

Université de Montréal

Association entre la dépression et le biais d'attribution hostile chez les individus hostiles et non hostiles : une étude de potentiels reliés aux événements

par Alexandre Gasse

Département de psychologie
Faculté des arts et des sciences

Mémoire présenté
en vue de l'obtention du grade de la maîtrise ès sciences (M. Sc.)
en psychologie

juillet, 2020

©Alexandre Gasse, 2020

Résumé

Contexte : Bien que l'hostilité et l'agressivité ne soient pas essentielles au diagnostic de la dépression, ces symptômes se sont révélés être très prévalents chez les patients déprimés et sont associés à des niveaux de comorbidité et de sévérité de la maladie plus élevés. En outre, des études ont montré que la dépression est associée de manière significative à des biais d'interprétation négatifs et que, selon les théories cognitives de la dépression, ces biais peuvent représenter un facteur de risque considérable pour l'apparition et/ou le maintien de la maladie. Alors que les biais d'interprétation négatifs soient reconnus comme un élément fondamental des modèles cognitifs de la dépression, peu d'études ont examiné les biais spécifiques du traitement de l'information, notamment le biais d'attribution d'intention hostile, chez des individus hostiles présentant des symptômes dépressifs. Méthodologie : À l'aide de données pré-recueillies auprès d'un échantillon de 72 (hommes = 41,6%, femmes = 58,3%) participants hostiles (n = 26) et non hostiles (n = 46), cette étude visait à examiner l'association entre la dépression et le biais d'attribution d'intention hostile en déterminant si les scores de niveau de dépression étaient liés de manière unique aux mesures électrophysiologiques du biais d'attribution hostile. Résultats : Le groupe hostile a montré des niveaux de dépression et d'agressivité réactive plus élevés que le groupe non hostile. En outre, les scores de dépression étaient des prédicteurs significatifs de l'effet N400 dans la condition non hostile, alors que l'agressivité réactive ne l'était pas, tandis que dans la condition hostile, le modèle global s'est avéré significatif, avec des niveaux de dépression et d'agressivité réactive montrant tous deux de fortes tendances vers la signification. Limites : Une petite taille d'échantillon a limité la portée des conclusions de l'étude. De plus, la sélection de l'échantillon n'a pas permis d'examiner les différences de groupe spécifiques concernant le biais d'attribution hostile dans les groupes déprimés et non déprimés. Conclusion :

Cette étude souligne l'importance de mieux comprendre les mécanismes de perception associés à la dépression et d'envisager l'application de techniques de restructuration cognitive pour contrer les processus d'interprétation biaisés dans des contextes où la dépression et l'agression se croisent. Cette étude suggère également que des alternatives aux méthodologies d'auto-évaluation pourraient être plus utiles lors de l'évaluation des processus automatiques et implicites.

Mots-clés : Dépression, Hostilité, Agressivité, Biais d'attribution d'intention hostile, Électrophysiologie, EEG, N400.

Abstract

Background: Hostility and aggression have been found to be highly prevalent among depressed patients and are associated with higher comorbidity and illness severity levels. Although negative interpretation biases are a fundamental element of cognitive models of depression, few studies have examined the specific biases in information processing, mainly the hostile attribution bias, found in hostile individuals who present depressive symptoms. Method: Using pre-collected data from a sample of 72 (male=41,6%, female=58,3%) undergraduate and community-based hostile (n=26) and non-hostile (n=46) adult participants, the authors aimed to examine the association between depression and the hostile attribution bias by determining whether depression level scores were uniquely related to electrophysiological measures of the hostile attribution bias. Results: The hostile group showed higher measured levels of depression and reactive aggression, compared to the non-hostile group. Also, depression scores were significant predictors of the N400 effect in the non-hostile condition, while reactive aggression was not, whereas in the hostile condition, the overall model was significant, with depression and reactive aggression levels both showing strong trends towards significance. Limitations: A small sample size limited the scope of our conclusions. Also, sample selection prevented the authors from examining specific group differences regarding the hostile attribution bias in depressed and non-depressed groups. Conclusion: This study highlights the importance of better understanding the perception mechanisms that are associated with depression and to consider the application of cognitive restructuring techniques to counter biased interpretation processes in settings where depression and aggression intersect. This study also suggests that alternatives to self-evaluative methodologies might be more useful when assessing automatic and implicit processes.

Key words: Depression, Aggression, Hostile attribution bias, EEG, ERP, N400.

Table de matières

Résumé.....	i
Abstract.....	iii
Liste des tableaux.....	v
Liste des figures.....	vi
Liste des abréviations.....	vii
Remerciements.....	viii
Chapitre I : Introduction générale.....	1
Introduction.....	2
Méthodologie.....	12
Chapitre II : Article scientifique.....	25
Abstract.....	27
Introduction.....	28
Method.....	35
Results.....	44
Discussion.....	53
Conclusion.....	57
References.....	58
Chapitre III : Conclusion générale.....	66
Discussion.....	67
Conclusion.....	74
Bibliographie.....	76
Annexe : Article scientifique publié dans Journal of Affective Disorders.....	86

Liste des tableaux

Table 1. Examples of Four Types of Possible Scenario Sentences.....	37
Table 2. Descriptive Statistics of Right Central Amplitudes in Microvolts (μV).....	48
Table 3. Sample Descriptives Using t-test for Equality of Means.....	49
Table 4. Means, Standard Deviations, and Correlations.....	50
Table 5. Repeated Measures ANOVA with Condition, Consistency, Hemisphere and Region as Factors, and with Standardized BSI Depression Scores as a Covariate and Reactive Aggression Scores as a Control Variable.....	51
Table 6. Multiple Regression Using Non-Hostile Right Central and Right Anterior N400 as the Criterion.....	52
Table 7. Multiple Regression Using Hostile Right Central and Right Anterior N400 as the Criterion.....	53

Liste des figures

Figure 1. Lateral electrode mean ERP waveforms for the match and mismatch conditions, and for the hostile and non-hostile conditions.....	45
Figure 2. Mean ERP difference waveforms (ERP mismatch minus ERP match) for the hostile and non-hostile conditions.....	46
Figure 3. Topographic maps of the mean ERP difference waveforms (ERP mismatch minus ERP match) following (a) the violation of non-hostile intention expectancy, and (b) the violation of hostile intention expectancy, while (c) outlines the mean of the hostile and non-hostile conditions for the 450-600 postcritical word timeframe.....	47

Liste des abréviations

Ag / AgCl : Argent / chlorure d'argent

ANOVA : Analyse de la variance

BAIH / HAB : Biais d'attribution d'intention hostile

EEG : Électroencéphalographie

EOG / HEOG / VEOG : Électrooculogramme / horizontal / vertical

Hz : Hertz

ms : milliseconde

PRE / ERP : Potentiels reliés aux événements

Remerciements

Je voudrais remercier mon superviseur de recherche, le Docteur Jean Gagnon, pour son appui constant et ses conseils avisés, mais également pour l'opportunité qu'il m'a offerte de faire partie de sa merveilleuse équipe, ainsi que Wan Seo Kim, pour sa grande part dans la réalisation de cette étude. Merci aussi à Jonathan James pour son apport à mon cheminement académique, et à tous les membres du laboratoire LENS pour leur soutien et leur camaraderie. Finalement, je dédie ce mémoire à Kavita Singh, car sans son soutien, son attention, son expertise et ses conseils, je suis certain que je n'y serais pas arrivé.

Chapitre I : Introduction générale

Introduction

Recension de la littérature

Dépression et agressivité

La dépression unipolaire est une maladie hétérogène composée de plusieurs sous-types cliniques, chacun présentant une symptomatologie spécifique et un pronostic distinct à long terme (Pedrelli et al., 2013). Ainsi, certains chercheurs suggèrent que l'identification de phénotypes spécifiques liés à un pronostic plus sévère de la maladie, conduisant ainsi à des stratégies de traitement adaptées à des sous-types particuliers du trouble, pourrait avoir un impact concret sur l'incidence globale de la dépression (Pedrelli et al., 2013). Un de ces phénotypes, identifié dans plusieurs études transversales à l'aide de diverses définitions conceptuelles et mesures, serait caractérisé par des niveaux de prévalence élevés d'irritabilité, de colère et d'hostilité (Judd et al., 2013). Par exemple, 95 des 254 patients dépressifs examinés par Benazzi et Akiskal (2005) présentaient des niveaux d'irritabilité et d'hostilité importants, alors que 37% des patients d'une étude de Snaith et Taylor (1985) montraient des signes d'irritabilité externalisée modérés à sévères. De plus, Fava (1982) et Riley (1989) ont tous deux observés des niveaux élevés de colère et d'hostilité chez les patients déprimés par rapport au groupe témoin, tandis que dans une étude de Pitt et ses collègues (1968), les auteurs ont remarqué que les nouvelles mères souffrant de dépression postpartum étaient souvent excessivement colériques. Finalement, dans deux études à site unique, des taux élevés d'attaques de colère – caractérisées principalement par de l'agressivité comportementale impulsive, soit une réaction agressive en réponse au comportement d'autrui perçu comme menaçant ou intentionnel (Fite et al., 2009) – ont été observés chez 127 (44%) et 164 (39%) patients souffrants d'un trouble dépressif majeur (Fava et al., 1993, 1996).

Dans une autre optique, plusieurs études ont également démontré que la dépression est une problématique prévalente chez les personnes agressives et violentes. Par exemple, Good (1978) a constaté qu'un état dépressif sous-jacent était souvent sous-diagnostiqué chez un échantillon d'hommes et de femmes incarcérés pour des comportements agressifs et antisociaux, alors que dans la revue de Kermani (1981), 29 des 53 sujets psychiatriques masculins présentant une forte tendance à la violence comportementale ont été décrits comme ayant une personnalité dépressive violente. En outre, les résultats d'une étude suédoise de 2015 ont révélé que les probabilités de commettre un crime violent, tel qu'un homicide, une tentative d'homicide ou une tentative de voies de fait graves, étaient trois fois plus élevées chez les personnes dépressives que chez les non dépressifs (Fazel et al., 2015). Enfin, il a été établi que pour chaque augmentation de 20% des symptômes dépressifs, la probabilité d'agresser physiquement un conjoint augmentait de 74% (Pan et al., 1994). Ainsi, pour plusieurs chercheurs, les résultats rapportés ci-dessus confirment l'idée que la dépression est une problématique plus répandue chez les personnes violentes qu'initialement crue (Maiuro et al., 1989).

Enfin, bien que l'irritabilité et la colère ne soient plus centrales au diagnostic du trouble dépressif majeur chez l'adulte depuis la révision du DSM-III (American Psychiatric Association, 1987) – l'irritabilité est aujourd'hui la seule référence diagnostique à l'hostilité vis-à-vis de la dépression majeure, et uniquement en tant que caractéristique associée – ces symptômes sont en réalité très répandus, et ceux qui les présentent souffrent d'une affection dépressive plus chronique, plus comorbide, et plus sévère du trouble (Judd et al., 2013). La dépression irritable/colérique est également associée au sexe féminin, à un plus jeune âge, à une plus faible satisfaction de vie, à plus de troubles anxieux et de contrôle de l'impulsivité, à un plus haut risque de suicide, à l'abus de substance, à un début précoce des troubles de l'humeur et à des

antécédents familiaux de bipolarité (Judd et al., 2013; Benazzi et Akiskal, 2005). Ainsi, l'importance d'identifier des phénotypes de dépression spécifiques liés à des pronostics sévères a conduit certains chercheurs à proposer l'idée que l'irritabilité et la colère soient reconnues comme des indicateurs clés d'un sous-type clinique important de trouble dépressif (Judd et al., 2013).

Dépression et biais interprétatifs

La dépression majeure est non seulement un trouble psychiatrique très répandu – elle affecte 4,7% des Canadiens chaque année et représente 10,9% de toutes les hospitalisations psychiatriques (Statistics Canada, 2012) – mais c'est aussi un trouble extrêmement récurrent qui pourrait être expliqué en partie par la présence de facteurs de vulnérabilité stables, exposant certaines personnes à des épisodes répétés (Dearing et al., 2009); certaines estimations suggèrent que 80% des Américains qui ont connu un épisode dépressif en connaîtront plusieurs au courant de leur vie (Kessler et al. 2003; Boland et Keller 2002). Par conséquent, des efforts de recherche considérables, guidés principalement par la théorie cognitive, ont été consacrés à l'identification des facteurs de risque potentiels associés à ce trouble (Gotlib et Krasnoperova, 1998).

Beck (1967, 1976, dans Gotlib et Krasnoperova, 1998) a formulé une théorie de la dépression attribuant l'apparition de cette maladie en grande partie à des biais cognitifs, soulignant que les processus cognitifs mésadaptés représentent un facteur de risque important de la dépression. Le modèle cognitif de la dépression de Beck stipule que les personnes qui ont connu de l'adversité dans leur enfance développent des schémas négatifs liés à la perte, à l'échec et à l'abandon. Plus tard, l'expérience d'événements de la vie qui déclenche le rappel du traumatisme initial réactivera les schémas sous la forme de croyances ou d'attitudes rigides et inappropriées à propos de soi-même et du monde, ainsi que de normes irréalistes et

perfectionnistes à l'égard desquelles le soi est jugé (Gotlib et Krasnoperova, 1998). Les schémas négatifs agissent ainsi comme des filtres au travers desquels l'environnement est perçu, évalué et stocké en mémoire, et les individus interprètent les informations de l'environnement en conséquence afin d'assurer une certaine cohérence avec ces schémas. De ce fait, les personnes déprimées, ainsi que celles susceptibles de souffrir de ce trouble, portent sélectivement attention aux stimuli négatifs, filtrent les stimuli positifs et perçoivent les informations négatives ou neutres de façon plus négative qu'elles ne le sont en réalité (Gotlib et Krasnoperova, 1998; Dearing et al., 2009). De plus, des chercheurs ont constaté que les individus déprimés sont également plus susceptibles d'interpréter certains indices ambigus, tels que des homophones et des phrases présentés en contexte expérimental, de façon menaçante, augmentant par ce fait la probabilité que ceux et celles qui sont prédisposés à la méfiance et à la perception de menace fassent preuve d'une plus grande hostilité et, par conséquent, se montrent davantage agressifs à l'égard des autres (Mogg et al., 2006; Moreno et al., 1994). Ainsi, les chercheurs suggèrent que le trouble dépressif majeur avec irritabilité pourrait favoriser une vision hostile du monde et une hostilité interpersonnelle accrue fondées en partie sur des processus cognitifs biaisés (Riley et al., 1989; Smith et al., 2016). (Riley et al., 1989).

Biais d'attribution d'intention hostile

Selon Berkowitz (1977) et Nasby (1980), les stimuli aversifs ne génèrent pas automatiquement des réactions agressives, mais que ce sont plutôt les processus cognitifs impliqués dans le traitement de l'information, tels que les évaluations, les attentes et le discours privé, qui jouent un rôle de médiateur dans la relation entre les stimuli aversifs et les réponses agressives. Ainsi, pour qu'une réaction aggressive se produise, un individu doit attribuer une intention aggressive ou hostile à un indice aversif pour que celui-ci serve de signal provocateur facilitant la réaction aggressive (Berkowitz, 1977). En outre, Novaco (1978) affirme que

l'évaluation cognitive des événements influence les réponses comportementales, et que ces réponses vont à leur tour modifier les évaluations ultérieures de ces événements. Ces perspectives sont ainsi cohérentes avec plusieurs théories dans le domaine du traitement de l'information sociale, en ce qui a trait au rôle de la cognition dans la colère et l'agressivité.

Une de ces théories est le modèle de traitement de l'information sociale de Dodge (1986), qui a été reformulé par la suite par Crick et Dodge (1994) et qui a servi de fondement théorique aux études les plus récentes sur la cognition sociale et les comportements agressifs. Bien que conçu initialement pour évaluer les processus impliqués dans l'adaptation sociale des enfants, le modèle a également servi dans l'évaluation des différences individuelles dans les comportements agressifs (Chen et al., 2012). Selon le modèle reformulé, pour réagir de manière adaptée à une situation sociale, les informations sociales doivent être traitées en étapes successives, consistant en des activités cognitives spécifiques. Ces étapes sont: (a) la perception, l'encodage et le décodage des stimuli environnementaux; (b) l'interprétation et la représentation des stimuli encodés; (c) la recherche de réponses possibles; (d) la prise de décision sur la réponse la plus appropriée; et (e) la mise en action de la réponse sélectionnée (Epps & Kendall, 1995). Chacune de ces étapes implique des processus cognitifs qui peuvent être inexacts, dysfonctionnels ou déficients, et toute distorsion ou tout déficit de traitement au cours de l'une des étapes ci-mentionnées peut expliquer la manifestation d'un comportement agressif quelconque (Epps & Kendall, 1995). Par exemple, un traitement interprétatif dysfonctionnel ou biaisé (étape b) pourrait potentiellement donner lieu à ce que Milich et Dodge (1984) appellent un biais d'attribution d'intention hostile (BAIH), qui représente la tendance à interpréter l'intention des autres comme hostile lorsque les signaux sociaux sont ambigus, ou n'indiquent pas une intention claire. Par conséquent, une des raisons pourquoi les individus agressifs se comportent de la sorte

est qu'ils attribuent plus fréquemment des intentions hostiles aux actes des autres que les individus non agressifs (Orobio de Castro et al., 2002).

Le BAIH est également associé à la tendance des individus agressifs à afficher des formes d'agressivité réactive, également appelée agressivité impulsive (Dodge & Coie, 1987), soit une réponse défensive et colérique qui se manifeste face à la frustration ou à la provocation (Crick & Dodge, 1996). Par exemple, une étude de Bailey et Ostrov (2008) a montré que la probabilité de présenter un BAIH lors de l'évaluation de vignettes de provocation ambiguës était plus grande chez ceux présentant une tendance à l'agressivité réactive dans des contextes sociaux. De plus, Crick et Dodge (1996) ont observé qu'en réponse à des situations de provocation ambiguës, les enfants agressifs de façon réactive attribuaient plus fréquemment une intention hostile aux pairs provocateurs que les enfants non agressifs. Enfin, plusieurs recherches empiriques ont démontré que seule l'agressivité réactive serait associée au BAIH, alors qu'aucune association n'a été observée avec l'agressivité proactive, vue ici comme une forme d'agressivité délibérée et intentionnelle émise dans le but d'atteindre un objectif souhaité (Crick & Dodge, 1996; Dodge & Coie, 1987).

Selon Dodge (2006), la tendance à développer un style d'attribution hostile s'explique par des différences individuelles dans la capacité à réagir impulsivement, par les expériences de vie traumatiques, et par l'absence d'attachements sécurisés auprès de figures parentales protectrices. Ces facteurs se cumulent et interagissent pour aboutir à des schémas hostiles qui sont stockés en mémoire, et qui donnent lieu à des interprétations dysfonctionnelles des intentions d'autrui (Dodge, 2006).

En somme, le BAIH représente un élément clé expliquant l'étiologie, le développement et la persistance dans le temps des problèmes liés à l'agressivité et à l'antisocialité, et est considéré

comme un facteur de risque important pour les troubles de comportement et les interactions sociales conflictuelles, principalement parce qu'il limite les possibilités d'apprentissage de comportements prosociaux par des moyens non agressifs (Orobio de Castro et al., 2002). Le BAIH est également interprété comme un exemple de « modèle de travail » conceptuel de l'environnement qui est façonné par l'expérience, comme mentionné précédemment, et qui influe sur le fonctionnement social général (Orobio de Castro et al., 2002).

Le BAIH : Traitement automatique et traitement délibéré

De nombreuses études ont montré que le traitement de l'information sociale se fait souvent de manière automatique, inconsciente ou implicite (Bargh, 1984; Shiffrin & Schneider, 1977). Ainsi, pour rendre compte des mécanismes impliqués dans le traitement de l'information sociale, des modèles à double processus ont été conceptualisés à l'aide d'un cadre à deux systèmes qui distinguent le mode automatique de traitement de l'information du mode conscient ou délibéré (Linder et al., 2010). D'abord, le système automatique repose sur un traitement qui est intuitif, associatif, sans effort et inconscient, et dans lequel les informations sont traitées rapidement et en parallèle. De plus, les processus automatiques ne requièrent pas de ressources attentionnelles contrôlées et sont le produit d'appariements stimulus-réponses qui sont fréquents et répétés (Linder et al., 2010). À l'opposé, le système conscient ou délibéré est basé plutôt sur un traitement réfléchi, intentionnel et déductif, et qui nécessite un déploiement d'effort mental. Dans ce système, les informations sont traitées de manière contrôlée et en étapes successives (Glöckner & Betsch, 2008).

Pour faire le lien avec le BAIH, Todorov et Bargh (2002) soulignent que les individus acquièrent des structures de connaissances hostiles inconsciemment développées à la suite d'expositions répétées à des stimuli agressifs, et que ces structures influencent le traitement des

informations sociales ultérieures. De plus, il a été démontré que les manifestations d'agressivité physique impulsives seraient liées à des biais attentionnels et attributionnels précis dans le traitement automatique de l'information (Linder et al., 2010). Par exemple, une étude a démontré que des individus agressifs interprétaient de manière inconsciente et hostile des phrases ambiguës dans des conditions expérimentales spontanées, ou automatiques, tandis que cet effet n'était pas constaté lors d'évaluations nécessitant des interprétations délibérées et contrôlées (Zelli et al., 1995). Selon Linder, le système automatique et le système contrôlé peuvent ainsi être interprétés comme contribuant indépendamment à la prédiction des différences individuelles dans l'étude de l'agressivité (Linder et al., 2010).

Bases conceptuelles de la présente étude

Bien que la dépression soit fréquemment liée à divers biais d'interprétation négatifs qui affectent l'évolution du trouble dépressif, le BAIH n'a jusqu'à présent pas été formellement conceptualisé dans la littérature comme un mécanisme cognitif inadapté par lequel les personnes déprimées qui manifestent de l'hostilité, de l'irritabilité et de la colère, traitent les informations sociales. Or, une étude qui a examiné le lien entre le BAIH et la dépression a démontré que les sujets déprimés interprétaient les scénarios interpersonnels ambigus comme plus hostiles que les sujets non déprimés, et que la dépression prédisait le BAIH indépendamment du trait de colère (Smith et al., 2016). Spécifiquement, les symptômes dépressifs étaient associés de manière unique à la tendance à attribuer une intention hostile aux comportements ambigus d'autrui dans des situations sociales, indépendamment du trait de colère. En se basant sur le modèle de génération de stress de Hammen (1991), décrit ci-bas, Smith et ses collègues ont également suggéré que le biais hostile puisse même être interprété comme un facteur de maintien important de la dépression, principalement en raison de son impact sur le fonctionnement social de l'individu. Le modèle de Hammen met en évidence le lien bidirectionnel entre la dépression et

les conflits interpersonnels et stipule que les individus dépressifs génèrent eux-mêmes un certain sous-type d'événements stressants. Ces situations problématiques, typiquement de nature interpersonnelle et conflictuelle, auraient un effet détériorant sur les relations proximales de l'individu déprimé, exacerbant dans le processus l'état dépressif préexistant (Gadassi & Rafaeli, 2015).

En ce qui a trait à la présente étude, celle-ci a été conçue pour confirmer les observations faites par Smith et ses collègues et a été réalisée au moyen d'une méthodologie électrophysiologique. Par conséquent, une contribution scientifique importante de cette étude réside dans l'utilisation des potentiels liés aux événements (PRE), saisis par enregistrement électroencéphalographique (EEG), pour évaluer le BAIH. En somme, les PRE contournent les mécanismes de contrôle conscients, délibérés et stratégiques qui facilitent ou inhibent le traitement de l'information, en fournissant une mesure précise et continue des processus cognitifs implicites et automatiques impliqués dans les étapes initiales du traitement de l'information sociale (Gagnon et al., 2016).

Plus précisément, les PRE représentent les variations du voltage électrique produit par le système nerveux en réponse à une stimulation sensorielle externe, ou à des stimuli cognitifs internes (Luck, 2014). Les formes d'onde PRE sont composées de séquences de déviations positives et négatives – appelées pics, ondes ou composantes – qui reflètent le flux d'informations circulant à travers le cerveau, tandis que le voltage à chaque moment dans la forme d'onde représente une mesure de l'activité neuronale à un instant précis dans le temps. Enfin, une composante PRE représente l'activité d'un agrégat neuronal qui est fonctionnellement distinct des autres amas neuronaux (Luck, 2014). Parmi les composantes PRE, celle d'intérêt pour ce projet est la N400, dont l'effet fut d'abord observé dans le cadre d'une étude conçue pour

examiner le rôle du contexte d'une phrase dans la reconnaissance de mots lus (Kutas et Hillyard, 1983). Dans cette étude, un mot incompatible avec le contexte de la phrase évoqua un potentiel à déflexion négative possédant une latence maximale d'environ 400 millisecondes, et un pic d'amplitude situé aux sites d'électrodes centro-pariétaux de l'hémisphère droit. Cette composante apparaît donc typiquement, mais pas exclusivement, en réponse à une violation des attentes sémantiques contextuelles du sujet (Gagnon et al., 2016; Kutas & Federmeier, 2011).

Pour mesurer le BAIH, Gagnon et al. (2016) ont développé une tâche PRE qui génère la composante N400 en tant qu'indicateur du biais. Dans cette tâche, rapportée plus en détail ci-dessous, des scénarios sociaux sont présentés au participant sous la forme de phrases décrivant d'abord un contexte social qui est soit hostile ou non hostile, suivi d'un comportement ambigu et enfin, d'une intention hostile ou non hostile qui concorde ou non avec le contexte social hostile ou non hostile initial. De plus, la condition de la tâche est caractérisée comme hostile ou non hostile dépendamment si l'intention est hostile (par exemple, Phrase 1-contexte : Vos amis sont serviables. Phrase 2-comportement : Lors de votre départ, ils vous apportent vos choses. Phrase 3-intention : Vos amis veulent vous précipiter.) ou non hostile (par exemple, Phrase 1-contexte : Vos amis sont serviables. Phrase 2-comportement : Lors de votre départ, ils vous apportent vos choses. Phrase 3-intention -Vos amis veulent vous aider.). Une étude récente utilisant cette tâche a révélé que les participants hostiles produisaient une onde N400 considérablement plus élevée dans la condition non hostile (contexte hostile suivi d'une intention non hostile), par rapport aux participants non hostiles (Gagnon et al., 2016). Ces résultats suggèrent que les personnes hostiles développent des attentes d'intention hostile plus élevées devant un comportement ambigu en contexte hostile que les personnes non hostiles, et que la violation de leurs attentes hostiles se traduit ici par une amplitude plus importante de la composante N400.

Hypothèses

La visée principale de cette étude est d'examiner la relation entre les symptômes dépressifs et le BAIH en utilisant des mesures neurophysiologiques du biais hostile. Ainsi, nous nous attendions d'abord (1) à ce que les scores de dépression corrèlent positivement avec les mesures autodéclarées du BAIH, selon les résultats de Smith. Nous avons également prévu (2) que le groupe hostile présenterait des scores de dépression plus élevés que le groupe non hostile, conformément à la littérature sur l'association entre la dépression et l'hostilité. Ensuite, nous avons prédit (3) que les scores de niveau de dépression seraient liés de façon unique au BAIH, comme avec les résultats de Smith, c'est-à-dire indépendamment des scores d'agressivité réactive, et que cette association existerait à la fois dans les conditions hostiles et non hostiles du paradigme de violation des attentes hostiles. Plus précisément, nous avons anticipé (3a) que les scores de dépression seraient corrélés négativement avec l'effet N400 dans la condition non hostile et discordante (contexte hostile et intention non hostile) par rapport à la condition non hostile concordante (contexte non hostile et intention non hostile), après avoir contrôlé pour l'agressivité réactive. Autrement dit, les participants présentant des niveaux de symptômes dépressifs plus élevés montreraient des niveaux de surprise plus élevés (violation plus forte des attentes hostiles) à la suite d'un mot critique révélant l'intention non hostile derrière le comportement provocateur ambigu du personnage dans un contexte hostile. Nous avons également anticipé (3b) que des scores de dépression plus élevés corrèleraient positivement avec l'effet N400 dans la condition hostile discordante (contexte non hostile et intention hostile) par rapport à la condition hostile concordante (contexte hostile et intention hostile), après avoir contrôlé pour l'agressivité réactive. Ici, si les niveaux de symptômes dépressifs sont liés au biais hostile, indépendamment de l'agressivité réactive, ceux qui présentent des niveaux de dépression

plus élevés devraient montrer moins de surprise (c'est-à-dire une violation plus petite des attentes hostiles comme le montre une amplitude N400 plus petite) après un mot critique révélant l'intention hostile du comportement ambigu émis dans un contexte non hostile.

Méthodologie

Les données utilisées pour ce projet ont été recueillies auprès de 87 participants dans le cadre d'une étude précédente sur l'agressivité et l'impulsivité intitulée « Cognitive and neurophysiological mechanisms associated with the "Urgency" personality trait: theoretical and empirical integration. » (Gagnon & Jolicœur, 2014).

Participants

72 participants ont été retenus pour la présente étude des 87 participants initialement recrutés dans trois populations différentes : i) des étudiants universitaires; (ii) la population générale par le biais d'annonces de journaux, sur Kijiji (un service de petites annonces en ligne) et sur les babillards communautaires; et (iii) des patients d'une clinique pour troubles de la personnalité de la région de Montréal. Des 10 participants retirés de l'étude, quatre appartenaient à la population étudiante et six à la population clinique, car les données sur les niveaux de dépression n'étaient pas disponibles pour ceux-ci. De plus, cinq participants de la population clinique ont été exclus pour éviter un biais de sélection potentiel, car ces sujets étaient plus susceptibles de présenter une symptomatologie dépressive et agressive. En somme, tous les participants cliniques ont été retirés de l'étude. Le groupe expérimental final était ainsi composé de 30 participants masculins (41,6%) et 42 féminins (58,3%) (N = 72, âge moyen = 26,9 ans, ET = 8,1).

Tous les répondants ont passé une entrevue de sélection téléphonique dans laquelle une brève description de l'étude a été fournie, suivie d'une évaluation par questionnaire des critères

d'inclusion et d'exclusion sur la base desquels les participants devaient être sélectionnés. Les personnes répondant aux critères, décrits ci-dessous, ont reçu des détails supplémentaires sur le déroulement de la procédure et ont été informées que l'étude, composée de deux séances expérimentales et d'une séance de sondage par questionnaire en ligne, prendrait environ trois heures pour être complétée. Ils devaient également s'abstenir de consommer de l'alcool 24 heures avant chaque séance expérimentale ou des drogues à usage récréatif une semaine avant chaque séance. Enfin, une compensation de 40 dollars (CAD) a été versée à chaque participant à la fin de la séance expérimentale.

Les répondants étaient exclus s'ils (i) n'avaient pas entre 18 et 55, (ii) n'avaient pas le français comme langue maternelle, (iii) avaient un niveau d'éducation inférieur à la sixième année, (iv) avaient déjà souffert d'une blessure grave à la tête, ou (v) avaient des antécédents de psychose.

Les participants ont été sélectionnés sur la base des scores obtenus à l'aide de la sous-échelle d'hostilité du Brief Symptom Inventory (BSI-II; Derogatis, 1993). Cette sous-échelle, décrite plus bas, a été administrée à chaque répondant lors de l'entretien téléphonique initial, tandis que le questionnaire entier a été complété en ligne via le site web Echo (www.questionnaireenligne.ca). Le BSI est un inventaire de symptômes composé de 53 items divisés en neuf sous-échelles mesurant la somatisation, l'obsession et la compulsion, la sensibilité interpersonnelle, la dépression, l'anxiété, l'hostilité, la phobie, l'idéation paranoïde et les tendances psychotiques. Cette forme abrégée du Symptoms Checklist-90-Revised (SCL-90-R; Derogatis, 1999) fournit également trois indices de détresse globaux : le Global Severity Index (GSI), qui représente le score moyen de tous les items; le Positive Symptoms Distress Index (PSDI), qui correspond au score moyen des items pour lesquels les participants ont

endossé un choix de réponse autre que zéro (non-zero rated items) ; et le Positive Symptom Total (PST), qui est un décompte des items endossés « autre que zéro ». Les participants évaluent les éléments sur une échelle de type Likert en 5 points, allant de « pas du tout » (0) à « extrêmement » (4). Le BSI possède également de bonnes propriétés psychométriques et une bonne validité de construit qui est empiriquement appuyée (Derogatis, 1993).

La sous-échelle de la dimension de l'hostilité comprend cinq items au total (par exemple, item 40 – « Avoir envie de battre ou blesser quelqu'un ») qui évaluent les composantes cognitives, émotionnelles et comportementales qui sont typiquement associées à l'hostilité. Les participants ayant obtenu un score supérieur au score-T standard de 65, selon l'échantillon normatif pour adultes non hospitalisés, ont été assignés au groupe hostile (26 participants; hommes=14, femmes=12), tandis que ceux ayant obtenu un score inférieur à 65 ont été assignés au groupe non-hostile (46 participants; hommes=16, femmes=30). Le coefficient alpha pour cette sous-échelle est de .85 (Derogatis & Melisaratos, 1983).

Matériaux et appareils

Paradigme de la violation des attentes hostiles

Le BAIH a été évalué à l'aide d'une tâche développée par Gagnon et al. (2016) à partir d'une adaptation du paradigme de Leuthold et al. (2012). Le paradigme de la violation des attentes hostile comprend 320 scénarios textuels décrivant des interactions sociales de la vie courante. Chaque scénario est composé de trois phrases : une première phrase dans laquelle un contexte social hostile ou non hostile est établi (phrase 1 - contexte); une deuxième phrase décrivant le comportement ambigu d'un personnage fictif (phrase 2 - comportement); et enfin, une troisième phrase dans laquelle le dernier mot (le mot cible) informe le lecteur de l'intention guidant le comportement ambigu du personnage (phrase 3 - intention). Chaque scénario possède

également une version partiellement identique qui diffère uniquement dans son contexte (1). Ainsi, le contexte sera hostile dans un scénario et non-hostile dans l'autre. De plus, la condition de la tâche est hostile ou non hostile dépendamment si l'intention (3) décrite dans la dernière phrase est hostile ou non, comme mentionné plus haut, tandis que la nature du contexte (1) du scénario (hostile ou non hostile) peut être concordant ou discordant avec la nature de l'intention (3) (hostile ou non hostile).

Toutes les phrases ont été composées de manière concise avec pas plus de 25 mots au total pour les deux phrases de contexte (1), et huit mots pour la phrase cible finale (3). Les phrases ont également été composées, à proportions égales, de manière affirmative (par exemple, « l'étranger souhaite vous provoquer. ») et négative (par exemple, « votre collègue ne veut pas vous aider. »). Ensuite, les 320 scénarios ont été divisés en deux ensembles de 160 scénarios chacun. Pour chaque ensemble, chaque scénario de chaque condition (hostile et non hostile) apparaît soit dans une forme concordante (par exemple, contexte et intention hostiles ou non hostiles), soit dans une forme discordante (par exemple, contexte hostile et intention non hostile ou vice versa) dans un ensemble donné, tout en apparaissant dans les deux formes de présentation (concordant et discordant) entre les deux ensembles (320 scénarios). Pour un ensemble (160) de scénarios donné, les conditions hostiles et non hostiles comprennent 80 items chacun, pour un total de 160 scénarios, et dans lequel 40 contextes hostiles (ou non hostiles) sont concordants avec la nature de l'intention, et 40 contextes hostiles (ou non hostiles) sont discordants avec la nature de l'intention. Finalement, un ensemble de 160 scénarios a été présenté à une première moitié du groupe expérimental, tandis que l'autre ensemble a été présenté à la seconde moitié.

Mesures auto-rapportées

Sous-échelle de dépression du Brief Symptoms Inventory (BSI-II; Derogatis, 1993). La sous-échelle de la dépression du BSI est une mesure des niveaux de symptômes cliniques de la dépression selon plusieurs dimensions de la dépression clinique, et est composée de six items au total (par exemple, item 18 – « Je ne ressente aucun intérêt pour les choses ») qui évaluent le degré auquel les répondants ressentent une humeur dépressive, une perte d'intérêt, une vulnérabilité aux critiques, la solitude, l'inutilité, et le désespoir, et ont des pensées suicidaires, avec des scores plus élevés représentant une symptomatologie plus sévère. Le coefficient alpha pour cette sous-échelle est de .85 (Derogatis & Melisaratos, 1983).

Reactive-Proactive Aggression Questionnaire (RPQ; Raine et al., 2006). Le RPQ est un questionnaire divisé en deux sous-échelles qui mesure à la fois l'agressivité réactive (11 items - par exemple, « combien de fois avez-vous réagi avec colère lorsque provoqué par autrui? ») et l'agressivité proactive (12 items - par exemple, « combien de fois avez-vous menacé ou intimidé quelqu'un? »). Les participants doivent évaluer leurs réactions à des scénarios agressifs fictifs en utilisant une échelle de type Likert à trois points allant de « jamais » (0) à « souvent » (2). Ici, le coefficient alpha est de .84 pour la sous-échelle réactive et .86 pour la sous-échelle proactive (Raine et al., 2006). Pour la présente étude, seule la sous-échelle d'agressivité réactive de la version adaptée en français (Gagnon et Rochat, 2017) a été utilisée, sur la base de données de recherche indiquant que seule l'agressivité réactive est associée au BAIH, comme mentionné plus haut.

Social Information Processing–Attribution and Emotional Response Questionnaire (SIP–AEQ; Coccaro et al., 2009). Comme mesure auto-rapportée du BAIH, la version adaptée en français (Gagnon et al., 2015) de ce questionnaire contient huit vignettes qui décrivent des

scénarios socialement ambigus dans lesquels une action aversive physique (par exemple, être frappé par quelqu'un) ou relationnelle (par exemple, être rejeté par quelqu'un) est dirigée envers un personnage fictif avec lequel le participant doit s'identifier. Le participant doit ensuite déterminer la probabilité du comportement ambigu en utilisant une échelle de type Likert à quatre points allant de « peu probable » (0) à « très probable » (3). Les items mesurent l'intention hostile directe (par exemple, « Cette personne voulait me blesser physiquement »), l'intention hostile indirecte (par exemple, « Cette personne voulait me faire mal paraître »), l'intention instrumentale non hostile (par exemple, « Cette personne voulait gagner le match »), et l'intention neutre ou bénigne (par exemple, « Cette personne a fait cela par accident »). Ces items sont ensuite suivis de deux autres items, cotés aussi sur la même échelle de Likert, qui évaluent les réponses émotionnelles négatives (par exemple, « Quelle est la probabilité que vous soyez en colère si cela vous arrivait? » et « Quelle est la probabilité que vous soyez contrarié(e) par vous-même si cela vous arrivait? »). Des scores globaux élevés indiquent une plus grande tendance à attribuer une intention selon les quatre types d'attributions. Les coefficients alpha pour les échelles d'intention hostile, instrumentale et bénigne sont de .82, .57 et .66, respectivement (Coccaro et al., 2009). De plus, la moyenne des scores combinés pour les formes d'intention hostile directe et indirecte a été utilisée pour les buts de cette étude.

Procédure

Avant de donner leur consentement verbal et écrit, les participants ont été informés de la procédure EEG et de la tâche expérimentale. On leur a demandé de lire attentivement toutes les consignes et d'éviter de faire des mouvements inutiles pendant la phase expérimentale EEG. Les participants ont ensuite été informés que la tâche expérimentale consistait à lire des scénarios décrivant des interactions sociales de la vie courante avec des personnages fictifs, tout en

imaginant qu'ils participaient eux-mêmes à de telles interactions. Ils ont également été invités à imaginer les pensées et les sentiments des personnages, et à inférer l'intention derrière le comportement ambigu du personnage provocateur.

Le déroulement des essais expérimentaux s'est fait dans une cage de Faraday dans laquelle le niveau de lumière ambiante était maintenu bas. Les stimuli textuels ont été présentés au centre d'un écran d'ordinateur de 17 pouces, en caractères Helvetica blancs à 14 points sur fond noir. Ensuite, une distance de visualisation des stimuli de 57 cm a été maintenue tout au long de la séance expérimentale à l'aide d'une mentonnière. Enfin, trois caractères sous-tendaient 1 ° d'angle visuel.

Quatre essais pratiques ont été suivis de dix blocs expérimentaux composés de 17 essais chacun (16 scénarios expérimentaux et une phrase de contrôle). Les blocs expérimentaux ont été divisés par une pause chronométrée dont la durée fut déterminée par le participant. Dans chaque essai, les deux premières phrases du scénario – la phrase contextuelle et la phrase décrivant le comportement ambigu – ont été présentées et affichées jusqu'à ce que le participant appuie sur la barre d'espace pour indiquer que la phrase a été lue, mais seulement après un temps d'affichage minimum de 1 500ms. Un intervalle sans stimulus a ensuite été montré pendant 500ms, suivi d'une croix de fixation présentée au centre de l'écran pendant 1 000ms. Ensuite, la phrase révélant l'intention a été présentée au centre de l'écran un mot à la fois pendant 300ms, suivi d'un intervalle sans stimulus de 200ms entre les mots. La prochaine phrase contextuelle a ensuite été présentée suivant un intervalle de 2 000ms au cours duquel le participant a été invité à garder son regard sur une croix de fixation. Pour s'assurer que les participants portent attention aux stimuli présentés, une phrase contrôle placée de manière aléatoire a été incluse parmi les 17 essais de chaque bloc, suivi d'une question de compréhension de type vrai ou faux. Le taux

moyen de réponses correctes pour les questions de compréhension était de 88%, indiquant que la plupart des participants portaient effectivement attention aux stimuli, tandis que ceux ayant un taux de réponses correctes inférieur à 60% étaient retirés de l'étude. Enfin, les participants ont été invités à compléter les versions informatisées du RPQ et du SIP-AEQ après la procédure EEG.

Enregistrement EEG

À l'aide du système Biosemi Active Two, le signal d'électroencéphalogramme (EEG) a été enregistré à partir de 64 électrodes Ag / AgCl montées sur un bonnet élastique conformément au système international 10-10 (Sharbrough et al., 1991), et référencé à la moyenne des mastoïdes gauche et droite. Les mouvements oculaires horizontaux ont été mesurés à l'aide d'un électro-oculogramme horizontal (HEOG) – enregistré comme la différence de voltage entre les électrodes placées latéralement par rapport au canthus externe – tandis que les clignements des yeux étaient détectés avec un électro-oculogramme vertical (VEOG), enregistré ici comme la différence entre le voltage des deux électrodes placées au-dessus et au-dessous de l'œil gauche. Au cours de l'enregistrement du signal, les EEG et EOG ont été numérisés à 512Hz et filtrés passe-bas à 134Hz, tandis que le signal continu a été filtré passe-haut à 0,01Hz et passe-bas à 30Hz lors des analyses de signal hors ligne. La fenêtre temporelle choisie pour segmenter l'EEG allait de 200ms avant l'apparition du mot cible (0ms) jusqu'à 1000ms après l'apparition de la cible. De plus, la correction référentielle (baseline correction) de l'EEG a été effectuée avec la moyenne du signal de -200ms à 0ms. De plus, la réduction des artefacts électro-oculaires et la correction des données ont été obtenues avec des procédures statistiques avancées telles qu'une analyse à composante indépendante (ICA – Independent component analysis) (voir Drisdelle, Aubin & Jolicœur, 2016). Les essais contenant des clignements de yeux (VEOG > 50mV dans un

intervalle de 150ms) et de grands mouvements oculaires horizontaux (HEOG > 35mV dans un intervalle de 300ms) ont été exclus de l'analyse, tandis que les essais montrant des déviations de l'EEG supérieures à 100mV pendant la fenêtre de segmentation prédéterminée sur une des 64 électrodes ou plus ont été sélectionnés pour une analyse supplémentaire. Si sept canaux ou moins démontrant une telle déviation ont été détectés au cours d'un essai donné, une interpolation sphérique par splines a été utilisée pour interpoler ces canaux à partir des canaux voisins, alors que les essais ont été rejetés s'il s'avérait que plus de sept canaux comportaient des artéfacts.

Analyse de données EEG

L'analyse des données EEG a été réalisée avec MATLAB et avec les fonctions de EEGLAB (Delorme & Makeig, 2004) et ERPLAB toolboxes (Lopez-Calderon & Luck, 2014). Le signal EEG au niveau de chaque site d'électrode a été moyenné séparément pour les quatre conditions expérimentales (hostile-concordant, hostile-discordant, non hostile-concordant, non hostile-discordant) et a été ancré temporellement (time-locked) pour viser le début de la présentation du mot cible. Après une première évaluation visuelle du signal global, une déviation négative a été identifiée entre 450ms et 650ms, ce qui correspond à l'intervalle de latence de l'effet N400 identifié dans la littérature (Kutas & Federmeier, 2011). Dans l'étude de Gagnon et al. (2017), les données d'amplitude des électrodes médianes ont été analysées séparément des données enregistrées des sites d'électrodes latérales, ces dernières étant regroupées en six régions identifiées par l'hémisphère gauche ou droit, et par les emplacements antérieurs, centraux ou postérieurs du cuir chevelu. L'hémisphère gauche était divisé en trois régions identifiées comme suit: antérieur gauche (AF3, AF7, F1, F3, F5, F7, FT7, FC1, FC3, FC5), centre gauche (TP7, T7, C1, C3, C5, CP1, CP3, CP5) et postérieur gauche (P1, P3, P5, P7, PO3, PO7, O1). L'hémisphère droit a été divisé en trois régions équivalentes identifiées par leurs électrodes hémisphériques

homologues: antérieur droit (AF4, AF8, F2, F4, F6, F8, FT8, FC2, FC4, FC6), centre droit (TP8, T8, C2, C4, C6, CP2, CP4, CP6) et postérieur droit (P2, P4, P6, P8, PO4, PO8, O2). Ils ont ensuite effectué des analyses statistiques sur les électrodes latérales au moyen d'une ANOVA à mesures répétées avec correction Huynh-Feldt avec les variables Condition (hostile, non hostile), Consistance (concordant, discordant), Hémisphère (gauche, droite), Localisation (antérieur, central, postérieur), et Groupe (hostile, non hostile), tandis que les électrodes médianes ont été soumises à une ANOVA avec les variables Condition, Consistance, Groupe et Localisation (antérieur: AFZ, Fz, FCz; central: Cz, CPz; postérieur: Pz, POz, Oz). Les résultats ont montré la présence d'une N400 suivant la présentation d'une intention non hostile, comme le révèle le mot cible, alors que le participant s'attendait à une intention hostile. Ici, l'amplitude de la composante N400 s'est avérée plus grande pour la condition non hostile que pour la condition hostile, pour le groupe hostile par rapport au groupe non hostile, ainsi que sur les sites d'électrodes centrales et postérieures par rapport aux sites antérieurs.

Analyses de la présente étude

Pour la présente étude, une analyse comparative (test t) a d'abord été effectuée pour comparer les groupes hostiles et non hostiles sur la sous-échelle de dépression du BSI, suivie d'une deuxième analyse comparant les scores d'agressivité réactive pour les deux groupes. Cette deuxième analyse comparative a été effectuée dans le but de confirmer les différences significatives observées entre les groupes dans l'étude de Gagnon et al. (2017), mais cette fois avec un plus petit échantillon dans lequel tous les participants cliniques, et potentiellement plus agressifs, ont été retirés. De plus, des analyses corrélationnelles ont été menées entre les variables d'intérêt dans le but d'identifier les associations significatives pertinentes.

Ensuite, nous avons testé l'hypothèse selon laquelle la dépression prédisait l'effet N400 indépendamment de l'agressivité réactive en effectuant d'abord une ANOVA à mesures répétées avec les variables Condition (hostile, non hostile), Consistance (concordant, discordant), Hémisphère (gauche, droite) et Localisation (antérieur, central, postérieur), et avec les scores de dépression comme covariable et les scores d'agressivité réactive comme variable de contrôle. Il est nécessaire de préciser que, comparativement à l'ANOVA effectuée dans l'étude de Gagnon et al. (2017) décrite ci-haut, la variable Groupe (hostile vs non hostile) n'a pas été retenue dans notre ANOVA étant donné que le but principal de notre étude était de prédire un indice du BAIH via la N400 à partir du score de la dépression dans l'échantillon total (participants hostiles et non hostiles). De plus, seules les électrodes latérales (et non pas médianes) ont été retenues pour l'ANOVA étant donné que celles-ci étaient suffisantes pour répondre au besoin de la sélection d'une région cérébrale indiquant la présence d'une N400. Une décomposition de l'analyse de variance a ensuite été effectuée pour isoler les interactions spécifiques impliquant les facteurs Condition, Consistance, Hémisphère, Localisation et la covariable dépression. L'analyse complète est décrite plus en détail dans l'article scientifique qui accompagne ce manuscrit.

Finalement, les analyses de régression nous ont permis de tester notre hypothèse principale selon laquelle les scores de dépression étaient liés au biais d'attribution hostile, après avoir contrôlé pour l'agressivité réactive, et d'examiner les effets différentiels des conditions hostiles et non hostiles de la tâche PRE sur l'amplitude de la composante N400. Ici, l'amplitude N400 a été mesurée comme la différence entre la forme (condition) discordante et la forme (condition) concordante, et cela pour les conditions hostile et non hostile. De plus, les régions du cuir chevelu de l'effet N400 ont été sélectionnées pour les analyses de régression sur la base des résultats de l'ANOVA à mesures répétées démontrant une interaction significative entre la

Condition, la Consistance, l'Hémisphère, la Localisation et la covariable dépression. Ainsi, nous nous attendions à ce que la région d'intérêt soit située principalement aux sites d'électrodes de la région centrale de l'hémisphère droit, à la lumière de la littérature de recherche suggérant que l'effet N440 est plus important sur les sites centropariétaux de cet hémisphère (Kutas et Federmeier, 2011).

Chapitre II : Article scientifique

Association between depression and hostile attribution bias in hostile and non-hostile individuals: an ERP study.

Authors: Alexandre Gasse^{1,2,3}, Wan Seo Kim^{1,2,3}, Jean Gagnon^{1,2,3}

1 Department of Psychology, University of Montreal, Canada

2 Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal, Canada

3 Laboratoire d'électrophysiologie en neuroscience sociale (LENS), University of Montreal, Canada

Corresponding author:

Alexandre Gasse

E-mail address: alexandre.gasse@umontreal.ca

Abstract

Background: Hostility and aggression have been found to be highly prevalent among depressed patients and are associated with higher comorbidity and illness severity levels. Although negative interpretation biases are a fundamental element of cognitive models of depression, few studies have examined the specific biases in information processing, mainly the hostile attribution bias, found in hostile individuals who present depressive symptoms. Method: Using pre-collected data from a sample of 72 (male=41,6%, female=58,3%) undergraduate and community-based hostile (n=26) and non-hostile (n=46) adult participants, the authors aimed to examine the association between depression and the hostile attribution bias by determining whether depression level scores were uniquely related to electrophysiological measures of the hostile attribution bias. Results: The hostile group showed higher measured levels of depression and reactive aggression, compared to the non-hostile group. Also, depression scores were significant predictors of the N400 effect in the non-hostile condition, while reactive aggression was not, whereas in the hostile condition, the overall model was significant, with depression and reactive aggression levels both showing strong trends towards significance. Limitations: A small sample size limited the scope of our conclusions. Also, sample selection prevented the authors from examining specific group differences regarding the hostile attribution bias in depressed and non-depressed groups. Conclusion: This study highlights the importance of better understanding the perception mechanisms that are associated with depression and to consider the application of cognitive restructuring techniques to counter biased interpretation processes in settings where depression and aggression intersect. This study also suggests that alternatives to self-evaluative methodologies might be more useful when assessing automatic and implicit processes.

Key words: Depression; Aggression; Hostile attribution bias; EEG; ERP; N400.

Introduction

Literature Review

Aggression and Depression

Unipolar depression is a heterogeneous disease composed of several clinical subtypes, each presenting its own symptomatology and distinct long-term prognosis. Researchers thus suggest that identifying explicit phenotypes linked to greater prognosis severity may positively impact overall disease incidence, by guiding the development of treatment strategies better adapted to specific subtypes of the disorder (Pedrelli et al., 2013). One such phenotype, characterized by high prevalence levels of irritability, hostility, and anger, has been identified in multiple cross-sectional studies using an array of conceptual definitions and measurement tools (Judd et al., 2013). For example, 95 of the 224 participants in a Benazzi and Akiskal (2005) study who experienced major depressive disorder, expressed high levels of irritability and hostility, while 37% of patients in a study by Snaith and Taylor (1985) showed signs of moderate to severe external irritability. In addition, Fava (1982) and Riley (1989) both observed high levels of anger and hostility in depressed patients compared to the control group, while Pitt and colleagues (1968) noted that new mothers with postpartum depression were often excessively angry.

Although irritability and anger are no longer central to the diagnosis of major depressive disorder in adults since the DSM-III (American Psychiatric Association, 1987) was revised, these symptoms are highly prevalent, and those who present them suffer from a more chronic, comorbid and severe form of the disorder (Judd et al., 2013). This form of 'irritable-hostile' depression is also associated with earlier mood disorder onset, younger age, atypical depressive features, hypomanic symptoms, and a family history of bipolar disorder (Benazzi & Akiskal, 2005, in Judd et al., 2013). Thus, the importance of identifying specific depression phenotypes

linked to severe prognoses has led some researchers to recommend that irritability and anger be recognized as key indicators of a major clinical subtype of depressive disorder (Judd et al., 2013).

Interpretative Biases in Depression

Not only is depression a highly prevalent psychiatric disorder – depressive disorders affect 4,7% of Canadians annually and represent 10,9% of all psychiatric hospitalizations (Canadian Community Health Survey on mental health, Statistics Canada, 2012) – it is also an extremely recurrent one whose incidence may be explained in part by stable vulnerability factors that expose certain people to repeated episodes (Gotlib et Krasnoperova, 1998; Dearing et al., 2009); it is said that 80% of Americans who have experienced a depressive episode will experience several such episodes during their lifetime (Kessler et al. 2003; Boland and Keller 2002). Consequently, considerable research efforts, guided in large part by cognitive theory, have been allocated to identifying the potential risk factors associated with the long-term incidence of the disorder (Gotlib and Krasnoperova, 1998).

Beck (1967, 1976, in Gotlib et Krasnoperova, 1998) was the first to emphasize the importance of maladaptive cognitive processes as risk factors for depression and formulated a theory of depression that largely attributed its appearance to cognitive biases in the form of negative schemas through which the environment is filtered, emphasizing that maladaptive cognitive processes are an important risk factor for depression. As such, depressive people selectively focus on negative stimuli, filter out positive stimuli and perceive negative or neutral information in a more negative way (Gotlib et Krasnoperova, 1998; Dearing et al., 2009). Furthermore, depressed individuals are more likely to interpret ambiguous cues and stimuli, such as homophones and sentences presented in an experimental setting, as threatening, increasing the

likelihood that those predisposed to mistrust and threat perception will experience more hostility, and consequently, act more aggressively towards others (Mogg et al., 2006; Moreno et al., 1994). Researchers therefore suggest that major depressive disorder with irritability may promote a hostile worldview and increased interpersonal hostility based in part on biased cognitive processes (Riley et al., 1989; Smith et al., 2016).

The Hostile Attribution Bias

According to earlier cognitive personality theorists, the cognitive appraisal of a cue is what determines its ability to evoke a response, not the cue per se (Berkowitz, 1977; Nasby et al., 1980). Consequently, aversive stimuli do not automatically generate aggressive outcomes, but rather, cognitive processes such as appraisals, expectations and private speech, play a mediating role in the relationship between aversive stimuli and the subsequent aggressive responses. For instance, for an aggressive reaction to occur, one must attribute an aggressive or hostile intent to an aversive stimulus for it to serve as a provocative cue (Berkowitz, 1977; Nasby et al., 1980).

As a result, in addition to negative or depressive interpretation biases, distorted interpretative processing may also lead to what Milich and Dodge (1984) call a hostile attributional bias (HAB), a key element in the etiology, development, and persistence of aggressive and antisocial behavior issues over time (Chen et al., 2012). Safer (2009) defines hostility as a directed and maintained attitude of ill-will, while the hostile attribution bias represents the tendency of aggressive individuals to interpret others' intent as hostile when social cues are ambiguous or fail to indicate a clear intent, and is generally displayed through reactive forms of aggression, which is an angry, defensive response to frustration, provocation or perception of threat (Bailey & Ostrov, 2008). Research has demonstrated that only reactive aggression is associated with HAB, whereas proactive aggression, defined as a deliberate

aggressive behavior with the aim of obtaining a desired goal, is not (Crick & Dodge, 1996; Dodge & Coie, 1987, in Wilkowski et al., 2008).

According to Dodge (2006), the propensity to develop a hostile attribution style can be explained by individual differences regarding impulsivity, by the experience of traumatic life events, and by the absence of a secure attachment to protective parental figures. The additive and interactive potential of these factors produce hostile patterns that are stored in memory, and which give rise to dysfunctional interpretations of others' intentions (Dodge, 2006).

HAB is largely accepted as a fundamental theoretical concept and is a key element in the etiology, development and persistence of aggressive and antisocial behavior issues over time. It is considered a significant risk factor for aggressive behavior problems and conflictual social interactions, mainly because it limits the opportunities to learn prosocial behaviors through nonaggressive means (Orobio de Castro et al., 2002).

HAB: Automatic VS deliberate processing

Many studies have shown that the processing of social information is often automatic, unconscious or implicit (Bargh, 1984; Shiffrin & Schneider, 1977). To account for and explain the mechanisms involved in social information processing, dual-process models have been conceptualized using a two-system framework that distinguishes the automatic mode of processing from the conscious or deliberate mode (Linder et al., 2010). First, the automatic system is based on processing which is intuitive, associative, effortless and unconscious, and in which information is processed quickly and in parallel. In addition, automatic processes do not require controlled attentional resources, and stem from frequent and repeated stimulus-response pairings (Linder et al., 2010). The conscious system, on the other hand, is based on thoughtful,

deliberate and deductive processing that requires mental effort, and in which information is processed in a controlled manner through successive steps (Glöckner & Betsch, 2008).

Todorov and Bargh (2002) argue that individuals acquire unconsciously developed hostile cognitive structures as a result of repeated exposure to aggressive stimuli, and that these structures influence the processing of subsequent social information. In addition, behavioral or physical aggression have been shown to be significantly linked to specific attentional and attributional biases in automatic information processing (Linder et al., 2010). For example, one study has shown that aggressive individuals unconsciously and hostilely interpret ambiguous sentences under spontaneous or automatic experimental conditions, while this effect was not seen in appraisals requiring deliberate and controlled interpretations (Zelli et al., 1995).

According to Linder (2010), automatic and controlled (or deliberate) processing can thus be seen as contributing independently to the prediction of individual differences in the study of aggression.

Theoretical Basis for the Current Study

Although depression is frequently linked to various negative interpretative biases that impact the course of the disorder, HAB has thus far not been formally conceptualized in the literature as a maladaptive cognitive mechanism by which depressed individuals who manifest hostility, irritability and anger, process social information. However, one study that did examine the link between HAB and depression showed that depressed subjects interpreted ambiguous interpersonal scenarios as hostile more frequently than did non-depressed subjects, and that depression predicted HAB independently of trait anger (Smith et al., 2016).

The current study was designed to confirm the observations made by Smith and her colleagues using a unique cognitive neuroscience approach. Therefore, an important scientific

contribution of this study lies in our use of electroencephalographically (EEG) recorded event-related potentials (ERP) to assess HAB.

ERPs effectively circumvent conscious, deliberate and strategic control mechanisms which facilitate or inhibit information processing, by providing a precise and continuous measure of the automatic and implicit cognitive processes involved in the initial steps of social information processing, such as the interpretation of information (Gagnon et al., 2016). More precisely, ERPs represent variations in the electrical voltage produced by the nervous system in response to external sensory stimulation, or internal cognitive stimuli (Luck, 2014). ERP waveforms consist of sequences of positive and negative deflections – called peaks, waves or components – which reflect the flow of information through the brain, whereas the voltage at each point in the waveform represents a measure of neural activity at a specific point in time. Finally, an ERP component represents the activity of a neural aggregate which is said to be functionally distinct from other neural clusters (Luck, 2014). The ERP component used for our study is the N400, a negative deflection component which has a peak latency of approximately 400 milliseconds in the centroparietal regions of the scalp, and appears typically, but not exclusively, in response to a violation of the subject's contextual semantic expectations (Gagnon et al., 2016; Kutas & Federmeier, 2011).

To measure HAB, Gagnon et al. (2016) developed the Hostile Expectancy Violation Paradigm, an ERP task comprising of social scenarios made up of three sentences that elicits the N400 component as an indicator of the hostile bias (see Table 1). The task is described in full below.

Study objective and hypotheses

The general objective of this study is to examine the relation between depressive symptoms and HAB using electrophysiological measures of HAB. First, we anticipated (1) that depression level scores would positively correlate with HAB self-reported measures (SIP-AEQ, described in full below), per Smith's findings. We also anticipated (2) that the hostile group would present higher depression level scores than the non-hostile group, in line with the literature on the association between depression and hostility. Next, we predicted (3) that depression level scores would be uniquely related to HAB, and therefore independent of reactive aggression scores, and that this association would exist in both the hostile and the non-hostile conditions of the hostile expectancy violation paradigm. More precisely, we hypothesized (3a) that depression level scores would correlate negatively with the N400 effect in the non-hostile mismatch condition (hostile context and non-hostile intention), after adjusting for reactive aggression. That is, participants presenting greater depressive symptom levels would show greater levels of surprise (stronger violation of hostile expectancy) following a critical word revealing the non-hostile intention behind the character's ambiguous provocative behavior in a hostile context. We also hypothesized (3b) that higher depression level scores would correlate positively with the N400 effect in the hostile mismatch condition, after adjusting for reactive aggression. Here, if depressive symptom levels are related to the hostile bias, independently of reactive aggression, then those presenting greater depression levels should show less surprise (i.e. smaller violation of hostile expectation as shown by a smaller N400 amplitude) following a critical word revealing the hostile intention of the ambiguous behavior expressed in a non-hostile context.

Method

The data used for this project was gathered as part of a prior study on aggression and impulsivity entitled “Cognitive and neurophysiological mechanisms associated with the "Urgency" personality trait: theoretical and empirical integration.” (Gagnon & Jolicoeur, 2014).

Participants

87 Participants were initially recruited from three different populations: (i) undergraduate university students; (ii) the general population using newspaper, Kijiji (an online classified advertising service) and community bulletin ads; and (iii) patients from a Montreal-area personality disorders clinic. All respondents completed a telephone screening interview and were excluded if they (i) were not between 18 and 55 years old, (ii) did not have French as their mother tongue, (iii) had less than a sixth-grade level of education, (iv) reported to have suffered from a serious head injury, or (v) had a history of psychosis. Those meeting the selection criteria were informed that the study would take approximately three hours to complete, for which they would receive a 40-dollar (CAD) compensation. Respondents were also directed to refrain from consuming alcohol 24h before each session and recreational drugs one week before each session.

For participant selection purposes, the Brief Symptoms Inventory (BSI-II; Derogatis, 1993) Hostility Dimension subscale was administered to each potential participant during the initial telephone interview, while the entire inventory (described below) was completed online via the Echo online questionnaire website (www.questionnaireenligne.ca). The Hostility Dimension subscale is comprised of five items in total (e.g. item 40 - Having urges to beat, injure, or harm someone) that assess the cognitive, emotional and behavioral components associated with hostility, using a 5-point Likert scale ranging from “not at all” (0) to “extremely” (4). Respondents scoring above the standard T-score of 65 in the adult nonpatient normative sample

were assigned to the hostile group (26 participants; male=14, female=12), whereas those scoring below this T-score were assigned to the non-hostile group (46 participants; male=16, female=30). This cut-off was selected according to the clinical marker guidelines found in the BSI user manual, along with the T-scores which are derived from comparing a given raw score to the normative scores calculated for a specific population (Stewart et al., 2010). The Cronbach's coefficient alpha for this subscale is .81 (Derogatis & Melisaratos, 1983)

For the current study, the final experimental group comprised of 30 male (41,6%) and 42 female (58,3%) participants (N=72, mean age=26,9 years, SD=8,1). Of the 10 participants that were removed from the study, four were from the student population and six were from the clinical population, because their depression level scores were not available. Also, the remaining five clinical participants were excluded to avoid a potential selection bias, as these subjects were more likely to present depressive symptomatology and show aggressive tendencies.

Materials and Apparatus

Hostile Expectancy Violation Paradigm

HAB was assessed using the Hostile Expectancy Violation paradigm developed by Gagnon et al. (2016) and consists of 320 scenarios depicting everyday social interactions. Each scenario is comprised of three sentences: an initial sentence establishing a hostile or non-hostile social context (1); followed by a second sentence describing a fictitious character's ambiguous behavior (2); and a third sentence in which a final target word informs the reader of the character's hostile or non-hostile intention (3) behind the ambiguous behavior (2). In this task, each scenario has a partially identical version that shares the same ambiguous behavior (2) sentence and the same intention (3) revealing sentence, but differs only in regards to the initial context (1) sentence, with one version depicting a hostile context, and the second, a non-hostile

context. Also, the nature of the context (1) (hostile or non-hostile) will either match or mismatch the nature of the character’s intention (3) (hostile or non-hostile). Finally, the experimental task condition (hostile or non-hostile) is dependent on whether the intention (3) is hostile or non-hostile (See Table 1 for a detailed description of the task format).

Table 1

<i>Examples of Four Types of Possible Scenario Sentences</i>			
First sentence (social context)	Second sentence (ambiguous behavior)	Third sentence (intention)	Condition
<u>Hostile</u> Your neighbor is vengeful.	He plays music all night.	<u>Hostile</u> He wants to <i>irritate</i> you. ^a	Hostile match
<u>Non-hostile</u> Your neighbor likes to organize parties.	He plays music all night.	<u>Hostile</u> He wants to <i>irritate</i> you.	Hostile mismatch
<u>Hostile</u> Before the exam, the students are in competition.	A friend walks by and does not talk to you.	<u>Non-hostile</u> He does not want to <i>distract</i> you. ^b	Non-hostile mismatch
<u>Non-hostile</u> Before the exam, the students study seriously.	A friend walks by and does not talk to you.	<u>Non-hostile</u> He does not want to <i>distract</i> you.	Non-hostile match

Translation into English of: ^a «Il veut vous *irriter*» and ^b «Il ne veut pas vous vous *distraindre*». In French, the pronoun "you" precedes the verb that describes the character's intention allowing for the target sentence to finish with the critical word.

Two sets of 160 scenarios were created so that each item of each condition (hostile or non-hostile) appeared in either the match (ex. hostile context and hostile intention, or vice versa) or mismatch (ex. hostile context and non-hostile intention, or vice versa) format in a given set, while appearing in both formats across both sets. For each set of scenarios, the hostile and non-hostile conditions consisted of 80 items each – for a total of 160 scenarios – which were then separated into 40 hostile contexts (and 40 non-hostile contexts) that match the nature of the

intention, and 40 hostile contexts (and 40 non-hostile contexts) that mismatch the nature of the intention. Finally, one set of scenarios was presented to one half of the experimental group, while the other set was presented to the other half.

Self-Report Measures

Brief Symptom Inventory Depression Dimension subscale (BSI-II; Derogatis, 1993). The BSI Depression subscale is a measure of depression symptom levels along multiple dimensions of clinical depression. It consists of a total of six items (e.g., item 18 - Feeling no interest in things) that rate the degree to which respondents experience depressive mood, loss of interest, vulnerability to criticism, loneliness, worthlessness, hopelessness, and thoughts of suicide, with higher scores representing more severe symptomatology. The alpha coefficient for this subscale is .85 (Derogatis & Melisaratos, 1983).

Reactive-Proactive Aggression Questionnaire (RPQ; Raine et al., 2006). The RPQ provides a measure of both reactive aggression (11 items - e.g. How often have you reacted angrily when provoked by others?) and proactive aggression (12 items - e.g. How often have you threatened and bullied someone?), and respondents must rate their reactions to fictional aggressive scenarios using a three-point Likert scale ranging from “never” (0) to “often” (2). The alpha coefficient for the reactive subscale is .84, and .86 for the proactive subscale (Raine et al., 2006). For the current study, only the reactive aggressive subscale of the French adapted version (Gagnon & Rochat, 2017) was used, based on research evidence indicating that only reactive aggression is associated with HAB, as mentioned above.

Social Information Processing—Attribution and Emotional Response Questionnaire (SIP—AEQ; Coccaro et al., 2009). As a self-report measure of HAB, this questionnaire contains eight vignettes that depict socially ambiguous scenarios in which a

physically (e.g., being struck by someone) or relationally (e.g., being rejected by someone) aversive action is directed towards a fictional character that the participant is asked to identify with. The participant must determine the likelihood of the ambiguous behavior using a four-point Likert scale ranging from “not at all likely” (0) to “very likely” (3). These items are then followed by two other Likert-scaled items that assess negative emotional responses. The SIP—AEQ measures direct hostile intent, indirect hostile intent, instrumental non-hostile intent, and neutral or benign intent. For this study, the average of the combined scores for the direct and indirect hostile intent forms of the French adapted version (Gagnon et al., 2015) of this questionnaire were used. Also, alpha coefficients for the hostile, instrumental, and benign intent scales are .82, .57, and .66, respectfully (Coccaro et al., 2009).

Procedure

Before giving their verbal and written consent, participants received information about the EEG procedure and the experimental task. They were instructed to read all materials carefully and to avoid making unnecessary movements during the EEG testing phase. Participants were told that the experimental task consisted in them reading scenarios depicting everyday social encounters with fictional characters while imagining that they were themselves taking part in such encounters. They were also instructed to imagine the characters’ thoughts and feelings, and the intent behind the characters’ ambiguous behaviors.

Participant testing was performed in an electrically shielded booth (Faraday cage) in which the ambient light level was minimized. The textual stimuli were presented at the center of a 17-inch computer monitor, in white 14-point Helvetica font on a black background. Also, viewing distance from the monitor was maintained at 23 inches (57 cm) throughout the experiment with a chin rest, which resulted in three characters subtending 1° of visual angle.

Four initial practice trials were followed by ten experimental blocks consisting of 17 trials (16 experimental scenarios and one filler scenario) each. The experimental blocks were each divided by a timed break, the duration of which was determined by the participant. In each trial, the first two scenario sentences – the context (1) sentence and the ambiguous behavior (2) sentence – were presented and displayed until the participant pressed the space bar to indicate that the sentence was read, but only after a minimum display time of 1 500ms. A blank screen interval was then shown for 500ms, followed by a fixation cross presented at the center of the screen for 1 000ms. Next, the intention (3) revealing sentence was shown one word at a time for 300ms, followed by a 200ms blank interval. The next context (1) sentence was then displayed after a 2 000ms interval during which the participant was instructed to maintain his or her gaze on a fixation cross. A filler sentence was included among the 17 trials in each block, followed by a true or false comprehension question, which resulted in a mean correct response rate of 88%. Those scoring below a correct response rate of 60% were removed from the study. Finally, participants were asked to complete computerized versions of the RPQ and the SIP-AEQ following the EEG procedure.

EEG Recording

Using the BioSemi Active Two system, the electroencephalogram (EEG) signal was recorded using 64 Ag / AgCl electrodes mounted on an elastic cap according to the International 10-10 System (Sharbrough et al., 1991), and referenced to the average of the left and right mastoids. Horizontal eye movements were measured using a horizontal electrooculogram (HEOG) – recorded as the difference in voltage between the electrodes placed laterally to the external canthi – while eye blinks were detected with a vertical electrooculogram (VEOG),

recorded here as the difference in voltage between the two electrodes placed above and below the left eye.

Next, EEG and EOG signals were both digitized at 512Hz and low-pass filtered at 134Hz, while the continuous signal was high-pass filtered at .01Hz and low-pass filtered at 30Hz during offline signal analysis. The EEG was epoched from 200ms before target onset (0ms) to 1000ms following target onset and was baseline corrected using the mean from -200ms to 0ms. Also, electro ocular artifact reduction and data correction were both achieved with an independent component analysis (ICA) statistical procedure (see Drisdelle, Aubin, & Jolicoeur, 2016). Trials containing eye blinks (i.e., VEOG>50 μ V within a 150ms interval) and large horizontal eye movements (i.e., HEOG>35 μ V within a 300ms interval) were excluded from the analysis, while trials showing EEG deflections greater than 100 μ V during the predetermined segmentation window on at least one of the 64 electrodes were selected for further analysis. If seven or fewer channels were detected in any given trial, spherical spline interpolation was used to interpolate these channels from neighboring channels, whereas trials were rejected if more than seven channels were found to contain artifacts.

EEG Data Analysis

EEG data analysis was done using MATLAB with functions from the EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes. The EEG signal at each electrode site was averaged separately for the four experimental conditions (hostile-match, hostile-mismatch, non-hostile-match, non-hostile-mismatch) and was time-locked to target word onset. Following an initial visual assessment, a significant negative ERP deflection was identified between 450ms and 650ms, which is within the literature-based N400 effect timeframe (Kutas & Federmeier, 2011). In the Gagnon et al., study (2017), midline electrode amplitude data

were analyzed separately from the data recorded from lateral electrode sites, the latter being pooled into six regions identified by either the left or right hemisphere and by the anterior, central or posterior scalp locations. The left hemisphere was divided into three regions that were identified as follows: left-anterior (AF3, AF7, F1, F3, F5, F7, FT7, FC1, FC3, FC5), left-central (TP7, T7 C1, C3, C5, CP1, CP3, CP5), and left-posterior (P1, P3, P5, P7, PO3, PO7, O1), while the right hemisphere was divided into three equivalent regions that were identified by their homologous hemispherical electrodes. Statistical analyses were carried out by Gagnon and colleagues on the lateral electrodes using a Huynh-Feldt corrected, repeated-measures ANOVA with Hostility (hostile, non-hostile), Consistency (match, mismatch), Hemisphere (left, right), Location (anterior, central, posterior), and Group (aggressive, nonaggressive) as the variables of interest. The midline electrodes were submitted to an ANOVA with variables Hostility, Consistency, Group, and Location (anterior: AFZ, Fz, FCz; central: Cz, CPz; posterior: Pz, POz, Oz). Results showed an N400 effect component following a mismatch between the participant's expectation of a hostile intention and the presented non-hostile intention, as revealed by the target word. Component amplitude was shown to be larger for the non-hostile condition compared to the hostile condition, for the hostile group compared to the non-hostile group, and over the central and posterior electrode sites compared to the anterior sites.

Current study analyses

For the current study, an independent sample t-test comparing scores on the BSI depression subscale for the hostile and non-hostile groups was first performed, followed by a second t-test comparing the two groups on their reactive aggression scores, as did Gagnon et al. (2017) in their study, with the goal of confirming the reactive aggression score group differences observed in their study, but in our case with a smaller sample in which all the clinical participants were

removed. Next, correlational analyses were conducted between the variables of interest for the purpose of identifying all significant associations between said variables.

Furthermore, we tested the hypothesis that depression would predict the N400 effect independently of reactive aggression by first conducting a repeated-measures ANOVA with Condition (hostile, non-hostile), Consistency (match, mismatch), Hemisphere (left, right) and Location (anterior, central, posterior) as the variables of interest, and with depression scores as a covariate and reactive aggression scores as a control variable. It must be noted that, compared to the ANOVA conducted by Gagnon and colleagues, which was described above, the Group variable (hostile vs non-hostile) was not utilized in our own ANOVA since the main goal of our study was to predict HAB, by means of the N400 amplitude as an indicator of the bias, from depression scores collected from the total sample (both hostile and non-hostile participants). Finally, only the lateral electrodes were selected for the ANOVA for the reason that these were found to be adequate for identifying which cerebral regions indicated the presence of an N400 effect. Subsequent decomposition of variance analysis was also performed to isolate specific interactions involving Condition, Consistency, Hemisphere, Location and depression level scores.

Finally, the regression analyses enabled us to test our main hypothesis that depression level scores were related to the hostile attribution bias, as measured by the N400 effect, after controlling for reactive aggression, and to examine the differential effects of the hostile and non-hostile ERP task conditions on the N400 amplitude. Here, the N400 amplitude was measured as the difference between the mean mismatch and match scenario forms (conditions). Also, scalp regions of the N400 effect were selected for the regression analyses based on the ANOVA's results that showed significant interactions between Condition, Consistency, Hemisphere,

Location and covariate depression scores. More precisely, we expected the region of interest to be centrally located and in the right hemisphere, in light of research literature suggesting that the N440 effect is largest over centroparietal sites of the right hemisphere (Kutas & Federmeier, 2011).

Results

Mean ERP waveforms, mean ERP difference waveforms and topographical maps

Lateral electrode (P3, C3, F3, P4, C4, F4) mean ERP waveforms for the match and mismatch conditions, and for the hostile and non-hostile conditions are presented in Figure 1, whereas the mean ERP difference waveforms (ERP mismatch minus ERP match) for the same electrodes, and for the hostile and the non-hostile conditions are presented in Figure 2. Finally, Figure 3 shows the topographic maps of the mean ERP difference waveforms (ERP mismatch minus ERP match) following (a) the violation of non-hostile intention expectancy, and (b) the violation of hostile intention expectancy, while (c) outlines the mean of the hostile and non-hostile conditions for the 450-600ms postcritical word timeframe. It must be noted that all ERP waveforms and topographic maps show results for the entire sample (no group differences are shown here). Also, only the lateral electrodes (P3, C3, F3, P4, C4, F4) were selected for our analysis, and are therefore the only ones reported in Figure 1 and Figure 2.

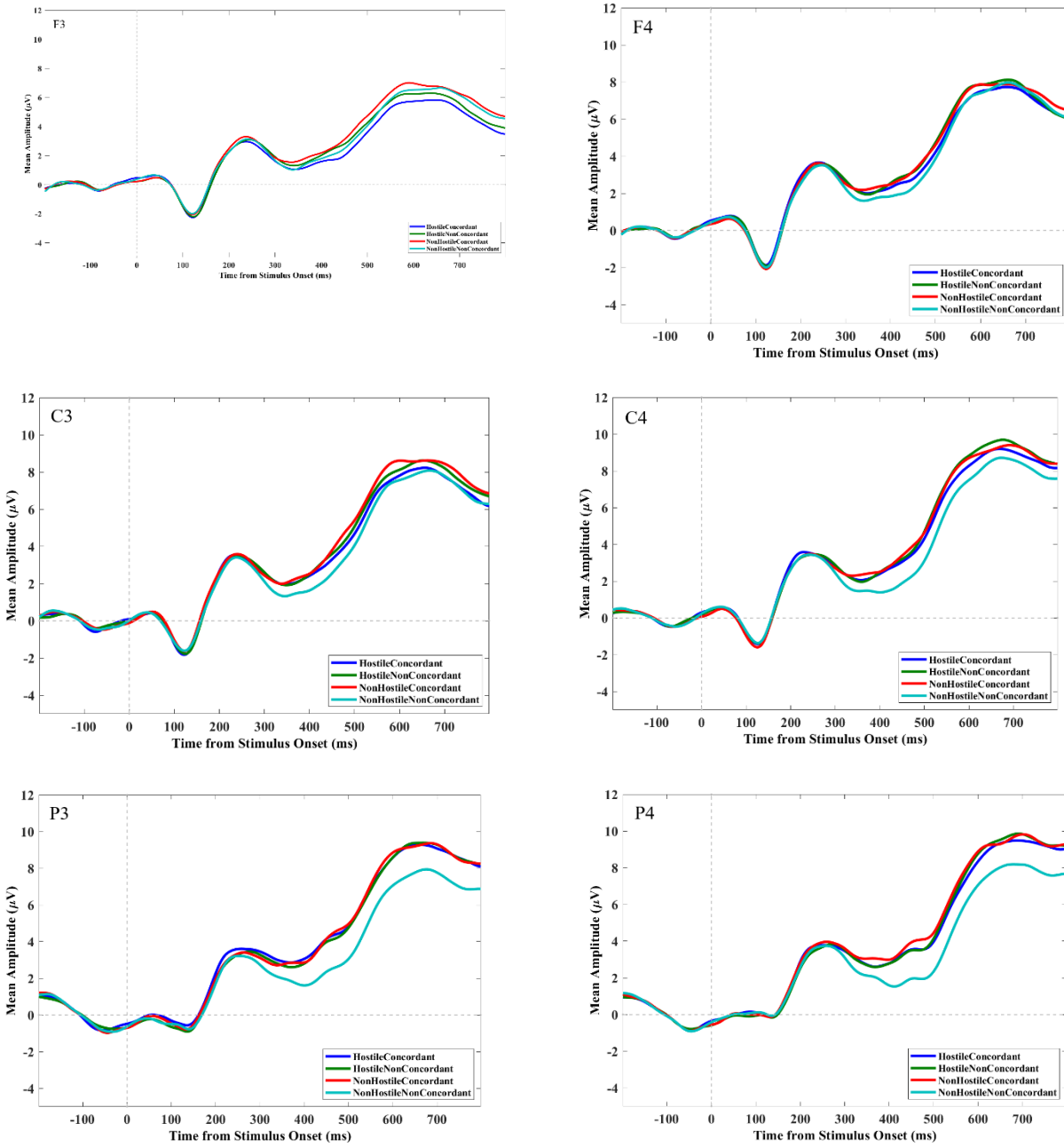


Fig. 1 Lateral electrode mean ERP waveforms for the match and mismatch conditions, and for the hostile and non-hostile conditions.

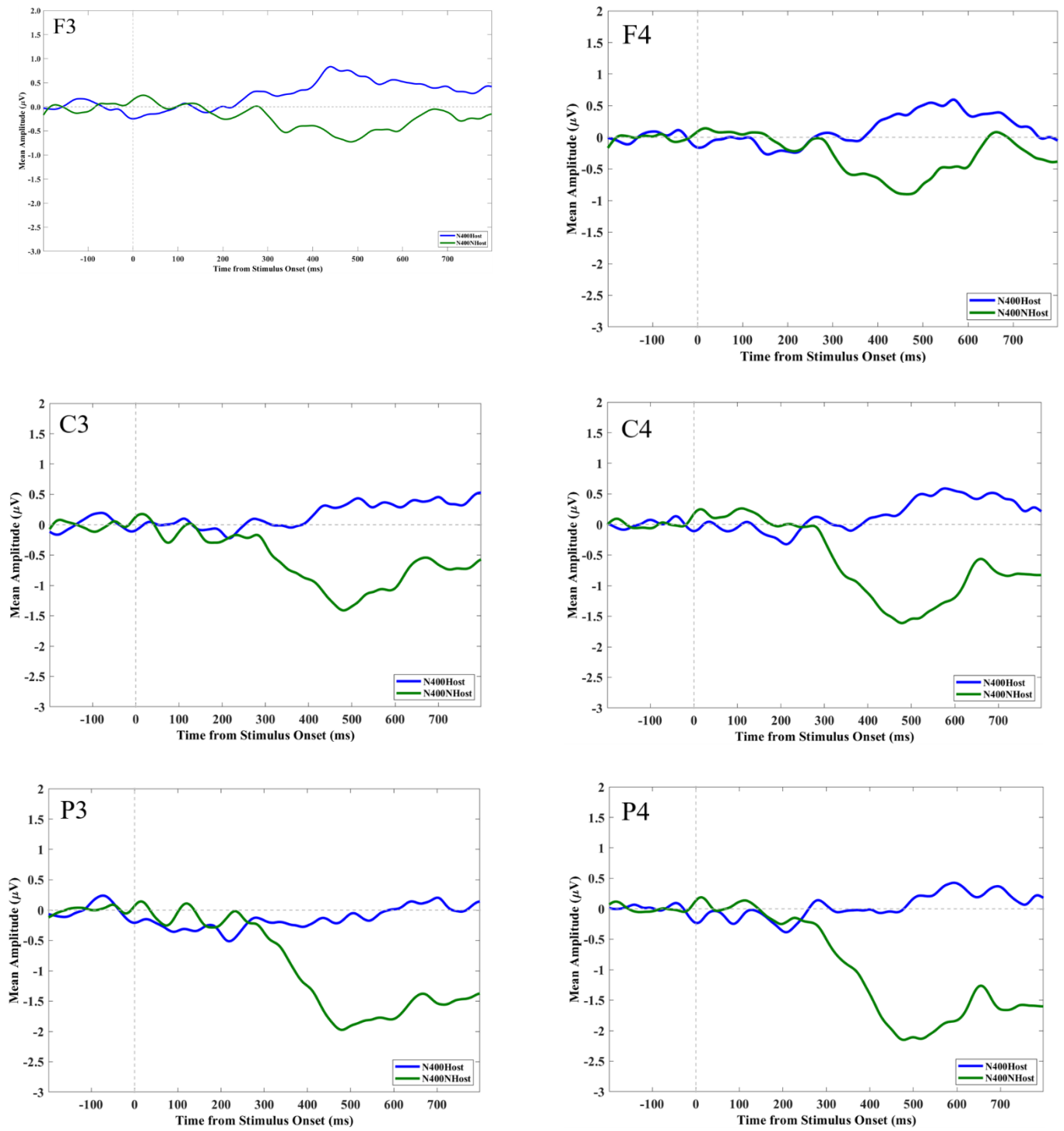
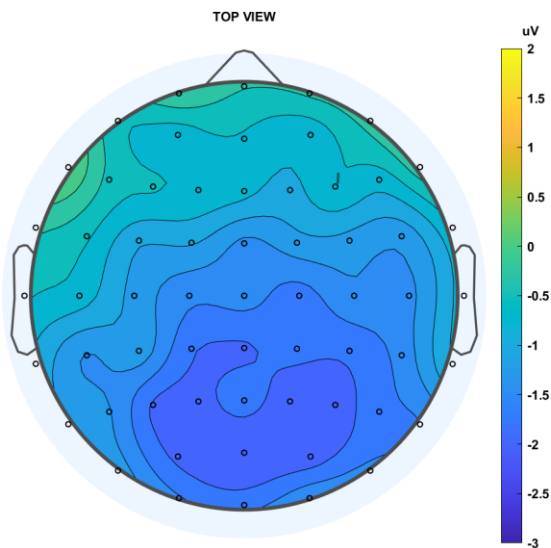
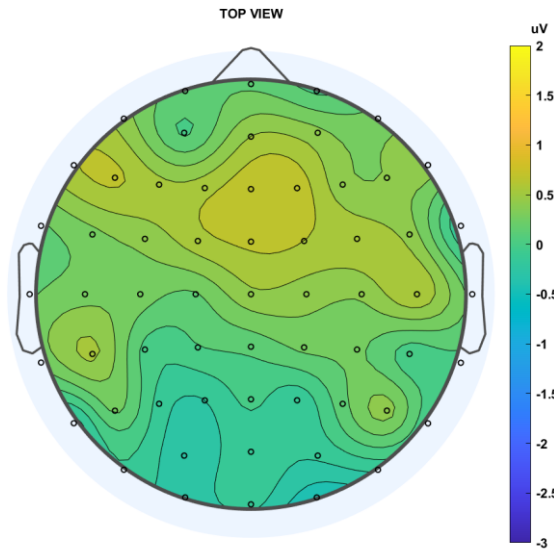


Fig. 2 Mean ERP difference waveforms (ERP mismatch minus ERP match) for the hostile and non-hostile conditions.

(a)



(b)



(c)

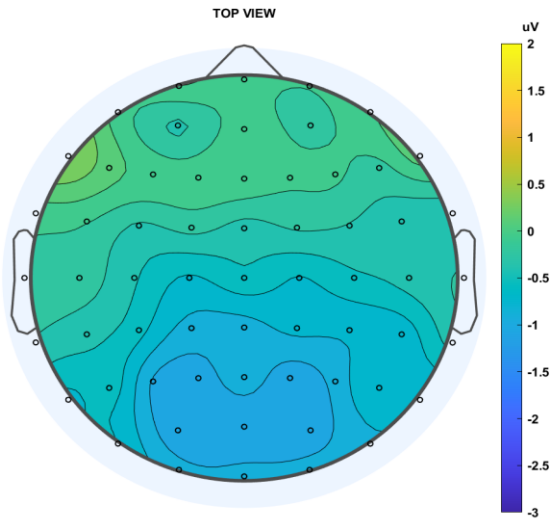


Fig. 3 Topographic maps of the mean ERP difference waveforms (ERP mismatch minus ERP match) following (a) the violation of non-hostile intention expectancy, and (b) the violation of hostile intention expectancy, while (c) outlines the mean of the hostile and non-hostile conditions for the 450-600 postcritical word timeframe.

Descriptive statistics

Table 2 outlines the means, standard deviations and ranges for right central amplitudes (in microvolts) for variables Condition (hostile and non-hostile) and Consistency (match and mismatch), and for the N400 effect in the hostile condition and the non-hostile condition. It must be noted that the right central and right anterior regions were identified with the ANOVA as the regions of interest with which to conduct our regression analyses. Furthermore, regression analyses revealed that the right anterior overall models for the hostile and non-hostiles conditions were non-significant. Therefore, data for the right anterior region were omitted in the descriptive statistics table.

Table 2

<i>Descriptive Statistics of Right Central Amplitudes in Microvolts (μV)</i>					
	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Hostile right central match	72	-1.83	16.33	5.56	3.26
Hostile right central mismatch	72	-1.97	13.89	5.83	3.41
Non-hostile right central match	72	.50	19.31	5.93	3.55
Non-hostile right central mismatch	72	-1.13	18.68	4.53	3.30
Hostile right central N400	72	-6.44	4.69	.27	2.00
Non-hostile right central N400	72	-5.29	2.45	-1.40	1.65

Note. *n*, *M* and *SD* are used to represent sample size, mean and standard deviation, respectively. *Min* and *Max* denote amplitude ranges.

T-tests

Table 3 outlines the comparative analyses performed on the hostile and non-hostile groups regarding depression level scores and reactive aggression scores.

Results of the first independent samples t-test showed that the scores on the BSI depression subscale were significantly higher for the hostile group, than for the non-hostile group.

Furthermore, as noted above, a second independent samples t-test was performed to confirm that the results obtained by Gagnon et al. (2017) regarding group differences in levels of reactive

aggression still held true with a smaller sample size in which all clinical patients were removed. In their study, comparative analysis results showed that the hostile group scored higher on the RPQ reactive aggression subscale than the non-hostile group. Our analyses also showed similar results, in that the hostile group showed significantly higher reactive aggression levels than the non-hostile group.

Table 3

Sample Descriptives Using t-test for Equality of Means

	Hostile		Non-hostile		<i>t</i> -test
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Standardized BSI depression score	.32	.12	.15	.14	-4.97**
RPQ reactive aggression score	1.11	.40	.71	.31	-4.63**

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. * indicates $p < .05$. ** indicates $p < .01$.

Correlations

Table 4 outlines correlational analysis results for all the examined variables of interest. Analyses results revealed significant positive relations between BSI depression subscale scores and RPQ reactive aggression scores, between BSI depression scores and BSI hostility scores, and between BSI hostility scores and RPQ reactive aggression scores. Also, a significant negative relation between BSI depression subscale scores and the right central N400 effect size for the non-hostile condition was observed. However, the correlation between the RPQ reactive aggression subscale scores and the right central N400 effect size for the non-hostile condition was found to be non-significant, although the association did show a trend towards significance. Furthermore, the association between the self-reported SIP-AEQ scores measuring the hostile attribution bias and the RPQ reactive aggression scores was found to be significant.

Table 4

Means, Standard Deviations, and Correlations

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7
1. Standardized BSI depression score	.21	.16							
2. RPQ reactive aggression score	.86	.39	.40**						
3. Non-hostile right central N400	-1.40	1.65	-.34**	-.23					
4. Hostile right central N400	.27	2.00	.31**	.33**	-.25*				
5. BSI hostility score	.92	.92	.58**	.70**	-.44**	.25*			
6. SIP-AEQ	1.13	.44	.07	.24*	-.06	.15	.14		
7. Sex	.58	.50	-.08	-.11	.06	-.11	-.14	-.07	
8. Age	26.53	8.15	-.13	-.04	.23	.01	-.04	.03	-.08

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. * indicates $p < .05$. ** indicates $p < .01$.

SIP-AEQ is the self-reported measure of HAB.

Analyses of variance

Table 5 outlines the repeated measures ANOVA main Consistency effect and the analysis results with Condition, Consistency, Hemisphere and Region as factors, and with standardized BSI depression scores as a covariate and reactive aggression scores as a control variable.

Because we were mainly interested in examining hostile expectancy violations in relation to depression, main effects and interactions involving the Consistency factor and depression were the focus of these analyses. First, the repeated measures ANOVA revealed a significant main Consistency effect, meaning that the waveforms were significantly more negative in the mismatch condition than in the match condition, which in turn confirms the presence of the N400

expectancy violation effect. Importantly, there was a significant interaction between Condition, Consistency, Hemisphere, Location and BSI depression scores. Decomposition of variance analysis (Condition X Consistency) for the six regions (Hemisphere by Location) revealed that BSI depression scores were significantly associated with more negative waveforms in the mismatch non-hostile condition measured in the right central site (Wilks' Lambda = .90, $F(1,69) = 7.45$, $p = .008$), and right anterior site (Wilks' Lambda = .94, $F(1,69) = 4.73$, $p = .033$). The other regions were found to be not significant (right posterior, Wilks' Lambda = .96, $F(1,69) = 2.69$, $p = .11$; left posterior, Wilks' Lambda = .96, $F(1,69) = 3.21$, $p = .080$; left central, Wilks' Lambda = .96, $F(1,69) = 2.65$, $p = .11$; and left anterior, Wilks' Lambda = .94, $F(1,69) = 3.46$, $p = .070$).

Table 5

Repeated Measures ANOVA with Condition, Consistency, Hemisphere and Region as Factors, and with Standardized BSI Depression Scores as a Covariate and Reactive Aggression Scores as a Control Variable

<i>Function</i>	<i>Wilks' lambda</i>	<i>F</i>	<i>Hyp df</i>	<i>Err df</i>	<i>p</i>
Consistency	.93	5.43	1	69	.02
Condition X Consistency X Hemisphere X Region X Standardized BSI depression score	.91	3.33	2	68	.04

Note. Hyp df=degrees of freedom for hypothesis, Err df=degrees of freedom for error.

Multiple Regression

Table 6 reports the multiple regression results using the N400 effect for the non-hostile condition as the criterion, whereas Table 7 reports results for the same analysis using the N400 effect for the hostile condition as the criterion for both the right central and right anterior regions.

We first tested for multicollinearity in the regression models and found that the variance inflation factors ranged between 1.022 and 1.214, indicating that collinearity was not a problem

in these models. Next, following the results of the repeated-measures ANOVA in which both the right central and right anterior regions showed significant interactions involving factors Condition, Consistency, and Depression, four standard multiple regressions were performed to measure the combined effect of depression, reactive aggression, age, and sex on the N400 amplitude for both the non-hostile and hostile conditions. The first analysis, which was performed using the right central N400 effect size for the non-hostile condition as the dependent variable, resulted in a significant overall model, in which depression was a significant predictor of the N400 effect, whereas reactive aggression, sex and age were not (see table 6). Furthermore, the multiple regression performed with the right central N400 effect size for the hostile condition revealed a significant overall model, but in which depression, reactive aggression, sex and age were all found to be nonsignificant (see table 7). Finally, multiple regressions were performed using the right anterior locations for the non-hostile ($F=2.46$, $p=.054$) and hostile conditions ($F=1.00$, $p=.42$), but neither showed significant overall models.

Table 6

<i>Multiple Regression Using Non-Hostile Right Central and Right Anterior N400 as the Criterion</i>									
Variable	<i>Standardized coefficients</i>				<i>Overall model fit</i>				
	<i>B</i>	<i>SE B</i>	β	<i>t</i>	<i>Adj. R²</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Right central N400					.11	67	7.61	3.16	.02
Sex	.15	.39	.05	.39					.70
Age	.04	.02	.20	1.72					.09
Standardized BSI depression score	-2.89	1.34	-.27	-2.16					.04
RPQ reactive aggression score	-.44	.53	-.11	-.83					.41
Right anterior N400					.08	67	5.15	2.46	.05

Note. *B*=unstandardized coefficient, *SE*=standard error, β =standardized regression coefficient, *MS*=mean square.

Table 7

<i>Multiple Regression Using Hostile Right Central and Right Anterior N400 as the Criterion</i>									
Variable	<i>Standardized coefficients</i>				<i>Overall model fit</i>				
	<i>B</i>	<i>SE B</i>	β	<i>t</i>	<i>Adj. R²</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Right central N400					.10	67	10.13	2.80	.03
Sex	-.27	.47	-.07	-.57					.57
Age	.01	.03	.04	.36					.72
Standardized BSI depression score	2.75	1.64	.22	1.69					.10
RPQ reactive aggression score	1.18	.65	.23	1.82					.07
Right anterior N400					.00	67	3.65	1.00	.42

Note. *B*=unstandardized coefficient, *SE*=standard error, β =standardized regression coefficient, *MS*=mean square.

Discussion

Drawing from Smith's study, in which the authors concluded that HAB is a separate construct from trait anger and that both characterize depression in distinctive ways, the purpose of this study was to examine the relation between depressive symptoms and HAB using electrophysiological measures of the hostile bias, namely, the N400 ERP component. Our decision to utilize the N400, and the paradigm to measure it, as a marker of the HAB is based on the characteristics that define the component. Although the N400 is commonly interpreted as a reflection of a violation of semantic expectations in the context of language comprehension, semantic anomalies are neither necessary nor sufficient to evoke an N400. The neural signature of the N400 is therefore not related to a simple representation of the violation of an arbitrary or over-learned pattern of cognition, but rather, is a fundamental aspect in the processing of what is salient (Kutas & Federmeier, 2011), with higher amplitudes of the N400 (more negative) being typically observed following more unexpected, or less salient, stimuli.

For the current study, we first hypothesized that depression level scores would correlate positively with HAB self-reported measures (SIP-AEQ) (Ho 1). Next, we anticipated that the hostile group would present higher depression level scores than the non-hostile group (Ho 2). Finally, we hypothesized that depression levels would be associated to the hostile attribution bias (as measured by the N400 effect following the presentation of a critical word that violated a hostile expectation), independently of reactive aggression (Ho 3), and that this association would be present in both the non-hostile (negative association) (Ho 3a) and hostile (positive association) (Ho 3b) conditions.

Regarding our first hypothesis (Ho 1), no relation was found between the self-reported measure of HAB (SIP-AEQ) and depressive symptom levels, which is contrary to Smith's findings and our hypothesis. A reason that could explain these divergent results is that Smith's measure of HAB may be more sensitive in its evaluation of the bias than our own measure.

For our second hypothesis (Ho 2), results of the comparative analysis (t-test) showed that the hostile and non-hostile groups differed markedly regarding their depression symptom levels, as expected, with the hostile group showing higher symptom levels than the non-hostile group. This is supported by research findings that suggest that overtly hostile, aggressive and violent individuals often manifest depressive symptomatology that is often chronic, severe and comorbid (Faizel et al., 2015; Judd et al., 2013).

Next, we conducted a repeated measures ANOVA with factors Condition, Consistency, Hemisphere and Location), and with depression as a covariate and reactive aggression as a control variable, to guide our subsequent regression analyses and confirm literature data recognizing the right central region as being associated with the hostile expectancy violation effect. Here, the repeated measures ANOVA revealed that the amplitude waveforms were

significantly more negative in the mismatch condition than in the match condition (Consistency), which is indicative of an N400 hostile expectancy violation effect. We also observed significant interactions between Condition, Consistency and Depression in the right-central and right-anterior regions. Hence, for the mismatch condition, a more negative amplitude in the non-hostile condition was observed in the central and anterior (to a lesser extent) regions of the right hemisphere when depression was factored in as a covariate. This result is therefore in line with research literature suggesting that the N440 effect is largest over centroparietal sites, as stated above, with a slight right hemisphere bias for written words in sentences (Kutas & Federmeier, 2011). However, to account for the presence of the N400 in the right anterior region, it is possible that depression may be associated with additional cognitive resources necessary to integrate non-hostile stimuli in a hostile context (Van Berkum et al. , 2008), but more studies would be needed to further explain this observation.

All in all, these results supported the notion that depression did indeed interact with the aforementioned variables and served to guide our subsequent regression analyses by identifying which scalp regions warranted further analysis, mainly the right central and right anterior regions. Although the overall regression model for the right anterior region turned out to be non-significant, our findings for the right central region supported the hypothesis that depressive symptom levels were negatively associated with the N400 amplitude for the non-hostile condition, after controlling for reactive aggression (Ho 3a). Hence, the more depressed one is, the more one expects hostile intentions underlying an ambiguous behavior in a hostile context, independently of reactive aggression. In this case, an N400 was elicited when a hostile context was followed by a non-hostile intention, due to a violation of hostile expectations. However, the component amplitude was greater in those who presented higher depression levels, in that the

hostile context served as a stronger primer for the expectation of a subsequent hostile intent. From this observation, we may posit that individuals with higher levels of depression react more strongly to the hostile context, due to a possible hypersensitivity towards negative social cues.

Next, our hypothesis that depressive symptom levels were positively related to the N400 amplitude for the hostile condition, after adjusting for reactive aggression (Ho 3b), was not confirmed. Although the overall model was found to be significant, neither depression levels nor reactive aggression scores correlated significantly with the N400 amplitude. However, reactive aggression showed a strong trend toward significance, as did depression, thus partially supporting the Gagnon et al. (2017) results which showed that the N400 effect in the hostile condition was associated with reactive aggression.

This study does however have some limitations, one of which is the lack of statistical power, which can be explained by the small sample size, another limitation of our study. Of course, an increase in sample size could potentially solve this problem. Also, because a small, nonclinical sample selected on hostility was used, the sample showed relatively low scores on the BSI depression subscale. Greater depression symptom levels could be expected in a larger clinical sample, which could validate our results and conclusions further. Another limitation comes from the participant sample selection. Because participants were selected based on their hostility level scores, we were unable to examine specific group differences regarding HAB in depressed and non-depressed individuals. Next, studies have suggested that anger, hostility, and irritability during depressive episodes may be interpreted as indicators of an undiagnosed or subthreshold bipolar disorder (Perlis et al., 2009), which we did not adjust for. Finally, as with Smith et al. (2016), we were unable to establish any causal relationship between the examined factors due to the correlational nature of our study.

Conclusion

In conclusion, although the results from the present study were obtained using a non-clinical sample, thus limiting the range of recommendations that could ultimately be made, several important implications for research and practice can nevertheless be ascertained. First, this study adds to our understanding of perception mechanisms in depression by suggesting that individuals who present higher levels of depression are more sensitive to hostile signals in social situations but will not perceive hostility when none is present. Also, it must be noted that because HAB can be regarded as a maintenance factor for depression due to its deteriorating effects on relationships, which can lead to social isolation and increased feelings of self-worthlessness (Smith et al., 2016), cognitive therapies should consider targeting hostile biases in depression to reduce the risk of depression recurrence in highly aggressive individuals. Furthermore, the current study is, to our knowledge, the first to examine HAB in the context of depression using an EEG methodology, while reemphasizing that ERPs represent a precise assessment tool for evaluating automatic cognitive processes. The strength of this methodology hence lies in its ability to accurately assess the existence and functioning of unconsciously activated schemas or associative networks. The current study therefore highlights the need for future research on interpretative biases to consider important theoretical and methodological factors, one of which is the use of a methodology and experimental paradigm that both assess automatic and implicit processes.

References

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., text rev.) Washington, DC: American Psychiatric Association.
- Bailey, C. A., & Ostrov, J. M. (2008). Differentiating forms and functions of aggression in emerging adults: Associations with hostile attribution biases and normative beliefs. *Journal of Youth and Adolescence*, 37(6), 713-722.
- Bargh, J. A. (1984). Automatic and conscious processing of social information. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition*, (Vol. 3, pp. 1-43). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper and Row.
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: New American Library.
- Benazzi, F., & Akiskal, H. (2005). Irritable-hostile depression: Further validation as a bipolar depressive mixed state. *Journal of Affective Disorders*, 84(2-3), 197-207.
- Berkowitz, L. (1977). Situational and personal conditions governing reactions to aggressive cues. In D. Magnusson & N. S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology* (pp.165-171). Hillsdale, NJ: Erlbaum.
- Boland, R. J., & Keller, M. B. (2002). Course and outcome of depression. In I. H. Gotlib, & C. L. Hammen (Eds.), *Handbook of depression* (pp. 43-60). New York, NY, USA: Guilford.

- Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development, 30*(5), 466-480.
- Chen, P., Coccaro, E. F., & Jacobson, K. C. (2012). Hostile attributional bias, negative emotional responding, and aggression in adults: Moderating effects of gender and impulsivity. *Aggressive Behavior, 38*(1), 47-63.
- Coccaro, E. F., Noblett, K. L., & McCloskey, M. S. (2009). Attributional and emotional responses to socially ambiguous cues: Validation of a new assessment of social/emotional information processing in healthy adults and impulsive aggressive patients. *Journal of Psychiatric Research, 43*, 915-925.
- Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*(1), 74-101.
- Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms on reactive and proactive aggression. *Child Development, 67*(3), 993-1002.
- Dearing, K. F., & Gotlib, I. H. (2009). Interpretation of ambiguous information in girls at risk for depression. *Journal of Abnormal Child Psychology, 37*(1), 79-91.
- 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.
- Derogatis, L. R. (1993). *BSI brief symptom inventory. Administration, scoring and procedures manual* (4th ed.). Minneapolis, MN:National Computer Systems.

- Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, **13**(3), 595–605.
- Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology*, *53*(6), 1146-1158.
- Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and Psychopathology*, *18*(3), 791-814.
- Drisdelle, B. L., Aubin, M., & Jolicoeur, P. (2016). Dealing with ocular artifacts on lateralized event-related potentials in studies of visuospatial attention and memory: ICA correction versus epoch rejection. *Psychophysiology*, *54*, 83-99.
- Fava, G.A., Kellner, R., Munari, F., Pavan, L., Pesarin, F. (1982). Losses, hostility, and depression. *The Journal of Nervous and Mental Disease*, *170*, 474-478.
- Faizel, S., Wolf, A., Chang, Z., Larsson, H., Goodwin, G. M., & Lichtenstein, P. (2015). Depression and violence: A Swedish population study. *The Lancet Psychiatry*, *2*(3), 224-232.
- Fite, P. J., Stoppelbein, L., & Greening, L. (2009). Proactive and reactive aggression in a child psychiatric inpatient population. *Journal of Clinical Child and Adolescent Psychology*, *38*(2), 199-205.

- Gagnon, J., & Jolicœur, P. (2014). Mécanismes cognitifs et neurophysiologiques associés au trait "Urgence" de la personnalité: intégration théorique et empirique. *Insight Development: Social Sciences and Humanities Research Council of Canada*.
- Gagnon, J., McDuff, P., Daelman, S., & Fournier, S. (2015). Is hostile attributional bias associated with negative urgency and impulsive behaviors? A social-cognitive conceptualization of impulsivity. *Personality and Individual Differences, 72*, 18-23.
- Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Bessette, M., and Jolicœur, P. (2016). Neural mechanisms underlying attribution of hostile intention in nonaggressive individuals: an ERP study. *International Journal of Psychophysiology 110*, 153-162.
- Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Brochu, A. F., Bessette, M., & Jolicœur, P. (2017). An ERP study on hostile attribution bias in aggressive and nonaggressive individuals. *Aggressive Behavior, 43*(3), 217-229.
- Gagnon, J., & Rochat, L. (2017). Relationships between hostile attribution bias, negative urgency, and reactive aggression. *Journal of Individual Differences, 38*(4), 211-219.
- Glöckner, A., & Betsch, T. (2008). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making, 3*(3), 215-228.
- Gotlib, I. H., & Krasnoperova, E. (1998). Biased information processing as a vulnerability factor for depression. *Behavior Therapy, 29*(4), 603-617.

- Judd, L.L., Schettler, P.J., Coryell, W., Akiskal, H.S., Fiedorowicz, J.G. (2013). Overt Irritability/Anger in Unipolar Major Depressive Episodes: Past and Current Characteristics and Implications for Long-term Course. *JAMA Psychiatry*, 70(11), 1171-1180.
- Kennedy, H. G. (1992). Anger and irritability. *The British Journal of Psychiatry*, 161, 145-153.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., et al. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association*, 289(23), 3095-3105.
- Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, 11, 539-550.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621-647.
- Linder, J., Werner, N., & Lyle, K. (2010). Automatic and controlled social information processing and relational aggression in young adults. *Personality and Individual Differences*, 49(7), 778-783.
- 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.
- Luck, S.J. (2014). *An Introduction to the Event-Related Potential Technique* (2nd ed.). Cambridge, MA: The MIT Press.
- Milich, R., & Dodge, K. A. (1984). Social information processing in child psychiatric populations. *Journal of Abnormal Child Psychology*, 12(3), 471-489.

- Miller, J. D., & Lynam, D. (2006). Reactive and proactive aggression: Similarities and differences. *Personality and Individual Differences, 41*, 1469-1480.
- Mogg, K., Bradbury, K. E., & Bradley, B. P. (2006). Interpretation of ambiguous information in clinical depression. *Behaviour Research and Therapy, 44*(10), 1411-1419.
- Moreno, J. K., Selby, M. J., Fuhriman, A., & Laver, G. D. (1994). Hostility in depression. *Psychological Reports, 75*(3, Pt 1), 1391-1401.
- Nasby, W., Hayden, B., & DePaulo, B. M. (1980). Attributional bias among aggressive boys to interpret unambiguous social stimuli as displays of hostility. *Journal of Abnormal Psychology, 89*(3), 459-468.
- Orobio de Castro, B., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2002). Hostile attribution of intent and aggressive behavior: A meta-analysis. *Child Development, 73*(3), 916-934.
- Pedrelli, P., Nyer, M., Holt, D., Bakow, B.R., Fava, M., Baer, L., ... Farabaugh, A. (2013). Correlates of irritability in college students with depressive symptoms. *The Journal of Nervous and Mental Disease, 201*, 953-958
- Perlis, R. H., Fava, M., Trivedi, M. H., Alpert, J., Luther, J. F., Wisniewski, S. R., & Rush, A. J. (2009). Irritability is associated with anxiety and greater severity, but not bipolar spectrum features, in major depressive disorder. *Acta Psychiatrica Scandinavica, 119*(4), 282-289.
- Pitt, B. (1968). "Atypical" depression following childbirth. *The British Journal of Psychiatry, 114*(516), 1325-1335.

- Raine, A., Dodge, K. A., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., Stouthamer-Loeber, M., & Liu, J. (2006). The reactiveproactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior*, 32, 159-171.
- Riley, W. T., Treiber, F. A., & Woods, M. G. (1989). Anger and hostility in depression. *Journal of Nervous and Mental Disease*, 177(11), 668-674.
- Safer, D. J. (2009). Irritable mood and the diagnostic and statistical manual of mental disorders. *Child and Adolescent Psychiatry and Mental Health*, 3, Article ID 35.
- Sharbrough, F., Chatrian, G.-E., Lesser, R. P., Luders, H., Nuwer, M., & Picton, T. W. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 8, 200-202.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127-190.
- Smith, H. L., Summers, B. J., Dillon, K. H., Macatee, R. J., & Couglen, J. R. (2016). Hostile interpretation bias in depression. *Journal of Affective Disorders*, 203, 9-13. doi: 10.1016/j.jad.2016.05.070
- Snaith, R. P., & Taylor, C. M. (1985). Irritability: Definition, assessment and associated factors. *The British Journal of Psychiatry*, 147, 127-136.
- Statistics Canada. (2012). Table: 105-1101 - *Canadian Community Health Survey Share File - Mental Health (CCHS)*, CANSIM (database).

Stewart, L. A., Harris, A., Wilton, G., Archambault, K., Cousineau, C., Varrette, S. (2010). An initial report on the results of the pilot of the computerized mental health intake screening system. Correctional Service of Canada, Ottawa, Ontario. No. R-218.

Todorov, A., & Bargh, J. A. (2002). Automatic sources of aggression. *Aggression and Violent Behavior, 7*, 53-68.

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20*(4), 580–591.

Wilkowski, B. M., & Robinson, M. D. (2008). The cognitive basis of trait anger and reactive aggression: An integrative analysis. *Personality and Social Psychology Review, 12*(1), 3-21.

Zelli, A., Huesmann, L. R., & Cervone, D. (1995). Social inference and individual differences in aggression: Evidence for spontaneous judgments of hostility. *Aggressive Behavior, 21*, 4015-4017.

Chapitre III : Conclusion générale

Discussion

S'appuyant sur l'étude de Smith dans laquelle les auteurs ont conclu que le BAIH est un construit distinct du trait de colère et que les deux caractérisent la dépression de manière indépendante, le but de notre étude était d'examiner la relation entre les symptômes dépressifs et le BAIH en utilisant une mesure électrophysiologique du biais, à savoir la composante PRE N400. Notre décision d'utiliser la composante N400, mesurée par le biais du paradigme de la violation des attentes hostiles, en tant que marqueur du BAIH est basé sur les caractéristiques particulières qui spécifient la composante. Ainsi, bien que la N400 soit généralement interprétée comme le reflet d'une violation des attentes sémantiques en lien avec la compréhension du langage, les anomalies sémantiques ne sont ni nécessaires ni suffisantes pour évoquer cette composante. La signature neuronale de la N400 n'est donc pas liée à une simple représentation de la violation d'un modèle cognitif arbitraire ou sur-appris, mais est plutôt un aspect fondamental dans le traitement de ce qui est saillant (Kutas & Federmeier, 2011), ce qui résulte généralement dans l'observation d'amplitudes plus élevées de la composante en réponse à des stimuli inattendus ou moins saillants.

En ce qui concerne la présente étude, nous nous attendions d'abord à ce que les scores obtenus sur la mesure auto-rapportée du BAIH (SIP-AEQ) corrèlent positivement avec les niveaux de dépression (Ho 1). Ensuite, nous nous attendions à ce que le groupe hostile présente des niveaux de dépression supérieurs au groupe non hostile (Ho 2). Finalement, nous avons anticipé que les scores de niveaux de dépression seraient liés au BAIH, indépendamment des scores d'agressivité réactive (Ho 3), et que cette association existerait à la fois dans la condition non-hostile (association négative) (Ho 3a) et la condition hostile (association positive) (Ho 3b) du paradigme de violation des attentes hostiles.

Pour la première hypothèse (Ho 1), aucune relation n'a été constatée entre la mesure auto-rapportée du BAIH (SIP-AEQ) et les niveaux de symptômes dépressifs. Ce résultat est contraire à celui de Smith et à notre hypothèse, et peut possiblement être expliqué par le fait que la mesure utilisée par Smith pour évaluer le BAIH soit plus sensible que la nôtre.

Pour ce qui est de notre deuxième hypothèse (Ho 2), les résultats de l'analyse comparative ont montré que les niveaux de symptômes dépressifs étaient sensiblement plus élevés pour le groupe hostile que pour le groupe non hostile, comme anticipé. Ceci est soutenu par des résultats de recherche qui suggèrent que les individus ouvertement hostiles, agressifs et violents manifestent une symptomatologie dépressive qui est souvent chronique, sévère et comorbide (Faizel et al., 2015; Judd et al., 2013).

Ensuite, des ANOVA à mesures répétées ont été effectuées pour guider nos analyses de régression ultérieures et confirmer les données de la littérature reconnaissant la région centrale droite comme étant associée à l'effet de violation des attentes hostile, soit l'effet N400. D'abord, les données ont révélé que les amplitudes étaient significativement plus négatives dans la condition discordante que dans la condition concordante (Consistance), ce qui suggère la présence d'un effet N400. Nous avons également observé des interactions significatives entre les facteurs Condition et Consistance, et la dépression, dans les régions centrale droite et antérieure droite. Par conséquent, pour la condition discordante (Consistance), une amplitude plus négative dans la condition non hostile (Condition) a été observée dans les régions centrale et antérieure (mais dans une moins grande mesure) de l'hémisphère droit lorsque la dépression a été prise en compte comme covariable. Ce résultat est ainsi conforme à la littérature de recherche suggérant que l'effet N440 est plus important sur les sites centropariétaux de l'hémisphère droit (Kutas et Federmeier, 2011). Toutefois, pour rendre compte de la présence de la N400 dans la région

antérieure droite, il se peut que la dépression soit associée à un déploiement de ressources cognitives supplémentaires pour intégrer des stimuli non hostiles dans un contexte hostile (Van Berkum et al., 2008), mais d'autres études seraient requises pour expliquer cette observation.

En somme, les résultats des ANOVA mentionnées ci-dessus soutiennent la notion qu'il existe effectivement une interaction entre la dépression et les variables Condition et Consistance, et servent à guider nos analyses de régression ultérieures en identifiant les régions du cuir chevelu qui méritent une analyse supplémentaire. Ainsi, bien que le modèle de régression pour la région antérieure droite se soit avéré non significatif, nos résultats pour la région centrale droite ont soutenu l'hypothèse selon laquelle les niveaux de symptômes dépressifs étaient négativement associés à l'amplitude N400 pour la condition non hostile (Ho 3a), après avoir contrôlé pour l'agressivité réactive. Autrement dit, plus les niveaux de symptômes dépressifs augmentent, plus l'amplitude de la N400 augmente (devient plus négative), ou encore, plus on est déprimé, plus on s'attend à des intentions hostiles sous-jacentes à un comportement ambigu dans un contexte hostile, indépendamment de l'agressivité réactive. Ici, une N400 a été évoquée lorsqu'un contexte hostile était suivi d'une intention non hostile, en raison d'une violation des attentes hostiles. De plus, l'amplitude de la composante était plus élevée chez ceux qui présentaient des niveaux de dépression plus élevés. Dans ce cas-ci, on peut dire que le contexte hostile a servi d'amorce plus forte dans l'anticipation d'une intention hostile subséquente. À partir de cette observation, une interprétation potentielle est que les individus présentant des niveaux de dépression plus élevés réagissent plus fortement au contexte hostile, possiblement en raison d'une hypersensibilité aux indices sociaux négatifs.

Ensuite, notre hypothèse selon laquelle les niveaux de symptômes dépressifs étaient positivement liés à l'amplitude N400 centrale droite pour la condition hostile, après avoir

contrôlé pour l'agressivité réactive (Ho 3b), n'a pas été confirmée. Bien que le modèle de régression se soit avéré significatif, ni les niveaux de dépression ni les scores d'agressivité réactive ne corrélaient de manière significative avec l'amplitude de la composante N400. Cependant, l'agressivité réactive a montré une forte tendance à la signification, tout comme la dépression, soutenant ainsi, du moins en partie, les résultats de Gagnon et al. (2017) qui ont montré que l'effet N400 était associé de manière significative à l'agressivité réactive dans la condition hostile.

Finalement, nous avons jugé pertinent l'idée de discuter des résultats ne faisant pas explicitement parties de nos hypothèses (analyses corrélationnelles secondaires et analyse comparant les niveaux d'agressivité réactive entre les groupes). Tout d'abord, notre analyse corrélacionnelle a démontré un lien significatif entre la mesure électrophysiologique du BAIH et les symptômes dépressifs, ce qui diffère de l'absence de lien significatif entre les scores obtenus sur la mesure auto-rapportée du BAIH (SIP-AEQ) et les niveaux de symptômes dépressifs, rapportée ci-haut (Ho 1). Une explication possible est que, comme indiqué auparavant, les PRE contournent les mécanismes de contrôle délibérés et conscients qui influencent le traitement de l'information, alors que les questionnaires auto-rapportés nécessitent l'activation de processus de réflexion délibérés pour endosser une réponse auto-déclarée (Linder et al., 2010).

De plus, les symptômes dépressifs étaient associés de manière significative aux scores d'agressivité réactive, ainsi qu'aux niveaux d'hostilité des participants, appuyant de ce fait les résultats de nombreuses études qui associent l'agressivité réactive à un état affectif négatif et à des niveaux élevés de tristesse et de mélancolie (Dodge & Coie, 1987; Card & Little, 2006; Day et al., 1992; Raine et al., 2006; Miller & Lynam, 2006).

Ensuite, nous avons constaté que les niveaux d'hostilité des participants étaient également associés de manière significative à leurs niveaux d'agressivité réactive, ce qui n'est pas surprenant du fait que l'hostilité et l'agressivité réactive sont des concepts connexes qui se chevauchent, mais qui présentent également des différences fondamentales qui les distinguent – l'hostilité étant un concept plus cognitif, tandis que l'agressivité réactive est un concept plus comportemental (Kennedy et al., 1992) – d'où la raison d'analyser le lien entre ces construits.

Par la suite, dans l'étude de Smith, une corrélation positive significative a été observée entre les scores de dépression et la sous-échelle des évaluations hostiles de leur mesure auto-déclarée du BAIH, alors qu'aucune corrélation significative n'a été constatée entre les scores de dépression et la sous-échelle des évaluations bénignes (Smith et al., 2016). Nous avons obtenu des résultats similaires, en ce sens que la présence de symptômes dépressifs n'a pas eu d'incidence sur l'interprétation des stimuli non hostiles (contexte non hostile), alors qu'ils en ont eu pour les stimuli hostiles (contexte hostile). On peut ainsi affirmer que nos résultats appuient ceux de Smith en fournissant une observation convergente via les PRE.

Aussi, nous avons vérifié le lien entre les scores d'agressivité réactive et la mesure du BAIH par électroencéphalogramme, et bien que les données aient confirmé l'existence d'une relation significative entre ces deux variables pour la condition expérimentale hostile dans la région centrale droite, la relation s'est avérée non significative pour la condition non hostile dans la même région. L'absence de lien significatif peut être expliquée par une faible puissance statistique, et une augmentation de la taille de l'échantillon pourrait potentiellement remédier à ce fait. Une autre explication pourrait être que, bien que le BAIH ait été lié de manière significative à l'agressivité réactive dans des études précédentes (Crick & Dodge, 1996; Dodge & Coie, 1987; Bailey & Ostrov, 2008), ceux qui attribuent une intention hostile à des indices sociaux ambigus

ne réagissent pas tous à ces indices de manière agressive (Gagnon et al., 2017). Par exemple, selon Crick et Dodge (1994), pour qu'un comportement se produise, les informations sociales doivent être traitées par étapes successives, dont l'une consiste en la prise de décision par rapport à la réponse comportementale la plus appropriée à la situation. Cependant, ils soulignent que le processus décisionnel peut être influencé par plusieurs facteurs, comme la croyance en sa propre capacité à émettre un comportement de manière efficace. Par conséquent, une personne peut attribuer une intention hostile aux comportements ambigus des autres, mais se croire néanmoins incapable de réagir efficacement à la provocation perçue, réduisant par ce fait la probabilité qu'un comportement agressif se produise.

Ensuite, aucune association n'a été observée entre les scores sur la mesure auto-déclarée du BAIH (SIP-AEQ) et les scores d'hostilité, ce qui est incompatible avec l'association significative rapportée entre la mesure électrophysiologique du BAIH et les niveaux d'hostilité. Aussi, aucun lien n'a été remarqué entre la mesure électrophysiologique et la mesure auto-déclarée du BAIH. Cette constatation peut être expliquée par le fait que les deux méthodes reflètent des systèmes fonctionnels fondamentalement différents, soit un système de traitement plus automatique et un système plus délibéré, comme mentionné plus haut. Ici, la méthode EEG serait plus indicatrice de l'activation d'un système automatique lié aux connaissances implicites (Gagnon et al., 2016), tandis que la méthode auto-évaluative nécessiterait plutôt une activité cognitive délibérée et consciente pour émettre une réponse. Selon Glöckner & Betsch (2008), de nombreuses fonctions cognitives automatiques opèrent sans ou même en opposition à un contrôle cognitif délibéré, ce qui pourrait expliquer, du moins en partie, l'absence de lien significatif entre deux méthodes qui sont censées mesurer sensiblement le même construit, soit le BAIH.

Enfin, les groupes hostiles et non hostiles ont différé quant à leurs scores d'agressivité réactive, conformément aux résultats obtenus dans l'étude de Gagnon et al. (2017), ainsi qu'aux résultats des analyses corrélationnelles discutés plus haut.

Il faut noter que cette étude présente également quelques limites, notamment une puissance statistique insuffisante en raison d'un trop petit échantillon, ce qui pourrait expliquer l'absence de corrélations significatives observée entre certaines variables clés. Bien que de fortes tendances à la signification aient été constatées, l'absence de signification statistique a néanmoins limité la portée de nos conclusions. Bien entendu, une augmentation de la taille de l'échantillon pourrait remédier à ce problème. De plus, parce qu'un petit échantillon non clinique sélectionné sur l'hostilité a été utilisé, l'échantillon a montré des scores relativement faibles sur la sous-échelle de dépression BSI. Ainsi, on pourrait s'attendre à ce que des niveaux plus élevés de symptômes de dépression soient observés avec un plus grand échantillon de type clinique, ce qui pourrait valider davantage nos résultats et conclusions. Une autre limite provient de la composition de l'échantillon des participants. Ainsi, parce que les participants ont été sélectionnés en fonction de leurs scores de niveau d'hostilité, comme mentionné plus ci-haut, nous n'avons pas pu examiner les différences de groupe spécifiques concernant les personnes déprimées et non déprimées. Des études ultérieures utilisant des échantillons cliniques constitués d'individus déprimés pourraient être utiles dans l'étude des processus cognitifs sous-jacents observés chez des individus souffrant de dépression qui sont également agressifs. Ensuite, des études ont suggéré que la présence de colère, d'hostilité et d'irritabilité lors d'épisodes dépressifs pourrait être interprétée comme une preuve d'un trouble bipolaire non diagnostiqué ou existant sous le seuil diagnostique (Perlis et al., 2009). Comme nous n'avons pas pu examiner cette possibilité, nous ne sommes pas en mesure de réfuter l'idée que les résultats actuels pourraient

être mieux expliqués par la présence d'un trouble bipolaire masqué. Enfin, comme avec Smith et al. (2016), nous n'avons pas pu établir de relations causales entre les facteurs examinés en raison de la nature corrélacionnelle de notre étude.

Conclusion

En conclusion, bien que nos résultats aient été obtenus en utilisant un échantillon non clinique, limitant ainsi le nombre et la portée des recommandations, plusieurs implications importantes pour la recherche et la pratique peuvent néanmoins être tirées de cette étude. D'abord, notre étude ajoute à notre compréhension des mécanismes de perception impliqués dans la dépression en suggérant que les personnes qui présentent plus de symptômes dépressifs sont plus sensibles aux signaux hostiles dans les situations sociales, mais qu'elles ne perçoivent pas l'hostilité quand aucune n'est présente. Aussi, parce que le BAIH peut être considéré comme un facteur de maintien de la dépression en raison de ses effets délétères sur les relations sociales (Smith et al., 2016), les thérapies cognitives devraient porter une attention particulière à la présence de biais hostiles dans la dépression afin de réduire le risque de récurrence de la dépression chez les individus très agressifs. Ensuite, la présente étude est, à notre connaissance, la première à examiner le BAIH dans le contexte de la dépression en utilisant une méthodologie EEG, tout en soulignant que les PRE représentent un outil d'évaluation précis pour évaluer les processus cognitifs automatiques qui surviennent tôt dans les étapes du traitement de l'information. La force de cette méthodologie réside ainsi dans sa capacité à évaluer l'existence et le fonctionnement de schémas ou de réseaux associatifs activés inconsciemment. Par ce fait, cette étude met en évidence la nécessité pour les recherches futures sur les biais interprétatifs de prendre en compte d'importants facteurs théoriques et méthodologiques, dont l'un est l'utilisation

d'une méthodologie et d'un paradigme expérimental qui évaluent à la fois les processus automatiques et implicites.

Bibliographie

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., text rev.) Washington, DC: American Psychiatric Association.
- Bailey, C. A., & Ostrov, J. M. (2008). Differentiating forms and functions of aggression in emerging adults: Associations with hostile attribution biases and normative beliefs. *Journal of Youth and Adolescence*, 37(6), 713-722.
- Bargh, J. A. (1984). Automatic and conscious processing of social information. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition*, (Vol. 3, pp. 1-43). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper and Row.
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: New American Library.
- Benazzi, F., & Akiskal, H. (2005). Irritable-hostile depression: Further validation as a bipolar depressive mixed state. *Journal of Affective Disorders*, 84(2-3), 197-207.
- Berkowitz, L. (1977). Situational and personal conditions governing reactions to aggressive cues. In D. Magnusson & N. S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology* (pp.165-171). Hillsdale, NJ: Erlbaum.
- Boland, R. J., & Keller, M. B. (2002). Course and outcome of depression. In I. H. Gotlib, & C. L. Hammen (Eds.), *Handbook of depression* (pp. 43-60). New York, NY, USA: Guilford.

- Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development, 30*(5), 466-480.
- Chen, P., Coccaro, E. F., & Jacobson, K. C. (2012). Hostile attributional bias, negative emotional responding, and aggression in adults: Moderating effects of gender and impulsivity. *Aggressive Behavior, 38*(1), 47-63.
- Coccaro, E. F., Noblett, K. L., & McCloskey, M. S. (2009). Attributional and emotional responses to socially ambiguous cues: Validation of a new assessment of social/emotional information processing in healthy adults and impulsive aggressive patients. *Journal of Psychiatric Research, 43*, 915-925.
- Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*(1), 74-101.
- Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms on reactive and proactive aggression. *Child Development, 67*(3), 993-1002.
- Dearing, K. F., & Gotlib, I. H. (2009). Interpretation of ambiguous information in girls at risk for depression. *Journal of Abnormal Child Psychology, 37*(1), 79-91.
- 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.

- Derogatis, L. R. (1993). *BSI brief symptom inventory. Administration, scoring and procedures manual* (4th ed.). Minneapolis, MN:National Computer Systems.
- Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine, 13*(3), 595–605.
- Derogatis, L. R., & Savitz, K. L. (1999). The SCL-90-R, Brief Symptom Inventory, and Matching Clinical Rating Scales. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (pp. 679-724). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology, 53*(6), 1146-1158.
- Dodge, K. A. (1986). A social information processing model of social competence in children. In M. Perlmutter (Ed.), *Minnesota Symposium on Child Psychology* (Vol. 18, pp.77-125). Hillsdale, NJ: Erlbaum.
- Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and Psychopathology, 18*(3), 791-814.
- Drisdelle, B. L., Aubin, M., & Jolicoeur, P. (2016). Dealing with ocular artifacts on lateralized event-related potentials in studies of visual spatial attention and memory: ICA correction versus epoch rejection. *Psychophysiology, 54*, 83-99.

- Epps, J., & Kendall, P. C. (1995). Hostile attributional bias in adults. *Cognitive Therapy and Research, 19*(2), 159-178.
- Fava, G.A., Kellner, R., Munari, F., Pavan, L., Pesarin, F. (1982). Losses, hostility, and depression. *The Journal of Nervous and Mental Disease, 170*, 474-478.
- Fava, M., Alpert, J., Nierenberg, A. A., Ghaemi, N., O'Sullivan, R., Tedlow, J., . . . Rosenbaum, J. (1996). Fluoxetine treatment of anger attacks: A replication study. *Annals of Clinical Psychiatry, 8*(1), 7-10.
- Fava, M., Rosenbaum, J. F., Pava, J. A., McCarthy, M. K., Steingard, R. J., & Bouffides, E. (1993). Anger attacks in unipolar depression: I. Clinical correlates and response to fluoxetine treatment. *The American Journal of Psychiatry, 150*(8), 1158-1163.
- Fazel, S., Wolf, A., Chang, Z., Larsson, H., Goodwin, G. M., & Lichtenstein, P. (2015). Depression and violence: A Swedish population study. *The Lancet Psychiatry, 2*(3), 224-232.
- Fite, P. J., Stoppelbein, L., & Greening, L. (2009). Proactive and reactive aggression in a child psychiatric inpatient population. *Journal of Clinical Child and Adolescent Psychology, 38*(2), 199-205.
- Gadassi, R., & Rafaeli, E. (2015). Interpersonal perception as a mediator of the depression-interpersonal difficulties link: A review. *Personality and Individual Differences, 87*, 1-7.
- Gagnon, J., & Jolicœur, P. (2014). Mécanismes cognitifs et neurophysiologiques associés au trait "Urgence" de la personnalité: intégration théorique et empirique. *Insight Development: Social Sciences and Humanities Research Council of Canada.*

- Gagnon, J., McDuff, P., Daelman, S., & Fournier, S. (2015). Is hostile attributional bias associated with negative urgency and impulsive behaviors? A social-cognitive conceptualization of impulsivity. *Personality and Individual Differences, 72*, 18-23.
- Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Bessette, M., and Jolicœur, P. (2016). Neural mechanisms underlying attribution of hostile intention in nonaggressive individuals: an ERP study. *International Journal of Psychophysiology 110*, 153-162.
- Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Brochu, A. F., Bessette, M., & Jolicœur, P. (2017). An ERP study on hostile attribution bias in aggressive and nonaggressive individuals. *Aggressive Behavior, 43*(3), 217-229.
- Gagnon, J., & Rochat, L. (2017). Relationships between hostile attribution bias, negative urgency, and reactive aggression. *Journal of Individual Differences, 38*(4), 211-219.
- Glöckner, A., & Betsch, T. (2008). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making, 3*(3), 215-228.
- Good, M. I. (1978). Primary affective disorder, aggression, and criminality. *Archives of General Psychiatry, 35*, 954-960.
- Gotlib, I. H., & Krasnoperova, E. (1998). Biased information processing as a vulnerability factor for depression. *Behavior Therapy, 29*(4), 603-617.
- Hammen, C. (1991). Generation of stress in the course of unipolar depression. *Journal of Abnormal Psychology, 100*(4), 555-561.

- Judd, L.L., Schettler, P.J., Coryell, W., Akiskal, H.S., Fiedorowicz, J.G. (2013). Overt Irritability/Anger in Unipolar Major Depressive Episodes: Past and Current Characteristics and Implications for Long-term Course. *JAMA Psychiatry*, 70(11), 1171-1180.
- Kennedy, H. G. (1992). Anger and irritability. *The British Journal of Psychiatry*, 161, 145-153.
- Kermani, E. J. (1981). Violent psychiatric patients: A study. *American Journal of Psychotherapy*, 35, 215-225.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., et al. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association*, 289(23), 3095-3105.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621-647.
- Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, 11, 539-550.
- Leuthold, H., Filik, R., Murphy, K., & Mackenzie, I. G. (2012). The online processing of socio-emotional information in prototypical scenarios: Inferences from brain potentials. *Social Cognitive and Affective Neuroscience*, 7, 457-466.
- Linder, J., Werner, N., & Lyle, K. (2010). Automatic and controlled social information processing and relational aggression in young adults. *Personality and Individual Differences*, 49(7), 778-783.

- 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.
- Luck, S.J. (2014). *An Introduction to the Event-Related Potential Technique* (2nd ed.). Cambridge, MA: The MIT Press.
- Maiuro, R. D., O'Sullivan, M. J., Michael, M. C., & Vitaliano, P. P. (1989). Anger, hostility, and depression in assaultive vs. suicide-attempting males. *Journal of Clinical Psychology*, *45*(4), 531-541.
- Milich, R., & Dodge, K. A. (1984). Social information processing in child psychiatric populations. *Journal of Abnormal Child Psychology*, *12*(3), 471-489.
- Miller, J. D., & Lynam, D. (2006). Reactive and proactive aggression: Similarities and differences. *Personality and Individual Differences*, *41*, 1469-1480.
- Mogg, K., Bradbury, K. E., & Bradley, B. P. (2006). Interpretation of ambiguous information in clinical depression. *Behaviour Research and Therapy*, *44*(10), 1411-1419.
- Moreno, J. K., Selby, M. J., Fuhriman, A., & Laver, G. D. (1994). Hostility in depression. *Psychological Reports*, *75*(3, Pt 1), 1391-1401.
- Nasby, W., Hayden, B., & DePaulo, B. M. (1980). Attributional bias among aggressive boys to interpret unambiguous social stimuli as displays of hostility. *Journal of Abnormal Psychology*, *89*(3), 459-468.
- Novaco, R. W. (1978). Anger and coping with stress: Cognitive behavioral intervention. In J. P. Foreyt & D. P. Rathjen (Eds.), *Cognitive behavioral therapy: Research and application* (pp.135-173). New York, NY: Plenum Press.

- Orobio de Castro, B., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2002). Hostile attribution of intent and aggressive behavior: A meta-analysis. *Child Development, 73*(3), 916-934.
- Pan, H. S., Neidig, P. H., & O'Leary, K. D. (1994). Male-female and aggressor-victim differences in the factor structure of the Modified Conflict Tactics Scale. *Journal of Interpersonal Violence, 9*(3), 366-382.
- Pedrelli, P., Nyer, M., Holt, D., Bakow, B.R., Fava, M., Baer, L., ... Farabaugh, A. (2013). Correlates of irritability in college students with depressive symptoms. *The Journal of Nervous and Mental Disease, 201*, 953-958
- Perlis, R. H., Fava, M., Trivedi, M. H., Alpert, J., Luther, J. F., Wisniewski, S. R., & Rush, A. J. (2009). Irritability is associated with anxiety and greater severity, but not bipolar spectrum features, in major depressive disorder. *Acta Psychiatrica Scandinavica, 119*(4), 282-289.
- Pitt, B. (1968). "Atypical" depression following childbirth. *The British Journal of Psychiatry, 114*(516), 1325-1335.
- Raine, A., Dodge, K. A., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., Stouthamer-Loeber, M., & Liu, J. (2006). The reactive proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32*, 159-171.
- Riley, W. T., Treiber, F. A., & Woods, M. G. (1989). Anger and hostility in depression. *Journal of Nervous and Mental Disease, 177*(11), 668-674.

Sharbrough, F., Chatrian, G.-E., Lesser, R. P., Luders, H., Nuwer, M., & Picton, T. W. (1991).

American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *8*, 200-202.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information

processing: II. Perceptual learning, automatic attending and a general theory.

Psychological Review, *84*(2), 127-190.

Smith, H. L., Summers, B. J., Dillon, K. H., Macatee, R. J., & Coughlen, J. R. (2016). Hostile

interpretation bias in depression. *Journal of Affective Disorders*, *203*, 9-13. doi:

10.1016/j.jad.2016.05.070

Snaith, R. P., & Taylor, C. M. (1985). Irritability: Definition, assessment and associated factors.

The British Journal of Psychiatry, *147*, 127-136

Statistics Canada. (2012). Table: 105-1101 - *Canadian Community Health Survey Share File -*

Mental Health (CCHS), CANSIM (database). Retrieved from

<https://www150.statcan.gc.ca/t1/tb11/en/tv.action?pid=1310046501>

Todorov, A., & Bargh, J. A. (2002). Automatic sources of aggression. *Aggression and Violent*

Behavior, *7*, 53-68.

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The

neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580–

591.

Zelli, A., Huesmann, L. R., & Cervone, D. (1995). Social inference and individual differences in aggression: Evidence for spontaneous judgments of hostility. *Aggressive Behavior*, *21*, 4015-4017.

Annexe : Article scientifique publié dans Journal of Affective Disorders

Association between depression and hostile attribution bias in hostile and non-hostile individuals: an ERP study.

Abstract

Background: Hostility and aggression have been found to be highly prevalent among depressed patients and are associated with higher comorbidity and illness severity levels. Although negative interpretation biases are a fundamental element of cognitive models of depression, few studies have examined the specific biases in information processing, mainly the hostile attribution bias, found in hostile individuals who present depressive symptoms.

Method: Using pre-collected data from a sample of 72 (male=41,6%, female=58,3%) undergraduate and community-based hostile (n=26) and non-hostile (n=46) adult participants, the authors aimed to examine the association between depression and the hostile attribution bias by determining whether depression level scores were uniquely related to electrophysiological measures of the hostile attribution bias.

Results: The hostile group showed higher measured levels of depression and reactive aggression compared to the non-hostile group. Also, depression scores were significant predictors of the N400 effect in the non-hostile task condition, while reactive aggression was not, whereas in the hostile condition, the overall model was significant, with depression and reactive aggression levels both showing strong trends towards significance.

Limitations: A small sample size limited the scope of our conclusions. Also, sample selection prevented us from examining specific group differences regarding the hostile attribution bias in depressed and non-depressed groups.

Conclusion: Clinical and research implications include the necessity to apply cognitive

restructuring techniques to counter biased interpretation processes in settings where depression and aggression intersect, and the need to consider alternatives to self-evaluative methodologies.

Key words: Depression; Aggression; Hostile attribution bias; EEG; ERP; N400

Introduction

Aggression and Depression

Unipolar depression is a heterogeneous disease composed of several clinical subtypes, each presenting its symptomatology and distinct long-term prognosis. Researchers thus suggest that identifying explicit phenotypes linked to greater prognosis severity may positively impact overall disease incidence, by guiding the development of treatment strategies better adapted to specific subtypes of the disorder (Pedrelli et al., 2013). One such phenotype, characterized here by high prevalence levels of irritability, hostility, and anger, has been identified in multiple cross-sectional studies using an array of conceptual definitions and measurement tools (Judd et al., 2013). For example, 37% of Benazzi and Akiskal's study sample who experienced a unipolar major depressive disorder, expressed high levels of irritability and hostility. This form of 'irritable-hostile' depression was associated with earlier mood disorder onset, younger age, atypical depressive features, hypomanic symptoms, and a family history of bipolar disorder (Benazzi & Akiskal, 2005, in Judd et al., 2013). Although irritability and anger are no longer central to the diagnosis of unipolar depression in adults since the DSM-III (American Psychiatric Association, 1987) was revised, these symptoms are highly prevalent, and those presenting them suffer from a more chronic, comorbid and severe form of the disorder (Judd et al., 2013).

Interpretative Biases in Depression

Not only is depression a highly prevalent psychiatric disorder – depressive disorders affect 4,7% of Canadians annually and represent 10,9% of all psychiatric hospitalizations (Canadian Community Health Survey on mental health, Statistics Canada, 2012) – it is also an extremely recurrent one whose incidence may be explained in part by stable vulnerability factors that expose certain people to repeated episodes (Gotlib & Krasnoperova, 1998; Dearing et al., 2009). Beck (1967, 1976, in Gotlib & Krasnoperova, 1998) was the first to emphasize the importance of maladaptive cognitive processes as risk factors for depression and formulated a theory of depression that largely attributed its appearance to cognitive biases in the form of negative schemas through which the environment is filtered. As such, depressive people selectively focus on negative stimuli, filter out positive stimuli and perceive negative or neutral information in a more negative way (Gotlib & Krasnoperova, 1998; Dearing et al., 2009). Furthermore, depressed individuals are more likely to interpret ambiguous homophones and sentences as threatening, increasing the likelihood that those predisposed to mistrust and threat perception will experience more hostility, and consequently, act aggressively towards others. Researchers, therefore, suggest that major depressive disorder with irritability may promote a hostile worldview and increased interpersonal hostility (Smith et al., 2016).

The Hostile Attribution Bias

According to earlier cognitive personality theorists, the cognitive appraisal of a cue is what determines its ability to evoke a response, not the cue per se (Nasby et al., 1980). Consequently, aversive stimuli do not automatically generate aggressive outcomes, but rather, cognitive processes such as appraisals, expectations and private speech, play a mediating role in the relationship between aversive stimuli and the subsequent aggressive responses. For instance, for

an aggressive reaction to occur, one must attribute an aggressive or hostile intent to an aversive stimulus for it to serve as a provocative cue (Berkowitz, 1977, in Nasby et al., 1980).

As a result, in addition to negative or depressive interpretation biases, distorted interpretative processing may also lead to what Milich and Dodge (1984) call a hostile attributional bias (HAB), a key element in the etiology, development, and persistence of aggressive and antisocial behavior issues over time (Chen et al., 2012). Safer (2009) defines hostility as a directed and maintained attitude of ill-will, while the hostile attribution bias represents the tendency of aggressive individuals to interpret others' intent as hostile when social cues are ambiguous or fail to indicate a clear intent, and is generally displayed through reactive forms of aggression, which is an angry, defensive response to frustration, provocation or perception of threat (Bailey & Ostrov, 2008).

Theoretical Basis for the Current Study

Although depression is frequently linked to various negative interpretative biases that impact the course of the disorder, HAB has thus far not been formally conceptualized in the literature as a maladaptive cognitive mechanism by which depressed individuals who manifest hostility, irritability, and anger, process social information. However, one study that did examine the link between HAB and depression showed that depressed subjects interpreted ambiguous interpersonal scenarios as hostile more frequently than did non-depressed subjects and that depression predicted HAB independently of trait anger (Smith et al., 2016).

The current study was designed to corroborate the observations made by Smith and her colleagues and was done using a unique cognitive neuroscience approach. An important scientific contribution of this study, therefore, lies in our use of electroencephalographically (EEG) recorded event-related potentials (ERP) to assess HAB. ERPs effectively circumvent

conscious, deliberate and strategic control mechanisms which facilitate or inhibit information processing, by providing a precise and continuous measure of the automatic and implicit cognitive processes involved in the initial steps of social information processing, such as the interpretation of information (Gagnon et al., 2016). The ERP component used for our study is the N400, which has a peak latency of approximately 400 milliseconds and appears typically, but not exclusively, in response to a violation of the subject's contextual semantic expectations (Gagnon et al., 2016; Kutas & Federmeier, 2011).

To measure HAB, Gagnon et al. (2016) developed the Hostile Expectancy Violation Paradigm, an ERP task comprising of social scenarios made up of three sentences that solicit the N400 component as an indicator of the hostile bias (see Table I). The task is described in detail below.

Study objective and hypotheses

The objective of this study is to examine the relation between depressive symptoms and HAB using electrophysiological measures of hostile bias. First, we anticipated (1) that depression level scores would positively correlate with HAB self-reported measures, per Smith's findings. We also anticipated (2) that the hostile group would present higher depression level scores than the non-hostile group, in line with the literature on the association between depression and hostility. Next, we predicted (3) that depression level scores would be uniquely related to HAB, as with Smith's findings, and therefore independent of reactive aggression scores, and that this association would exist in both the hostile and the non-hostile conditions of the hostile expectancy violation paradigm. More precisely, we hypothesized (3a) that depression level scores would correlate negatively with the N400 effect in the non-hostile mismatch condition (hostile context and non-hostile intention), after adjusting for reactive aggression. That

is, participants presenting greater depressive symptom levels would show greater levels of surprise (stronger violation of hostile expectation) following a critical word revealing the non-hostile intention behind the character's provocative ambiguous behavior in a hostile context. We also hypothesized (3b) that higher depression level scores would correlate positively with the N400 effect in the hostile mismatch condition, after adjusting for reactive aggression. Here, if depressive symptom levels are independently related to the hostile bias, those presenting greater depression levels should show less surprise (i.e. smaller violation of hostile expectation as shown by a smaller N400 amplitude) following a critical word revealing the hostile intention behind the character's provocative ambiguous behavior in a non-hostile context.

Methods

Participants

The data used for this project were gathered from 87 participants as part of a prior study on aggression and impulsivity entitled "Cognitive and neurophysiological mechanisms associated with the "Urgency" personality trait: theoretical and empirical integration." (Gagnon & Jolicoeur, 2014). Out of the 10 participants who were removed from the study, four were from the student population and six from the clinical population because data on depression levels were not available. From that number, five participants, all recruited from a clinical population, were excluded to avoid potential selection bias, as these clinical subjects were more likely to present depressive symptomatology. The final experimental group was composed of 30 male (41,6%) and 42 female (58,3%) participants (N=72, mean age=26,9 years, SD=8,1) that were recruited from three different populations: (i) undergraduate university students; (ii) the general population using newspaper, Kijiji and community bulletin ads; and (iii) patients from a Montreal-area personality disorders clinic. Those meeting the selection criteria were told that the study would

take approximately three hours to complete, for which they would receive a forty-dollar compensation. Respondents were also directed to refrain from consuming alcohol 24 hours before each session, and recreational drugs one week before each session.

Respondents were excluded from participation if they (i) were not between 18 and 55 years old, (ii) did not have French as their mother tongue, (iii) had less than a sixth-grade level of education, (iv) reported a previous serious head injury, or (v) had a history of psychosis.

For participant selection purposes, the Brief Symptoms Inventory (BSI-II; Derogatis, 1993) Hostility Dimension subscale was administered to each potential participant during the initial telephone interview, while the entire inventory was completed online via the Echo online questionnaire website (www.questionnaireenligne.ca). The Hostility Dimension subscale is comprised of five items (e.g. item 40 - Having urges to beat, injure, or harm someone) that assess the cognitive, emotional and behavioral components that are characteristically associated with hostility, using a 5-point Likert scale ranging from 'not at all' (0) to 'extremely'(4). Respondents scoring above the standard T-score of 65 in the adult nonpatient normative sample were assigned to the hostile group (26 participants; male=14, female=12), whereas those scoring below were assigned to the nonhostile group (46 participants; male=16, female=30). The Cronbach's coefficient alpha for this subscale is .81 (Derogatis & Melisaratos, 1983)

Materials and Apparatus

Hostile Expectancy Violation Paradigm

HAB was assessed using the Hostile Expectancy Violation paradigm developed by Gagnon et al. (2016) and consists of 320 scenarios depicting everyday social interactions (see Table I). Each scenario is comprised of three sentences: (1) an initial sentence establishing a hostile or non-hostile social context, followed by (2) a second sentence describing a fictitious character's

ambiguous behavior, and (3) a third sentence in which a final target word informs the reader of the character's hostile or non-hostile intention behind the ambiguous behavior. In this task, each scenario has a semi identical version that shares the same (2) ambiguous behavior sentence and the same (3) intention revealing sentence, but differs only regarding the initial (1) context sentence, with one version depicting a hostile context, and the second, a non-hostile context. The nature of the (1) context (hostile or non-hostile) will either match or mismatch the nature of the character's (3) intention. Also, the experimental task condition (hostile or non-hostile) is dependent on the hostile or non-hostile nature of the (3) intention.

Self-Report Measures

Brief Symptom Inventory Depression Dimension subscale (BSI-II; Derogatis, 1993). For this study, the BSI-II Depression subscale was used to determine the participants' depression levels. This subscale represents multiple dimensions of clinical depression and consists of six items (e.g., item 18 - Feeling no interest in things) that rate the degree to which patients experience depressive mood, loss of interest, vulnerability to criticism, loneliness, worthlessness, hopelessness, and thoughts of suicide, with higher scores representing more severe symptomatology. The alpha coefficient for this subscale is .85 (Derogatis & Melisaratos, 1983).

Reactive-Proactive Aggression Questionnaire (RPQ; Raine et al., 2006). With the RPQ, participants must rate their reactions to fictional aggressive scenarios using a three-point Likert scale ranging from never (0) to often (2). 23 items are divided into two subscales measuring both reactive aggression (11 items - e.g. How often have you reacted angrily when provoked by others?) and proactive aggression (12 items - e.g. How often have you threatened and bullied someone?). The alpha indexes for the reactive subscale are .84, and .86 for the proactive subscale

(Raine et al., 2006). For the current study, only the reactive aggressive subscale of the French adapted version (Gagnon & Rochat, 2017) was used.

Social Information Processing—Attribution and Emotional Response

Questionnaire (SIP—AEQ; Coccaro et al., 2009). As a measure of HAB, this questionnaire contains eight vignettes that depict socially ambiguous scenarios in which a physically (e.g., being struck by someone) or relationally (e.g., being rejected by someone) aversive action is directed towards a fictional character. The participant must determine the likelihood of the ambiguous behavior using a four-point Likert scale ranging from 0 (not at all likely) to 3 (very likely). These items are then followed by two other Likert-scaled items that assess negative emotional responses. The SIP—AEQ measures direct hostile intent, indirect hostile intent, instrumental non-hostile intent, and neutral or benign intent. For this study, the average of the combined scores for the direct and indirect hostile intent forms of the French adapted version (Gagnon et al., 2015) of this questionnaire was used. Also, alpha coefficients for the hostile, instrumental, and benign intent scales are .82, .57, and .66, respectfully (Coccaro et al., 2009).

Procedure

Before giving their verbal and written consent, participants received information about the EEG procedure and the experimental task. They were instructed to read all materials carefully and to avoid making unnecessary movements during the EEG testing phase. Participants were told that the experimental task consisted of them reading scenarios depicting everyday social encounters with fictional characters while imagining that they were taking part in such encounters. They were also instructed to imagine the thoughts and feelings experienced by the characters and the intent behind the characters' ambiguous behaviors.

Participant testing was performed in an electrically shielded booth (Faraday cage) in which the ambient light level was kept low. The textual stimuli were presented at the center of a 17-inch computer monitor, in white 14-point Helvetica font on a black background. Also, viewing distance from the monitor was maintained at 23 inches (57 cm) throughout the experiment with a chin rest, which resulted in three characters subtending 1° of visual angle.

Four initial practice trials were followed by ten experimental blocks consisting of 17 trials (16 experimental scenarios and one filler scenario) each. The experimental blocks were each divided by a timed break, the duration of which was determined by the participant. In each trial, the first two scenario sentences – the (1) context sentence and the (2) ambiguous behavior sentence – were presented and displayed until the participant pressed the space bar to indicate that the sentence was read, but only after a minimum display time of 1500ms. A blank screen interval was then shown for 500ms, followed by a fixation point presented at the center of the screen for 1000ms. Next, the third (3) intention revealing sentence was shown one word at a time for 300ms, followed by a 200ms blank interval. The next (1) context sentence was then displayed after a 2000ms interval during which the participant was instructed to maintain his or her gaze on the fixation point. A filler sentence was included among the 17 trials in each block, followed by a true or false comprehension question, which resulted in a mean correct response rate of 88%. Finally, participants were asked to complete computerized versions of the RPQ and the SIP-AEQ following the EEG procedure.

EEG Recording

Using the BioSemi Active Two system, the electroencephalogram (EEG) signal was recorded using 64 Ag / AgCl electrodes mounted on an elastic cap according to the International 10-10 System (Sharbrough et al., 1991), and referenced to the average of the left and right

mastoids. Horizontal eye movements were measured using a horizontal electrooculogram (HEOG), recorded as the difference in voltage between the electrodes placed laterally to the external canthi, while eye blinks were detected with a vertical electrooculogram (VEOG), recorded here as the difference in voltage between the two electrodes placed above and below the left eye.

Next, EEG and EOG signals were both digitized at 512Hz and low-pass filtered at 134Hz, while the continuous signal was high-pass filtered at .01Hz and low-pass filtered at 30Hz during offline signal analysis. The EEG was epoched from 200ms before target onset (0ms) to 1000ms following target onset and was baseline corrected using the mean from -200ms to 0ms. Also, electro ocular artifact reduction and data correction were both achieved with an independent component analysis (ICA) statistical procedure (see Drisdelle, Aubin, & Jolicoeur, 2016). Trials containing eye blinks (i.e., VEOG > 50 μ V within a 150ms interval) and large horizontal eye movements (i.e., HEOG > 35 μ V within a 300ms interval) were excluded from the analysis, while trials showing EEG deflections greater than 100 μ V during the predetermined segmentation window on at least one of the 64 electrodes were selected for further analysis. If seven or fewer channels were detected in any given trial, spherical spline interpolation was used to interpolate these channels from neighboring channels, whereas trials were rejected if more than seven channels were found to contain artifacts.

EEG Data Analysis

EEG data analysis was done with MATLAB with functions from the EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes. The EEG signal at each electrode site was averaged separately for the four experimental conditions (hostile-match, hostile-mismatch, non-hostile-match, non-hostile-mismatch) and was time-locked to target word

onset. After an initial visual assessment of the global ERP, a significant negative ERP deflection was identified between 450ms and 650ms, which is within the literature-based N400 effect timeframe (Kutas & Federmeier, 2011). Lateral electrode sites were pooled into six regions identified by either the left or right hemisphere and by the anterior, central, or posterior scalp locations. The left hemisphere was divided into three regions that were identified as follows: left-anterior (AF3, AF7, F1, F3, F5, F7, FT7, FC1, FC3, FC5), left-central (TP7, T7 C1, C3, C5, CP1, CP3, CP5), and left-posterior (P1, P3, P5, P7, PO3, PO7, O1), while the right hemisphere was divided into three equivalent regions that were identified by their homologous hemispherical electrodes. The original data analyses performed by Gagnon et al. (2016) showed an N400 effect component following a mismatch between the participant's expectation of a hostile intention and the presented non-hostile intention, as revealed by the target word. Component amplitude was shown to be larger for the non-hostile condition compared to the hostile condition, for the hostile group compared to the non-hostile group, and over the central and posterior electrode sites compared to the anterior sites.

For the current study, we first tested the hypothesis that depression would interact with the N400 effect independently of reactive aggression by conducting a repeated-measures ANOVA with Condition (hostile, non-hostile), Consistency (match, mismatch), Hemisphere (left, right) and Location (anterior, central, posterior) as the variables of interest, as did Gagnon et al. in their 2016 study, but with depression scores as a covariate and reactive aggression scores as a control variable. Subsequent decomposition of variance analysis was performed to isolate specific interactions involving Condition, Consistency, and Depression.

Next, the regression analyses would enable us to test our main hypothesis that depression level scores were uniquely related to the hostile attribution bias, as measured by the N400 effect,

after adjusting for reactive aggression, and to examine the differential effects of the hostile and non-hostile conditions of the ERP task on the N400 amplitude.

Results

T-tests

An independent sample t-test was first performed and the results showed that the scores on the BSI Depression subscale were significantly higher for the hostile group ($M=.32$, $SD=.12$, $t(70)=-4.97$, $df=70$, $p<.01$) than for the non-hostile group ($M=.15$, $SD=.14$). Furthermore, a second independent samples t-test was performed to confirm the results obtained by Gagnon et al. (2017) regarding group differences in levels of reactive aggression. In their study, t-test analysis confirmed that the hostile group ($M=1.19$; $SD=.39$; $t(85)=7.16$, $p<.001$) scored significantly higher on the RPAQ reactive aggression subscale than the non-hostile group ($M=.69$; $SD=.26$). Our findings also showed that the scores on the RPQ reactive aggression scale were significantly higher for the hostile group ($M=1.11$, $SD=.40$, $t(70)=-4.63$, $df=70$, $p<.01$) than for the non-hostile group ($M=.71$, $SD=.31$). (See Table 2).

Correlations

Results from the correlational analyses are reported in Table 3

Analyses of variance

Because we were mainly interested in examining hostile expectancy violations in relation to depression, the main effects and interactions involving the Consistency factor and depression were the focus of these analyses. First, the repeated measures ANOVA revealed a significant main Consistency effect (Wilks' Lambda=.93, $F(1, 69)=5.43$, $p=.023$), meaning that the waveforms were significantly more negative in the mismatch condition than in the match condition, which in turn confirms the presence of the N400 expectancy violation effect. More

importantly, a significant interaction was observed between Condition, Consistency, Hemisphere, Location, and BSI depression scores (Wilks' Lambda=.91, $F(2,68)=3.33$, $p=.042$). Decomposition of variance analysis (Condition x Consistency) for the six regions (Hemisphere by Location) revealed that BSI depression score was significantly associated with more negative waveforms in the mismatch condition measured in the right central site (Wilks' Lambda=.90, $F(1,69)=7.45$, $p=.008$), and right anterior site (Wilks' Lambda=.94, $F(1,69)=4.73$, $p=.033$). Moreover, decomposition analysis for each Condition (hostile, non-hostile) revealed a significant interaction between Consistency and depression in the non-hostile condition (Wilks' Lambda=.90, $F(1,74)=8.21$, $p=.005$).

Multiple Regression

We first tested for multicollinearity in the regression models and found that the variance inflation factors ranged between 1.022 and 1.214, indicating that collinearity was not a problem in these models. Next, following the results of the repeated-measures ANOVA in which both the right central and right anterior regions showed significant interactions involving factors Condition, Consistency, and Depression, four standard multiple regressions were performed to measure the combined effect of depression, reactive aggression, age, and sex on the N400 amplitude for both the non-hostile and hostile conditions. The first analysis, which was performed using the right central N400 effect size for the non-hostile condition as the dependent variable, resulted in a significant overall model ($F=3.16$, $p=.02$), in which depression was a significant predictor of the N400 effect ($\beta=-.27$, $t=-2.16$, $df=67$, $p=.04$), whereas reactive aggression ($p=.41$), sex ($p=.70$) and age ($p=.09$) were not (See Table 4). Further, the multiple regression performed with the right central N400 effect size for the hostile condition revealed a significant overall model ($F=2.80$, $p=.03$), but in which depression ($p=.10$), reactive aggression

($p=.07$), sex ($p=.57$) and age ($p=.72$) were all found to be nonsignificant (See Table 5). Finally, multiple regressions were performed using the right anterior locations for the non-hostile ($F=2.46$, $p=.054$) and hostile conditions ($F=1.00$, $p=.42$), but neither showed significant overall models.

Discussion

Drawing from Smith's study, in which the authors concluded that HAB is a separate construct from trait anger and that both characterize depression in distinctive ways, the purpose of this study was to examine the relation between depressive symptoms and HAB using electrophysiological measures of the hostile bias, namely, the N400 ERP component. Our decision to utilize the N400, and the paradigm to measure it, as a marker of the HAB is based on a certain understanding of the component's characteristics. Although the N400 is commonly interpreted as a reflection of a violation of semantic expectations in the context of language comprehension, semantic anomalies are neither necessary nor enough to evoke an N400. The neural signature of the N400 is therefore not related to a simple representation of the violation of an arbitrary or over-learned pattern of cognition, but rather, is a fundamental aspect in the processing of what is significant (Kutas & Federmeier, 2011), with higher amplitudes of the N400 being typically observed following less expected, or less significant, stimuli.

Regarding the results obtained for the current study, no relation was found between the self-reported measures of HAB and depressive symptom levels, which is contrary to Smith's findings and our hypothesis. On the other hand, the results did show that the hostile and non-hostile groups differed markedly regarding their depression symptom levels, as expected, with the hostile group showing higher symptom levels than the non-hostile group. This is supported by research findings that suggest that overtly hostile, aggressive, and violent individuals often

manifest depressive symptomatology that is often chronic, severe, and comorbid (Faizel et al., 2015; Judd et al., 2013).

Next, the repeated measures ANOVA revealed that the amplitude waveforms were significantly more negative in the mismatch condition than in the match condition, which is indicative of an N400 hostile expectancy violation effect. Also, for the mismatch condition, a more negative amplitude was observed in the central and anterior regions of the right hemisphere, as well as in the non-hostile condition, when depression was factored in as a covariate. Here, the results showed that depression still interacted with all of the aforementioned factors and that this interaction was most significant in the right central region, which is in line with research literature suggesting that the N440 effect is largest over centroparietal sites, with a slightly right hemisphere bias for written words in sentences (Kutas & Federmeier, 2011). To explain the presence of an ERP component in the right anterior region, depression might possibly be associated with additional effort to integrate non-hostile stimuli in a hostile context (Van Berkum et al., 2008), but further studies would be needed to account for this result.

All in all, these results supported the notion that depression did indeed interact with the aforementioned variables and served to guide our subsequent regression analyses by identifying which scalp regions warranted further analysis. Although the overall regression model for the right anterior region turned out to be not significant, our findings for the right central region supported the hypothesis that depressive symptom levels were negatively associated with the N400 amplitude for the non-hostile condition, after adjusting for reactive aggression. Hence, the more depressed one is, the more one expects hostile intentions underlying an ambiguous behavior in a hostile context, independently of reactive aggression. In this case, when a hostile context was followed by a non-hostile intention, the N400 was evoked in all participants due to a

violation of hostile expectations. However, the component amplitude was greater in those who presented higher depression levels, in that the hostile context served as a stronger primer towards expecting a subsequent hostile intent. From this observation, we may posit that individuals with higher levels of depression react more strongly to the hostile context, due to a possible hypersensitivity towards negative social cues.

Next, our hypothesis that depressive symptom levels were positively related to the N400 amplitude for the hostile condition, after adjusting for reactive aggression, was not confirmed. The overall model was found to be significant, although neither depression levels nor reactive aggression scores correlated significantly with the N400 amplitude. However, reactive aggression showed a strong trend toward significance, as did depression. This result is consistent with the Gagnon et al. (2016) results showing that the N400 effect in the hostile condition was associated with reactive aggression.

In the Smith study, the authors reported a significant positive correlation between depression scores and the hostile rating sub-scale of the self-reported measure of HAB, while no significant correlation was found between depression scores and the benign evaluation subscale (Smith et al., 2016). We obtained similar results in our study, in that depressive symptoms did not influence the interpretation of stimuli in the non-hostile context but did so regarding the hostile context. This validates Smith's idea, detailed below, that HAB can be interpreted as a maintenance factor in depression because of its deteriorating effect on social relationships.

However, this study also has some limitations, one of which is the absence of significant correlations between certain key variables, which may partially be explained by the small sample size, another limitation of our study. Although strong tendencies towards significance were at times observed, the lack of statistical significance nevertheless limited the scope of our

conclusions. Of course, an increase in sample size could remedy this problem. Also, because a small, nonclinical sample selected on hostility was used, the sample showed relatively low scores on the BSI depression subscale. Greater depression symptom levels could be expected in a larger clinical sample, which could validate our results and conclusions further. Another limitation comes from the participant sample selection. Because participants were selected based on their hostility level scores, we were unable to examine specific group differences regarding HAB in depressed and non-depressed individuals. Next, studies have suggested that anger, hostility, and irritability during depressive episodes may be interpreted as indicators of an undiagnosed or subthreshold bipolar disorder (Perlis et al., 2009), which we did not adjust for. Furthermore, the high intercorrelation between BSI hostility and depression scores (.58) indicates that the dimensions may share a significant proportion of variance and therefore not be as independent as expected. Using a different depression measure would be suggested for future research. Finally, as with Smith et al. (2016), we were unable to establish any causal relationship between the examined factors due to the correlational nature of our study.

Despite these limitations, our study suggests that depressive individuals are more sensitive to hostile signals in social situations but will not perceive hostility when none is present, which adds to our understanding of perception mechanisms in depression. Also, the current study is, to our knowledge, the first to examine HAB in the context of depression using an EEG methodology, while emphasizing that ERPs represent a precise assessment tool for evaluating automatic cognitive processes. The strength of this methodology hence lies in its ability to accurately assess the existence and functioning of unconsciously-activated schemas or associative networks, in a manner self-evaluative measures are incapable of accomplishing,

mainly because these measures solicit more deliberate and controlled mechanisms when appraising implicit cognitive processes (Gotlib & Krasnoperova, 1998).

Conclusion

In conclusion, although the results from the present study were obtained using a nonclinical sample, thus limiting the range of recommendations that can ultimately be made, several important implications for research and practice can nevertheless be ascertained. First, because HAB can be regarded as a maintenance factor for depression due to its deteriorating effects on relationships, which can lead to social isolation and increased feelings of self-worthlessness (Smith et al., 2016), cognitive therapies should contemplate targeting hostile biases in depression, to improve treatment outcomes and reduce the risk of depression recurrence in highly aggressive individuals. Furthermore, the current study highlights the need for future research on interpretative biases to consider important theoretical and methodological factors, one of which is the use of a methodology and experimental paradigm that both assess automatic and implicit processes, and control for biases associated with self-evaluative measures.

References

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., text rev.) Washington, DC: American Psychiatric Association.
- Bailey, C. A., & Ostrov, J. M. (2008). Differentiating forms and functions of aggression in emerging adults: Associations with hostile attribution biases and normative beliefs. *Journal of Youth and Adolescence*, 37(6), 713-722.
- Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper and Row.
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: New American Library.
- Benazzi, F., & Akiskal, H. (2005). Irritable-hostile depression: Further validation as a bipolar depressive mixed state. *Journal of Affective Disorders*, 84(2-3), 197-207.
- Berkowitz, L. (1977). Situational and personal conditions governing reactions to aggressive cues. In D. Magnusson & N. S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology* (pp.165-171). Hillsdale, NJ: Erlbaum.
- Chen, P., Coccaro, E. F., & Jacobson, K. C. (2012). Hostile attributional bias, negative emotional responding, and aggression in adults: Moderating effects of gender and impulsivity. *Aggressive Behavior*, 38(1), 47-63.
- Coccaro, E. F., Noblett, K. L., & McCloskey, M. S. (2009). Attributional and emotional responses to socially ambiguous cues: Validation of a new assessment of social/emotional

- information processing in healthy adults and impulsive aggressive patients. *Journal of Psychiatric Research*, 43, 915-925.
- Dearing, K. F., & Gotlib, I. H. (2009). Interpretation of ambiguous information in girls at risk for depression. *Journal of Abnormal Child Psychology*, 37(1), 79-91.
- 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.
- Derogatis, L. R. (1993). *BSI brief symptom inventory. Administration, scoring and procedures manual* (4th ed.). Minneapolis, MN:National Computer Systems.
- Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, 13(3), 595–605.
- Drisdelle, B. L., Aubin, M., & Jolicoeur, P. (2016). Dealing with ocular artifacts on lateralized event-related potentials in studies of visuospatial attention and memory: ICA correction versus epoch rejection. *Psychophysiology*, 54, 83-99.
- Faizel, S., Wolf, A., Chang, Z., Larsson, H., Goodwin, G. M., & Lichtenstein, P. (2015). Depression and violence: A Swedish population study. *The Lancet Psychiatry*, 2(3), 224-232.
- Gagnon, J., & Jolicoeur, P. (2014). Mécanismes cognitifs et neurophysiologiques associés au trait "Urgence" de la personnalité: intégration théorique et empirique. *Insight Development: Social Sciences and Humanities Research Council of Canada*.

- Gagnon, J., McDuff, P., Daelman, S., & Fournier, S. (2015). Is hostile attributional bias associated with negative urgency and impulsive behaviors? A social-cognitive conceptualization of impulsivity. *Personality and Individual Differences, 72*, 18-23.
- Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Bessette, M., and Jolicœur, P. (2016). Neural mechanisms underlying attribution of hostile intention in nonaggressive individuals: an ERP study. *International Journal of Psychophysiology 110*, 153-162.
- Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Brochu, A. F., Bessette, M., & Jolicœur, P. (2017). An ERP study on hostile attribution bias in aggressive and nonaggressive individuals. *Aggressive Behavior, 43*(3), 217-229.
- Gagnon, J., & Rochat, L. (2017). Relationships between hostile attribution bias, negative urgency, and reactive aggression. *Journal of Individual Differences, 38*(4), 211-219.
- Gotlib, I. H., & Krasnoperova, E. (1998). Biased information processing as a vulnerability factor for depression. *Behavior Therapy, 29*(4), 603-617.
- Judd, L.L., Schettler, P.J., Coryell, W., Akiskal, H.S., Fiedorowicz, J.G. (2013). Overt Irritability/Anger in Unipolar Major Depressive Episodes: Past and Current Characteristics and Implications for Long-term Course. *JAMA Psychiatry. 70*(11), 1171-1180.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.
- 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.

- Milich, R., & Dodge, K. A. (1984). Social information processing in child psychiatric populations. *Journal of Abnormal Child Psychology*, *12*(3), 471-489.
- Nasby, W., Hayden, B., & DePaulo, B. M. (1980). Attributional bias among aggressive boys to interpret unambiguous social stimuli as displays of hostility. *Journal of Abnormal Psychology*, *89*(3), 459-468.
- Pedrelli, P., Nyer, M., Holt, D., Bakow, B.R., Fava, M., Baer, L., ... Farabaugh, A. (2013). Correlates of irritability in college students with depressive symptoms. *The Journal of Nervous and Mental Disease*, *201*, 953-958
- Perlis, R. H., Fava, M., Trivedi, M. H., Alpert, J., Luther, J. F., Wisniewski, S. R., & Rush, A. J. (2009). Irritability is associated with anxiety and greater severity, but not bipolar spectrum features, in major depressive disorder. *Acta Psychiatrica Scandinavica*, *119*(4), 282-289.
- Raine, A., Dodge, K. A., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., Stouthamer-Loeber, M., & Liu, J. (2006). The reactiveproactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior*, *32*, 159-171.
- Safer, D. J. (2009). Irritable mood and the diagnostic and statistical manual of mental disorders. *Child and Adolescent Psychiatry and Mental Health*, *3*, Article ID 35.
- Sharbrough, F., Chatrian, G.-E., Lesser, R. P., Luders, H., Nuwer, M., & Picton, T. W. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *8*, 200-202.

Smith, H. L., Summers, B. J., Dillon, K. H., Macatee, R. J., & Coughen, J. R. (2016). Hostile interpretation bias in depression. *Journal of Affective Disorders, 203*, 9-13. doi:

10.1016/j.jad.2016.05.070

Statistics Canada. (2012). Table: 105-1101 - *Canadian Community Health Survey Share File - Mental Health (CCHS)*, CANSIM (database).

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20*(4), 580–591.