviezer, H. (2018). Perceiving emotionally incongruent cues from faces and bodies: Older adults get the whole picture. *Psychology and Aging*, *33*(4), 660-666. <https://doi.org/10.1037/pag0000255>
- Abramson, L., Petranker, R., Marom, I., & Aviezer, H. (2021). Social interaction context shapes emotion recognition through body language, not facial expressions. *Emotion*, *21*(3), 557-568. <https://doi.org/10.1037/emo0000718>
- Almeida, P. R., Ferreira-Santos, F., Chaves, P. L., Paiva, T. O., Barbosa, F., & Marques-Teixeira, J. (2016). Perceived arousal of facial expressions of emotion modulates the N170, regardless of emotional category: Time domain and time-frequency dynamics. *International Journal of Psychophysiology*, *99*, 48-56. <https://doi.org/10.1016/j.ijpsycho.2015.11.017>
- Amting, J. M., Greening, S. G., & Mitchell, D. G. (2010). Multiple mechanisms of consciousness: The neural correlates of emotional awareness. *Journal of Neuroscience*, *30*(30), 10039-10047. <https://doi.org/10.1523/JNEUROSCI.6434-09.2010>
- Anderson, A. K., Christorf, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. (2013). Neural correlates of the automatic processing of threat facial signals. *Journal of Neuroscience*, *23*(13), 5627-5633. <https://doi.org/10.1523/JNEUROSCI.23-13-05627.2003>
- App, B., Reed, C. L., & McIntosh, D. N. (2012). Relative contributions of face and body configurations: Perceiving emotional state and motion intention. *Cognition & Emotion*, *26*(4), 690-698. <https://doi.org/10.1080/02699931.2011.588688>
- Arnold, M. B. (1960). *Emotion and personality (Vol. I). Psychological aspects*. Columbia University Press
- 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. <https://doi.org/10.1068/p5096>

- Aviezer, H., Bentin, S., Dudarev, V., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion, 11*(6), 1406-1414. <https://doi.org/10.1037/a0023578>
- 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. <https://doi.org/10.1111/j.1467-9280.2008.02148.x>
- Aviezer, H., Trope, Y., & Todorov, A. (2012a). Holistic person processing: Faces with *103*(1), 20-37. <https://doi.org/10.1037/a0027411>
- Aviezer, H., Trope, Y., & Todorov, A. (2012b). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225-1229. <https://doi.org/10.1126/science.1224313>
- Baier, D., Kempkes, M., Ditye, T., & Ansorge, U. (2022). Do subliminal fearful facial expressions capture attention? *Frontiers in Psychology, 13*, Article 840746. <https://doi.org/10.3389/fpsyg.2022.840746>
- Balconi, M., & Carrera, A. (2011). Cross-modal integration of emotional face and voice in congruous and incongruous pairs: The P2 ERP effect. *Journal of Cognitive Psychology, 23*(1), 132-139. <https://doi.org/10.1080/20445911.2011.473560>
- Bargh, J. A. (2014). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. Dans R. Wyer, Jr. & T. K. Srull (Éds), *Handbook of social cognition* (2^e éd., pp. 1-40). Lawrence Erlbaum Associates, Inc.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*(2), 230-244. <https://doi.org/10.1037/0022-3514.71.2.230>
- Bargh, J. A., & Tota, M. E. (1988). Context-dependent automatic processing in depression: Accessibility of negative constructs with regard to self but not others. *Journal of Personality and Social Psychology, 54*(6), 925-939. <https://doi.org/10.1037/0022-3514.54.6.925>
- Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research, 17*(3), 613-620. [https://doi.org/10.1016/s0926-6410\(03\)00174-5](https://doi.org/10.1016/s0926-6410(03)00174-5)
- Borgomaneri, S., Bolloni, C., Sessa, P., & Avenanti, A. (2020). Blocking facial mimicry affects recognition of facial and body expressions. *PLoS One, 15*(2), Article e0229364. <https://doi.org/10.1371/journal.pone.0229364>

- Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences*, *10*(4), 159-166. <https://doi.org/10.1016/j.tics.2006.02.004>
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*(8), 641-651. <https://doi.org/10.1038/nrn1724>
- Calvo, M. G., Gutiérrez-García, A., Fernández-Martín, A., & Nummenmaa, L. (2014). Recognition of facial expressions of emotion is related to their frequency in everyday life. *Journal of Nonverbal Behavior*, *38*(4), 549-567. <https://doi.org/10.1007/s10919-014-0191-3>
- Calvo, M. G., Marrero, H., & Beltran, D. (2013). When does the brain distinguish between genuine and ambiguous smiles? An ERP study. *Brain and Cognition*, *81*(2), 237-246. <https://doi.org/10.1016/j.bandc.2012.10.009>
- Campanella, S., Quinet, P., Bruyer, R., Crommelinck, M., & Guerit, J. M. (2002). Categorical perception of happiness and fear facial expressions: An ERP study. *Journal of Cognitive Neuroscience*, *14*(2), 210-227. <https://doi.org/10.1162/089892902317236858>
- Cao, F., Zeng, K., Li, W., Liu, S., Zhang, L., Katembu, S., & Xu, Q. (2022). Influence of scene-based expectation on facial expression perception: The moderating effect of cognitive load. *Biological Psychology*, *168*, Article 108247. <https://doi.org/10.1016/j.biopsycho.2021.108247>
- Carretie, L., Hinojosa, J. A., Martin-Loeches, M., Mercado, F., & Tapia, M. (2004). Automatic attention to emotional stimuli: Neural correlates. *Human Brain Mapping*, *22*(4), 290-299. <https://doi.org/10.1002/hbm.20037>
- Chen, T., Sun, Y., Feng, C., & Feng, W. (2022). In identifying the source of the incongruent effect. *Journal of Psychophysiology*, *36*(3), 167-176. <https://doi.org/10.1027/0269-8803/a000290>
- Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, *28*, 117-139. <https://doi.org/10.1023/B:JONB.0000023655.25550.be>
- Cunningham, D. W., & Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. *Journal of Vision*, *9*(13), Article 7. <https://doi.org/10.1167/9.13.7>

- Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion, 12*(5), 1085-1101. <https://doi.org/10.1037/a0025737>
- Darwin, C. (1872). *The expression of the emotions in man and animals*. John Murray.
- Darwin, C., & Prodger, P. (1998). *The expression of the emotions in man and animals*. Oxford University Press.
- Dawel, A., Miller, E. J., Horsburgh, A., & Ford, P. (2022). A systematic survey of face stimuli used in psychological research 2000-2020. *Behavior Research Methods, 54*(4), 1889-1901. <https://doi.org/10.3758/s13428-021-01705-3>
- Dawel, A., Wright, L., Irons, J., Dumbleton, R., Palermo, R., O'Kearney, R., & McKone, E. (2017). Perceived emotion genuineness: Normative ratings for popular facial expression stimuli and the development of perceived-as-genuine and perceived-as-fake sets. *Behavior Research Methods, 49*(4), 1539-1562. <https://doi.org/10.3758/s13428-016-0813-2>
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience, 7*(3), 242-249. <https://doi.org/10.1038/nrn1872>
- de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3475-3484. <https://doi.org/10.1098/rstb.2009.0190>
- de Gelder, B., de Borst, A. W., & Watson, R. (2015). The perception of emotion in body expressions. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(2), 149-158. <https://doi.org/10.1002/wcs.1335>
- de Gelder, B., Meeren, H. K., Righart, R., van den Stock, J., van de Riet, W. A., & Tamietto, M. (2006). Beyond the face: Exploring rapid influences of context on face processing. *Progress in Brain Research, 155*, 37-48. [https://doi.org/10.1016/S0079-6123\(06\)55003-4](https://doi.org/10.1016/S0079-6123(06)55003-4)
- de Gelder, B., & van den Stock, J. (2011). The Bodily Expressive Action Stimulus Test (BEAST). Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in Psychology, 2*, Article 181. <https://doi.org/10.3389/fpsyg.2011.00181>
- Deffke, I., Sander, T., Heidenreich, J., Sommer, W., Curio, G., Trahms, L., & Lueschow, A. (2007). MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus. *Neuroimage, 35*(4), 1495-1501. <https://doi.org/10.1016/j.neuroimage.2007.01.034>

- Durston, A. J., & Itier, R. J. (2021). The early processing of fearful and happy facial expressions is independent of task demands – Support from mass univariate analyses. *Brain research*, 1765, Article 147505. <https://doi.org/10.1016/j.brainres.2021.147505>
- Durston, A. J., & Itier, R. J. (2025). Event-related potentials to facial expressions are related to stimulus-level perceived arousal and valence. *Psychophysiology*, 62(3), Article e70045. <https://doi.org/10.1111/psyp.70045>
- Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES: A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, 42(1), 351-362. <https://doi.org/10.3758/BRM.42.1.351>
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99(3), 550-553. <https://doi.org/10.1037/0033-295X.99.3.550>
- Ekman, P. (1999). Basic emotions. Dans T. Dalgleish & M. J. Power (Éds), *Handbook of cognition and emotion* (pp. 45-60). John Wiley & Sons Ltd. <https://doi.org/10.1002/0470013494.ch3>
- Ekman, P., & Friesen, W. V. (1967). Head and body cues in the judgment of emotion: A reformulation. *Perceptual and Motor Skills*, 24, 711-724. <https://doi.org/10.2466/pms.1967.24.3.711>
- Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology & Nonverbal Behavior*, 1(1), 56-75. <https://doi.org/10.1007/BF01115465>
- Ensenberg-Diamant, N., Hassin, R. R., & Aviezer, H. (2025). Profound individual differences in contextualized emotion perception. *Journal of Experimental Psychology: General*, 154(5), 1236-1249. <https://doi.org/10.1037/xge0001692>
- Erthal, F. S., De Oliveira, L., Mocaiber, I., Pereira, M. G., Machado-Pinheiro, W., Volchan, E., & Pessoa, L. (2005). Load-dependent modulation of affective picture processing. *Cognitive, Affective, & Behavioral Neuroscience*, 5(4), 388-395. <https://doi.org/10.3758/CABN.5.4.388>
- Fernandes-Magalhaes, R., Carpio, A., Ferrera, D., van Ryckeghem, D., Pelaez, I., Barjola, P., De Lahoz, M. E., Martin-Buro, M. C., Hinojosa, J. A., van Damme, S., Carretie, L., & Mercado, F. (2023). Pain E-motion Faces Database (PEMF): Pain-related micro-clips for emotion research. *Behavior Research Methods*, 55(7), 3831-3844. <https://doi.org/10.3758/s13428-022-01992-4>

- Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science, 18*(12), 1050-1057. <https://doi.org/10.1111/j.1467-9280.2007.02024.x>
- Frijda, N. H. (1987). Emotion, cognitive structure, and action tendency. *Cognition & Emotion, 1*(2), 115-143. <https://doi.org/10.1080/02699938708408043>
- Frijda, N. (1988). *The laws of emotion*. Psychology Press.
- Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology, 57*(2), 212-228. <https://doi.org/10.1037/0022-3514.57.2.212>
- Gajewski, P. D., & Stoerig, P. (2011). N170 – An index of categorical face perception? *Journal of Psychophysiology, 25*(4), 174-179. <https://doi.org/10.1027/0269-8803/a000057>
- Gao, X., Maurer, D., & Nishimura, M. (2010). Similarities and differences in the perceptual structure of facial expressions of children and adults. *Journal of Experimental Child Psychology, 105*(1-2), 98-115. <https://doi.org/10.1016/j.jecp.2009.09.001>
- Gauthier, I., & Curby, K. M. (2005). A perceptual traffic jam on highway N170: Interference between face and car expertise. *Current Directions in Psychological Science, 14*(1), 30-33. <https://doi.org/10.1111/j.0963-7214.2005.00329.x>
- Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 431-446. <https://doi.org/10.1037/0096-1523.28.2.431>
- Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform “face area” is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience, 12*(3), 495-504. <https://doi.org/10.1162/089892900562165>
- Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion, 14*(2), 251-262. <https://doi.org/10.1037/a0036052>
- Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: A validation study. *Cognition & Emotion, 22*(6), 1094-1118. <https://doi.org/10.1080/02699930701626582>

- Grimshaw, G. M., Kranz, L. S., Carmel, D., Moody, R. E., & Devue, C. (2018). Contrasting reactive and proactive control of emotional distraction. *Emotion, 18*(1), 26-38. <https://doi.org/10.1037/emo0000337>
- Gu, Y., Mai, X., & Luo, Y. J. (2013). Do bodily expressions compete with facial expressions? Time course of integration of emotional signals from the face and the body. *PLoS One, 8*(7), Article e66762. <https://doi.org/10.1371/journal.pone.0066762>
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Current Biology, 13*(24), 2201-2205. <https://doi.org/10.1016/j.cub.2003.11.049>
- Hawk, S. T., Fischer, A. H., & van Kleef, G. A. (2012). Face the noise: Embodied responses to nonverbal vocalizations of discrete emotions. *Journal of Personality and Social Psychology, 102*(4), 796-814. <https://doi.org/10.1037/a0026234>
- Hietanen, J. K., & Nummenmaa, L. (2011). The naked truth: The face and body sensitive N170 response is enhanced for nude bodies. *PLoS One, 6*(11), Article e24408. <https://doi.org/10.1371/journal.pone.0024408>
- Hinojosa, J. A., Mercado, F., & Carretie, L. (2015). N170 sensitivity to facial expression: A meta-analysis. *Neuroscience & Biobehavioral Reviews, 55*, 498-509. <https://doi.org/10.1016/j.neubiorev.2015.06.002>
- Hu, Y., Baragchizadeh, A., & O'Toole, A. J. (2020). Integrating faces and bodies: Psychological and neural perspectives on whole person perception. *Neuroscience & Biobehavioral Reviews, 112*, 472-486. <https://doi.org/10.1016/j.neubiorev.2020.02.021>
- Jacques, C., & Caharel, S. (2022). The time course of categorical perception of facial expressions. *Neuropsychologia, 177*, Article 108424. <https://doi.org/10.1016/j.neuropsychologia.2022.108424>
- James, W. (1894). Discussion: The physical basis of emotion. *Psychological Review, 1*(5), 516-529. <https://doi.org/10.1037/h0065078>
- James, W. (1994). The physical basis of emotion. *Psychological Review, 101*(2), 205-210. <https://doi.org/10.1037/0033-295X.101.2.205>
- 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. <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>

- Karaaslan, A., Durmuş, B., & Amado, S. (2020). Does body context affect facial emotion perception and eliminate emotional ambiguity without visual awareness? *Visual Cognition*, 28(10), 605-620. <https://doi.org/10.1080/13506285.2020.1846649>
- Khalid, S., Horstmann, G., Ditye, T., & Ansorge, U. (2017). Measuring the emotion-specificity of rapid stimulus-driven attraction of attention to fearful faces: Evidence from emotion categorization and a comparison with disgusted faces. *Psychological Research*, 81(2), 508-523. <https://doi.org/10.1007/s00426-016-0743-8>
- Kiss, M., & Eimer, M. (2008). ERPs reveal subliminal processing of fearful faces. *Psychophysiology*, 45(2), 318-326. <https://doi.org/10.1111/j.1469-8986.2007.00634.x>
- Kret, M. E., & de Gelder, B. (2013). When a smile becomes a fist: The perception of facial and bodily expressions of emotion in violent offenders. *Experimental Brain Research*, 228(4), 399-410. <https://doi.org/10.1007/s00221-013-3557-6>
- Kret, M. E., Pichon, S., Grezes, J., & de Gelder, B. (2011). Similarities and differences in perceiving threat from dynamic faces and bodies. An fMRI study. *Neuroimage*, 54(2), 1755-1762. <https://doi.org/10.1016/j.neuroimage.2010.08.012>
- Kret, M. E., Roelofs, K., Stekelenburg, J. J., & de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience*, 7, Article 810. <https://doi.org/10.3389/fnhum.2013.00810>
- Kret, M. E., Stekelenburg, J. J., Roelofs, K., & de Gelder, B. (2013). Perception of face and body expressions using electromyography, pupillometry and gaze measures. *Frontiers in Psychology*, 4, Article 28. <https://doi.org/10.3389/fpsyg.2013.00028>
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, 24(8), 1377-1388. <https://doi.org/10.1080/02699930903485076>
- Lazarus, R. S. (2000). *Evolution of a model of stress, coping, and discrete emotions*. Dans V. H. Rice (Éd.), *Handbook of stress, coping, and health: Implications for nursing research, theory, and practice* (2^e éd., pp. 199-223). Sage Publications, Inc.
- Lecker, M., Dotsch, R., Bijlstra, G., & Aviezer, H. (2020). Bidirectional contextual influence between faces and bodies in emotion perception. *Emotion*, 20(7), 1154-1164. <https://doi.org/10.1037/emo0000619>

- LeDoux, J. E. (1990). Fear pathways in the brain: Implications for a theory of the emotional brain. Dans P. F. Brain, S. Parmigiani, R. J. Blanchard, & D. Mainardi (Éds), *Fear and defence* (pp. 163-177). Harwood Academic Publishers.
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Lee, S. A., Kim, C. Y., Shim, M., & Lee, S. H. (2017). Gender differences in neural responses to perceptually invisible fearful face: An ERP study. *Frontiers in Behavioral Neuroscience, 11*, Article 6. <https://doi.org/10.3389/fnbeh.2017.00006>
- Lenzoni, S., Bozzoni, V., Burgio, F., de Gelder, B., Wennberg, A., Botta, A., Pegoraro, E., & Semenza, C. (2020). Recognition of emotions conveyed by facial expression and body postures in myotonic dystrophy (DM). *Cortex, 127*, 58-66. <https://doi.org/10.1016/j.cortex.2020.02.005>
- Li, X. (2021). Recognition characteristics of facial and bodily expressions: Evidence from ERPs. *Frontiers in Psychology, 12*, Article 680959. <https://doi.org/10.3389/fpsyg.2021.680959>
- Lichtenstein-Vidne, L., Okon-Singer, H., Cohen, N., Todder, D., Aue, T., Nemets, B., & Henik, A. (2017). Attentional bias in clinical depression and anxiety: The impact of emotional and non-emotional distracting information. *Biological Psychology, 122*, 4-12. <https://doi.org/10.1016/j.biopsycho.2016.07.012>
- Lien, M.-C., Taylor, R., & Ruthruff, E. (2013). Capture by fear revisited: An electrophysiological investigation. *Journal of Cognitive Psychology, 25*(7), 873-888. <https://doi.org/10.1080/20445911.2013.833933>
- Lima, C. F., Anikin, A., Monteiro, A. C., Scott, S. K., & Castro, S. L. (2019). Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion, 19*(2), 219-233. <https://doi.org/10.1037/emo0000429>
- Lopez, L. D., Reschke, P. J., Knothe, J. M., & Walle, E. A. (2017). Postural communication of emotion: Perception of distinct poses of five discrete emotions. *Frontiers in Psychology, 8*, Article 710. <https://doi.org/10.3389/fpsyg.2017.00710>
- Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT Press.
- Luo, W., Feng, W., He, W., Wang, N. Y., & Luo, Y. J. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *Neuroimage, 49*(2), 1857-1867. <https://doi.org/10.1016/j.neuroimage.2009.09.018>

- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122-1135. <https://doi.org/10.3758/s13428-014-0532-5>
- Magnée, M. J., Stekelenburg, J. J., Kemner, C., & de Gelder, B. (2007). Similar facial electromyographic responses to faces, voices, and body expressions. *Neuroreport*, *18*(4), 369-372. <https://doi.org/10.1097/WNR.0b013e32801776e6>
- Maurage, P., & Campanella, S. (2013). Experimental and clinical usefulness of crossmodal paradigms in psychiatry: An illustration from emotional processing in alcohol-dependence. *Frontiers in human neuroscience*, *7*, Article 394. <https://doi.org/10.3389/fnhum.2013.00394>
- Maurage, P., & Campanella, S. (2014). Towards a crossmodal exploration of cognitive deficits in psychopathology. *Psychologica Belgica*, *54*(3), 282-297. <https://doi.org/10.5334/pb.as>
- Maurage, P., Campanella, S., Philippot, P., Pham, T. H., & Joassin, F. (2007). The crossmodal facilitation effect is disrupted in alcoholism: A study with emotional stimuli. *Alcohol Alcohol*, *42*(6), 552-559. <https://doi.org/10.1093/alcalc/agm134>
- Maurage, P., Joassin, F., Pesenti, M., Grandin, C., Heeren, A., Philippot, P., & de Timary, P. (2013). The neural network sustaining crossmodal integration is impaired in alcohol-dependence: An fMRI study. *Cortex*, *49*(6), 1610-1626. <https://doi.org/10.1016/j.cortex.2012.04.012>
- Maurage, P., Philippot, P., Joassin, F., Pauwels, L., Pham, T., Prieto, E. A., Palmero-Soler, E., Zanow, F., & Campanella, S. (2008). The auditory-visual integration of anger is impaired in alcoholism: An event-related potentials study. *Journal of Psychiatry and Neuroscience*, *33*(2), 111-122. <https://doi.org/10.1016/j.ijpsycho.2008.05.059>
- Maurage, P., Rombaux, P., & de Timary, P. (2013). Olfaction in alcohol-dependence: A neglected yet promising research field. *Frontiers in Psychology*, *4*, Article 1007. <https://doi.org/10.3389/fpsyg.2013.01007>
- Maxwell, J. W., Joseph, M., & Ruthruff, E. (2021). Capacity-free automatic processing of facial expressions of emotion. *Emotion*, *21*(8), 1771-1780. <https://doi.org/10.1037/emo0000965>
- Maxwell, J. W., Sanchez, D. N., & Ruthruff, E. (2023). Infrequent facial expressions of emotion do not bias attention. *Psychological Research*, *87*(8), 2449-2459. <https://doi.org/10.1007/s00426-023-01844-6>

- McCrackin, S. D., & Itier, R. J. (2021). I can see it in your eyes: Perceived gaze direction impacts ERP and behavioural measures of affective theory of mind. *Cortex*, *143*, 205-222. <https://doi.org/10.1016/j.cortex.2021.05.024>
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, *102*(45), 16518-16523. <https://doi.org/10.1073/pnas.0507650102>
- Mondloch, C. J., Nelson, N. L., & Horner, M. (2013). Asymmetries of influence: Differential effects of body postures on perceptions of emotional facial expressions. *PLoS One*, *8*(9), Article e73605. <https://doi.org/10.1371/journal.pone.0073605>
- Moody, E. J., Reed, C. L., van Bommel, T., App, B., & McIntosh, D. N. (2017). Emotional mimicry beyond the face? *Social Psychological and Personality Science*, *9*(7), 844-852. <https://doi.org/10.1177/1948550617726832>
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297-326. <https://doi.org/10.1037/0033-2909.132.2.297>
- Morningstar, M., Gilbert, A. C., Burdo, J., Leis, M., & Dirks, M. A. (2021). Recognition of vocal socioemotional expressions at varying levels of emotional intensity. *Emotion*, *21*(7), 1570-1575. <https://doi.org/10.1037/emo0001024>
- Mumenthaler, C., & Sander, D. (2015). Automatic integration of social information in emotion recognition. *Journal of Experimental Psychology: General*, *144*(2), 392-399. <https://doi.org/10.1037/xge0000059>
- Neal, D. T., & Chartrand, T. L. (2011). Embodied emotion perception. *Social Psychological and Personality Science*, *2*(6), 673-678. <https://doi.org/10.1177/1948550611406138>
- Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Visual Cognition*, *25*(4-6), 563-574. <https://doi.org/10.1080/13506285.2017.1301615>
- Niedenthal, P. M. (2007). Embodying emotion. *Science*, *316*(5827), 1002-1005. <https://doi.org/10.1126/science.1136930>
- O'Reilly, H., Pigat, D., Fridenson, S., Berggren, S., Tal, S., Golan, O., Bolte, S., Baron-Cohen, S., & Lundqvist, D. (2016). The EU-Emotion Stimulus Set: A validation study. *Behavior Research Methods*, *48*(2), 567-576. <https://doi.org/10.3758/s13428-015-0601-4>

- Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience*, 2(3-4), 167-178. <https://doi.org/10.1080/17470910701391943>
- Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, 80(3), Article 381. <https://doi.org/10.1037/0022-3514.80.3.381>
- Ong, J. H., Leung, F. Y. N., & Liu, F. (2023). The Reading Everyday Emotion Database (REED): A set of audio-visual recordings of emotions in music and language. *Lang Resources & Evaluation*, 59, 27-49. <https://doi.org/10.1007/s10579-023-09698-5>
- Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, 45(1), 75-92. <https://doi.org/10.1016/j.neuropsychologia.2006.04.025>
- Pegna, A. J., Khateb, A., Lazeyras, F., & Seghier, M. L. (2005). Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nature Neuroscience*, 8(1), 24-25. <https://doi.org/10.1038/nn1364>
- Pessoa, L., Japee, S., & Ungerleider, L. G. (2005). Visual awareness and the detection of fearful faces. *Emotion*, 5(2), 243-247. <https://doi.org/10.1037/1528-3542.5.2.243>
- Pessoa, L., Kastner, S., & Ungerleider, L. G. (2002). Attentional control of the processing of neutral and emotional stimuli. *Cognitive Brain Research*, 15(1), 31-45. [https://doi.org/10.1016/S0926-6410\(02\)00214-8](https://doi.org/10.1016/S0926-6410(02)00214-8)
- Pessoa, L., Padmala, S., & Morland, T. (2005). Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation. *Neuroimage*, 28(1), 249-255. <https://doi.org/10.1016/j.neuroimage.2005.05.048>
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128-2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Pourtois, G., Debatisse, D., Despland, P. A., & de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Cognitive Brain Research*, 14(1), 99-105. [https://doi.org/10.1016/S0926-6410\(02\)00064-2](https://doi.org/10.1016/S0926-6410(02)00064-2)
- Puce, A., McNeely, M. E., Berrebi, M. E., Thompson, J. C., Hardee, J., & Brefczynski-Lewis, J. (2013). Multiple faces elicit augmented neural activity. *Frontiers in Human Neuroscience*, 7, Article 282. <https://doi.org/10.3389/fnhum.2013.00282>

- Puffet, A.-S., & Rigoulot, S. (2025a). Validation of the Emotionally Congruent and Incongruent Face-Body Static Set (ECIFBSS). *Behavior Research Methods*, 57(1), Article 41. <https://doi.org/10.3758/s13428-024-02550-w>
- Puffet, A.-S., & Rigoulot, S. (2025b). The role of cognitive load in automatic integration of emotional information from face and body. *Scientific Reports*, 15(1), Article 28184. <https://doi.org/10.1038/s41598-025-12511-8>
- Puffet, A.-S., Saleh, G., & Rigoulot, S. (2025). Exploring the impact of facial-body emotion incongruence: Insight from a systematic review. *Journal of Nonverbal Behavior*. <https://doi.org/10.1007/s10919-025-00488-x>
- Puls, S., & Rothermund, K. (2018). Attending to emotional expressions: No evidence for automatic capture in the dot-probe task. *Cognition and Emotion*, 32(3), 450-463. <https://doi.org/10.1080/02699931.2017.1314932>
- Reschke, P. J., & Walle, E. A. (2021). The unique and interactive effects of faces, postures, and scenes on emotion categorization. *Affective Science*, 2(4), 468-483. <https://doi.org/10.1007/s42761-021-00061-x>
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178. <https://doi.org/10.1037/h0077714>
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, 115(1), 102-141. <https://doi.org/10.1037/0033-2909.115.1.102>
- Russell, J. A., & Fernández-Dols, J. M. (1997). *The psychology of facial expression*. Cambridge University Press; Editions de la Maison des Sciences de l'Homme. <https://doi.org/10.1017/CBO9780511659911>
- Sander, D., & Scherer, K. (2019). *Traité de psychologie des émotions*. Dunod.
- Santos, I. M., Iglesias, J., Olivares, E. I., & Young, A. W. (2008). Differential effects of object-based attention on evoked potentials to fearful and disgusted faces. *Neuropsychologia*, 46(5), 1468-1479. <https://doi.org/10.1016/j.neuropsychologia.2007.12.024>
- Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, 23(7), 1307-1351. <https://doi.org/10.1080/02699930902928969>

- Schindler, K., van Gool, L., & de Gelder, B. (2008). Recognizing emotions expressed by body pose: A biologically inspired neural model. *Neural Networks*, *21*(9), 1238-1246. <https://doi.org/10.1016/j.neunet.2008.05.003>
- Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex*, *130*, 362-386. <https://doi.org/10.1016/j.cortex.2020.06.010>
- Schindler, S., Richter, T. S., Bruchmann, M., Busch, N. A., & Straube, T. (2022). Effects of task load, spatial attention, and trait anxiety on neuronal responses to fearful and neutral faces. *Psychophysiology*, *59*(11), Article e14114. <https://doi.org/10.1111/psyp.14114>
- Schindler, S., Zell, E., Botsch, M., & Kissler, J. (2017). Differential effects of face-realism and emotion on event-related brain potentials and their implications for the uncanny valley theory. *Scientific reports*, *7*, Article 45003. <https://doi.org/10.1038/srep45003>
- Schutter, D. J., de Haan, E. H., & van Honk, J. (2004). Functionally dissociated aspects in anterior and posterior electrocortical processing of facial threat. *International Journal of Psychophysiology*, *53*(1), 29-36. <https://doi.org/10.1016/j.ijpsycho.2004.01.003>
- Shaw, K., Lien, M.-C., Ruthruff, E., & Allen, P. A. (2011). Electrophysiological evidence of emotion perception without central attention. *Journal of Cognitive Psychology*, *23*(6), 695-708. <https://doi.org/10.1080/20445911.2011.586624>
- Smith, M. L. (2012). Rapid processing of emotional expressions without conscious awareness. *Cerebral Cortex*, *22*(8), 1748-1760. <https://doi.org/10.1093/cercor/bhr250>
- Sollfrank, T., Kohnen, O., Hilfiker, P., Kegel, L. C., Jokeit, H., Brugger, P., Loertscher, M. L., Rey, A., Mersch, D., Sternagel, J., Weber, M., & Grunwald, T. (2021). The effects of dynamic and static emotional facial expressions of humans and their avatars on the EEG: An ERP and ERD/ERS study. *Frontiers in Neuroscience*, *15*, Article 651044. <https://doi.org/10.3389/fnins.2021.651044>
- Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: An ERP study on the body-inversion effect. *Neuroreport*, *15*(5), 777-780. <https://doi.org/10.1097/01.wnr.0000119730.93564.e8>
- Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, *12*(1), 43-47. <https://doi.org/10.1111/1467-9280.00308>

- Teufel, C., Westlake, M. F., Fletcher, P. C., & von dem Hagen, E. (2019). A hierarchical model of social perception: Psychophysical evidence suggests late rather than early integration of visual information from facial expression and body posture. *Cognition*, *185*, 131-143. <https://doi.org/10.1016/j.cognition.2018.12.012>
- Thierry, G., Pegna, A. J., Dodds, C., Roberts, M., Basan, S., & Downing, P. (2006). An event-related potential component sensitive to images of the human body. *Neuroimage*, *32*(2), 871-879. <https://doi.org/10.1016/j.neuroimage.2006.03.060>
- Thoma, P., Soria Bauser, D., & Suchan, B. (2013). BESST (Bochum Emotional Stimulus Set): A pilot validation study of a stimulus set containing emotional bodies and faces from frontal and averted views. *Psychiatry Research*, *209*(1), 98-109. <https://doi.org/10.1016/j.psychres.2012.11.012>
- Tomasik, D., Ruthruff, E., Allen, P. A., & Lien, M. C. (2009). Nonautomatic emotion perception in a dual-task situation. *Psychonomic Bulletin & Review*, *16*(2), 282-288. <https://doi.org/10.3758/PBR.16.2.282>
- Tomkins, S. S. (1995). *Exploring affect: The selected writings of Silvan S Tomkins*. Cambridge University Press.
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242-249. <https://doi.org/10.1016/j.psychres.2008.05.006>
- Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion*, *8*(1), 81-95. <https://doi.org/10.1037/1528-3542.8.1.81>
- Tzelgov, J. (1997). Specifying the relations between automaticity and consciousness: A theoretical note. *Consciousness and Cognition*, *6*(2-3), 441-451. <https://doi.org/10.1006/ccog.1997.0303>
- Utama, N. P., Takemoto, A., Koike, Y., & Nakamura, K. (2009). Phased processing of facial emotion: An ERP study. *Neuroscience Research*, *64*(1), 30-40. <https://doi.org/10.1016/j.neures.2009.01.009>
- Vakli, P., Nemeth, K., Zimmer, M., & Kovacs, G. (2017). The electrophysiological correlates of integrated face and body-part perception. *Quarterly Journal of Experimental Psychology*, *70*(1), 142-153. <https://doi.org/10.1080/17470218.2015.1127981>

- van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E.-J., Derks, K., Dablander, F., Gronau, Q. F., Kucharský, Š., Gupta, A. R. K. N., Sarafoglou, A., Voelkel, J. G., Stefan, A., Ly, A., Hinne, M., Matzke, D., & Wagenmakers, E.-J. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *L'Année psychologique*, *120*(1), 73-96. <https://doi.org/10.3917/anpsy1.201.0073>
- van den Bergh, D., Wagenmakers, E.-J., & Aust, F. (2023). Bayesian repeated-measures analysis of variance: An updated methodology implemented in JASP. *Advances in Methods and Practices in Psychological Science*, *6*(2). <https://doi.org/10.1177/25152459231168024>
- van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*(3), 487-494. <https://doi.org/10.1037/1528-3542.7.3.487>
- van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion*, *11*(4), 907-920. <https://doi.org/10.1037/a0023853>
- van Doorn, J., van den Bergh, D., Bohm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharsky, S., Ly, A., Marsman, M., Matzke, D., Gupta, A., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E. J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*(3), 813-826. <https://doi.org/10.3758/s13423-020-01798-5>
- van Heijnsbergen, C. C., Meeren, H. K., Grezes, J., & de Gelder, B. (2007). Rapid detection of fear in body expressions: An ERP study. *Brain Research*, *1186*, 233-241. <https://doi.org/10.1016/j.brainres.2007.09.093>
- van Kleef, G. A. (2009). How emotions regulate social life: The emotions as social information (EASI) model. *Current Directions in Psychological Science*, *18*(3), 184-188. <https://doi.org/10.1111/j.1467-8721.2009.01633.x>
- Victeur, Q., Huguet, P., & Silvert, L. (2020). Attentional allocation to task-irrelevant fearful faces is not automatic: Experimental evidence for the conditional hypothesis of emotional selection. *Cognition and Emotion*, *34*(2), 288-301. <https://doi.org/10.1080/02699931.2019.1622512>
- Vuilleumier, P. (2002). Facial expression and selective attention. *Current Opinion in Psychiatry*, *15*(3), 291-300. <https://doi.org/10.1097/00001504-200205000-00011>

- Wallbott, H. G. (1992). Effects of distortion of spatial and temporal resolution of video stimuli on emotion attributions. *Journal of Nonverbal Behavior*, *16*, 5-20. <https://doi.org/10.1007/BF00986876>
- Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, *28*(6), 879-896. [https://doi.org/10.1002/\(SICI\)1099-0992\(199811\)28:6<879::AID-EJSP901>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1099-0992(199811)28:6<879::AID-EJSP901>3.0.CO;2-W)
- Wang, L., Xia, L., & Zhang, D. (2017). Face-body integration of intense emotional expressions of victory and defeat. *PLoS One*, *12*(2), Article e0171656. <https://doi.org/10.1371/journal.pone.0171656>
- Watanabe, S., Kakigi, R., & Puce, A. (2003). The spatiotemporal dynamics of the face inversion effect: A magneto- and electro-encephalographic study. *Neuroscience*, *116*(3), 879-895. [https://doi.org/10.1016/s0306-4522\(02\)00752-2](https://doi.org/10.1016/s0306-4522(02)00752-2)
- Williams, C. L., Milanak, M. E., Judah, M. R., & Berenbaum, H. (2018). The association between PTSD and facial affect recognition. *Psychiatry Research*, *265*, 298-302. <https://doi.org/10.1016/j.psychres.2018.04.055>
- Willis, M. L., Palermo, R., & Burke, D. (2011). Judging approachability on the face of it: The influence of face and body expressions on the perception of approachability. *Emotion*, *11*(3), 514-523. <https://doi.org/10.1037/a0022571>
- Winkielman, P., Niedenthal, P. M., & Oberman, L. (2008). The embodied emotional mind. Dans G. R. Semin & E. R. Smith (Éds), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 263-288). Cambridge University Press. <https://doi.org/10.1017/CBO9780511805837.012>
- Wirth, B. E., & Wentura, D. (2020). It occurs after all: Attentional bias towards happy faces in the dot-probe task. *Attention, Perception, & Psychophysics*, *82*(5), 2463-2481. <https://doi.org/10.3758/s13414-020-02017-y>
- Yang, Y. F., Brunet-Gouet, E., Burca, M., Kalunga, E. K., & Amorim, M. A. (2020). Brain processes while struggling with evidence accumulation during facial emotion recognition: An ERP study. *Frontiers in Human Neuroscience*, *14*, Article 340. <https://doi.org/10.3389/fnhum.2020.00340>
- Zhang, D., Zhao, T., Liu, Y., & Chen, Y. (2015). Comparison of facial expressions and body expressions: An event-related potential study. *Acta Psychologica Sinica*, *47*(8), 963-970. <https://doi.org/10.3724/sp.J.1041.2015.00963>

Zhang, M., Liu, T., Jin, Y., He, W., Huang, Y., & Luo, W. (2019). The asynchronous influence of facial expressions on bodily expressions. *Acta Psychologica, 200*, Article 102941. <https://doi.org/10.1016/j.actpsy.2019.102941>

Appendice
Certificats d'éthique

CERTIFICAT ÉTHIQUE POUR L'ÉTUDE 1 ET 2

3968



CERTIFICAT D'ÉTHIQUE DE LA RECHERCHE - PSYCHOLOGIE ET PSYCHOÉDUCATION

En vertu du mandat qui lui a été confié par l'Université, le Comité d'éthique de la recherche - Psychologie et Psychoéducation a analysé et approuvé pour certification éthique le protocole de recherche suivant :

Titre : L'automatisme de l'effet de congruence dans la reconnaissance d'expressions faciales émotionnelles

Chercheur(s) : Anne-Sophie Puffet
Département de psychologie

Organisme(s) : Subvention découverte (CRSNG)

N° DU CERTIFICAT CERPPE-22-06-07.03

PÉRIODE DE VALIDITÉ : Du 07 avril 2025 au 07 avril 2026

En acceptant le certificat éthique, le chercheur s'engage à :

- Aviser le CER par écrit des changements apportés à son protocole de recherche avant leur entrée en vigueur;
- Procéder au renouvellement annuel du certificat tant et aussi longtemps que la recherche ne sera pas terminée;
- Aviser par écrit le CER de l'abandon ou de l'interruption prématurée de la recherche;
- Faire parvenir par écrit au CER un rapport final dans le mois suivant la fin de la recherche.

Charles Viau-Quesnel
Président du comité

Vicky Montambault
Secrétaire du comité

CERTIFICAT ÉTHIQUE POUR L'ÉTUDE 3

4853



CERTIFICAT D'ÉTHIQUE DE LA RECHERCHE - PSYCHOLOGIE ET PSYCHOÉDUCATION

En vertu du mandat qui lui a été confié par l'Université, le Comité d'éthique de la recherche - Psychologie et Psychoéducation a analysé et approuvé pour certification éthique le protocole de recherche suivant :

Titre : Traitement automatique de l'effet contextuel lors de la reconnaissance d'expressions faciales et corporelles émotionnelles

Chercheur(s) : Anne-Sophie Puffet
Département de psychologie

Organisme(s) : CRSNG – Subvention découverte

N° DU CERTIFICAT : CERPPE-24-36-07.05

PÉRIODE DE VALIDITÉ : Du 23 janvier 2025 au 23 janvier 2026

En acceptant le certificat éthique, le chercheur s'engage à :

- Aviser le CER par écrit des changements apportés à son protocole de recherche avant leur entrée en vigueur;
- Procéder au renouvellement annuel du certificat tant et aussi longtemps que la recherche ne sera pas terminée;
- Aviser par écrit le CER de l'abandon ou de l'interruption prématurée de la recherche;
- Faire parvenir par écrit au CER un rapport final dans le mois suivant la fin de la recherche.

Charles Viau-Quesnel
Président du comité

Vicky Montambault
Secrétaire du comité