

Université de Montréal

**Negativity bias and instability in spontaneous and deliberate evaluations of others: The
role of borderline personality disorder features**

Par
Félix Gauthier Mongeon

Département de psychologie
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures en vue de l'obtention du grade de
Philosophiae Doctor (Ph.D.)
en psychologie recherche et intervention
option clinique
25 novembre 2016

© Felix Gauthier Mongeon, 2017

Résumé

Les dysfonctions interpersonnelles sont au cœur du trouble de la personnalité limite (TPL). La recherche passée a examiné différents mécanismes sous-jacents aux dysfonctions interpersonnelles dans la personnalité limite, incluant la dérégulation émotionnelle et l'impulsivité, les déficits dans la reconnaissance des expressions faciales émotionnelles, la mentalisation faible, la confiance et la coopération faibles, les difficultés concernant l'empathie et l'intimité ainsi que l'instabilité affective et interpersonnelle. Ainsi, la présente thèse vise à développer les connaissances sur les mécanismes socio-cognitifs sous-jacents aux dysfonctions interpersonnelles dans la personnalité limite. Dans ce but, cette thèse examine le rôle des évaluations d'autrui dans la personnalité limite. En effet, chacun forme des évaluations plus ou moins positives ou négatives à partir des premières interactions avec une nouvelle personne et ces impressions, tout en évoluant, demeurent fondamentales dans les interactions sociales, incluant dans les relations amoureuses. Ces évaluations peuvent se manifester de deux manières : en effet, il y a maintenant amplement de preuves empiriques suggérant que les évaluations peuvent résulter de processus relativement réfléchis et délibérés (c.-à.-d., *évaluations délibérées*) ou qu'elles peuvent émerger à partir de processus plus instantanés où il y a aucune ou peu d'intention consciente d'évaluer le stimulus (c.-à.-d., *évaluations spontanées*). Dans cette thèse, deux études mesurent les traits de la personnalité limite (PL) sur un continuum de sévérité et examinent si ces traits impliquent des évaluations d'autrui plus négatives ou clivées (c.-à.-d., instables et polarisées) aux niveaux spontané et délibéré.

L'étude 1 ($N = 204$) examinait les évaluations d'un personnage de film à deux temps de mesure, c.-à.-d., après un extrait de film positif et après un extrait de film négatif. Les évaluations spontanées étaient mesurées à l'aide d'une tâche d'amorçage évaluatif et les évaluations délibérées étaient mesurées à l'aide d'une mesure auto-rapportée. L'étude 1 examinait deux hypothèses principales : si les traits du PL étaient liés à (1) plus d'instabilité et (2) plus de négativité dans les évaluations du personnage de film aux niveaux délibéré et spontané. Les résultats montraient que les participants évaluaient généralement positivement le personnage sur les deux types d'évaluations. Contrairement à l'hypothèse de l'instabilité, les traits du PL n'impliquaient pas d'alternance entre des évaluations positives et négatives ou entre des évaluations négatives et positives. Cela était vrai pour les évaluations spontanée et

délibérée. Toutefois, les deux types d'évaluations impliquaient des biais négatifs contextuels. Spécifiquement, les traits PL étaient reliés à des évaluations relativement plus négatives du personnage de film. Ce patron de résultat est discuté à la lumière des considérations additionnelles suivantes : (1) il émergeait dans différentes conditions de l'étude selon le type d'évaluations, (2) il semblait avant tout déterminé par davantage de positivité chez les individus avec des traits PL faibles et (3) il était basé sur une interaction marginale à trois voies pour la mesure implicite. Pris dans leur ensemble, les résultats obtenus avec les mesures explicite et implicite dans l'étude 1 convergent partiellement avec les appuis empiriques et les théories cliniques sur les biais négatifs dans le TPL, mais il importe de nuancer les résultats sur la mesure implicite considérant qu'ils sont basés sur une interaction marginale.

L'étude 2 ($N = 292$) examinait les évaluations du partenaire après que les participants aient imaginé soit (a) que leur partenaire les rejetait ou (b) qu'il cherchait une connexion (c) ou sans qu'ils aient imaginé aucun scénario (condition contrôle). Les évaluations spontanées du partenaire étaient mesurées avec une tâche de lettre-nom (*name-letter task*) et les évaluations délibérées du partenaire étaient mesurées à l'aide d'une mesure auto-rapportée. L'étude 2 examinait deux hypothèses principales : si les traits PL étaient liés (1) à des évaluations plus polarisées du partenaire (c.-à.-d., alternance entre les valences positive et négative) ou (2) à des évaluations plus négatives du partenaire. Ces deux hypothèses étaient examinées au niveau des évaluations spontanées et délibérées. D'abord, conformément à la recherche passée, les résultats montraient que les participants avaient des évaluations à prédominance positives à l'endroit de leur partenaire et d'eux-mêmes. Deuxièmement, contrairement aux hypothèses, les manipulations de rejet et de connexion n'influençaient pas les évaluations spontanées du partenaire et leur influence sur les évaluations délibérées du partenaire était marginale et petite. Troisièmement, les résultats ne supportaient pas l'hypothèse de la polarité : les individus avec des traits PL élevés ne manifestaient pas davantage de positivité après la condition de connexion et ils ne manifestaient pas davantage de négativité après la condition de rejet. Plutôt, les individus avec des traits PL élevés avaient une préférence réduite pour les initiales de leur partenaire et rapportaient des réactions moins positives à l'endroit de leur partenaire indépendamment des conditions de l'étude, conformément à un biais négatif généralisé sur les deux types d'évaluations. La discussion nuance les implications des résultats pour prendre en compte le fait que les conditions expérimentales n'ont pas produit les effets

principaux attendus sur les évaluations du partenaire. Prises dans leur ensemble, les deux études ont trouvé des évaluations généralement positives et les traits PL modulaient ces effets : les deux études comportent des preuves préliminaires suggérant que les individus avec des traits PL pourraient évaluer relativement plus négativement les autres et que ces biais négatifs pourraient se présenter au niveau de leurs évaluations spontanées et délibérées. Ces résultats suggèrent que de potentiels biais négatifs à l'endroit d'autrui pourraient être issus d'une combinaison de processus à la fois relativement précoces et tardifs dans les étapes de traitement de l'information. Toutefois, cette interprétation est nuancée et considérée avec prudence étant donné (a) qu'elle est basée en partie sur des effets marginaux, (b) que certains effets n'ont pas été répliqués à travers les conditions de chaque étude et d'une étude à l'autre et (c) que certains des effets principaux attendus des conditions expérimentales n'ont pas été observés dans l'étude 2.

Mots clefs: personnalité limite; évaluations spontanées; évaluations délibérées; biais négatif; instabilité; cognition sociale; mesure implicite; mesure explicite

Abstract

Interpersonal dysfunctions are central to borderline personality disorder (BPD). Past research has examined different mechanisms underlying interpersonal dysfunctions in borderline personality, including emotion dysregulation and impulsivity, deficits in facial emotion recognition, low mentalization, low trust and cooperation, impairments in empathy and intimacy as well as affective and interpersonal instability. The present thesis also aims at obtaining deeper insights into the interplay of socio-cognitive mechanisms underlying interpersonal dysfunctions in borderline personality. Toward this goal, this thesis examined the role of evaluations of others in borderline personality. In fact, evaluations of others exist from the first moment when we meet someone and it remains one of the most fundamental impressions we have in our social interactions, including in romantic interactions. Such evaluations can present at two levels of evaluations; in fact, there is ample evidence that evaluations can be the result of relatively thoughtful, deliberate processes (i.e. *deliberate evaluations*) or of relatively instantaneous processes where there is no or little conscious intention to evaluate the stimuli (i.e. *spontaneous evaluations*). In this thesis, two studies assessed borderline personality (BP) features on a continuum of severity and examined whether these features are related with negative or split-off (i.e., unstable and polarized) evaluations of others at the spontaneous and deliberate levels of evaluations.

Study 1 ($N = 204$) examined evaluations of a movie character at two different times, that is, after a positive clip and after a negative clip. Spontaneous evaluations were measured with an evaluative priming task and deliberate evaluations were measured with a self-report measure. Study 1 examined two main hypotheses: whether BP features are related with (1) more unstable or (2) more negative evaluations of the movie character at the spontaneous and deliberate levels of evaluations. Results showed that participants had overall positive evaluations of the movie character on both types of evaluations. Contrary to the instability hypotheses, BP features involved no switch from positive to negative evaluations or from negative to positive evaluations on either type of evaluations. However, we found evidence for context-specific negativity biases for both spontaneous and deliberate evaluations. Specifically, BP features were related with relatively more negative evaluations of the movie character. This effect was interpreted in light of the following additional considerations: (1) it emerged in different conditions of the study depending on the type of evaluations, (2) it

appeared to be driven by greater positivity in individuals with low BP features and (3) it relied on a marginal three-way interaction regarding the implicit measure. Together, the results obtained on the explicit and implicit measures in Study 1 partially converge with empirical evidence and clinical theorizing about negativity biases in BPD and the result on the implicit measure should be considered with prudence given that it relies on a marginal three-way interaction.

Study 2 ($N = 292$) examined partner evaluations after participants imagined either (a) that their partner rejected them or (b) sought connection or (c) imagined no scenario (control condition). Spontaneous partner evaluations were measured with a name-letter task and deliberate partner evaluations were measured with a self-report measure. Study 2 examined two main hypotheses: whether BP features are related with (1) more polarized partner evaluations (i.e., alternation between negativity and positivity) or (2) more negative partner evaluations at the spontaneous and deliberate levels of evaluations. First, replicating prior research, results showed that individuals had predominantly positive evaluations of themselves and of their partner. Second, contrary to expectations, the rejection and closeness manipulations did not influence spontaneous partner evaluations and their influence on deliberate partner evaluations was marginal and small. Third, the results did not support the polarity hypothesis: individuals with high BP features did not display greater positivity after the closeness condition and greater negativity after the rejection condition. Instead, individuals with high BP features showed a reduced preference for their partner's initials and reported less positive reactions to their partner independent of study conditions, in line with a generalized negativity bias on both types of evaluations. The discussion nuances the implications of the results to take in consideration the fact that the experimental conditions did not produce the expected main effects on partner evaluations.

Taken together, both studies found overall more positive evaluations, but BP features modulated this effect: we found preliminary evidence that individuals with high BP features had more negative evaluations (i.e., negativity biases) on both kinds of evaluations and in both studies. These results suggest that potential negativity biases toward others may be shaped both at relatively early and late processing stages in borderline personality. However, this interpretation is considered with prudence given that (a) it is partly based on marginal effects, (b) that some effects did not replicate across conditions of each study or across the two studies

and (c) that the expected main effects of experimental conditions were not observed in Study 2.

Keywords: borderline personality; spontaneous evaluation; deliberate evaluation; negativity bias; instability; social cognition; implicit measure; explicit measure

Liste des tableaux

Introduction

Table 1 Concepts hypothesized to potentially characterize evaluations of others in BPD and their operationalization. **11**

Article 1

Table 1 List of positive and negative words used in the explicit and implicit measures **39**

Table 2 Summary of GLMs on spontaneous and deliberate evaluations. Error!
Bookmark not defined.

Table 3 Summary of ANCOVAs on spontaneous and deliberate evaluations. Error!
Bookmark not defined.

Table 4 Contrasts between positive and negative indices of deliberate evaluations **46**

Article 2

Table 1 Socio-demographic statistics. **67**

Table 2 List of positive and negative words used in the explicit measure. **68**

Table 3 Summary of regressions on spontaneous and deliberate partner evaluations. **71**

Table des matières

Résumé	i
Abstract	iv
Liste des tableaux	vii
Introduction	1
Article 1	29
Negativity bias and instability in spontaneous and deliberate evaluations of others: The role of borderline personality disorder features.....	29
Article 2	58
Negativity bias and polarization in spontaneous and deliberate partner evaluations: The role of borderline personality disorder features.....	58
Conclusion	85
Références citées dans l'introduction et dans la conclusion	106
Annexe 1 : Consent form for Study 1	120
Annexe 2 : Ethics approval for Study 1	123
Annexe 3 : Material for Study 1	124
Annexe 4 : Consent form for Study 2	132
Annexe 5 : Ethics approval for Study 2	133
Annexe 6 : Material for Study 2	140

Introduction

Borderline personality disorder (BPD) is a disorder characterized by significant interpersonal dysfunctions. Suicidal and self-damaging behaviors (Soloff, Lis, Kelly, Cornelius, & et al., 1994) and a drastic increase in prevalence in clinical settings exemplify the treatment complexities associated with this disorder: despite a prevalence of 1-2% in the general population (Swartz, Blazer, George, & Winfield, 1990), BPD has been shown to occur in 20% of intern patients in psychiatry (Zanarini, Frankenburg, Khera, & Bleichmar, 2001). In addition, a hallmark feature of this disorder is the difficulty maintaining positive or stable evaluations of other people. The notions of dichotomous thinking and splitting both capture a similar tendency of individuals with BPD to evaluate other people through mutually exclusive or polarized categories (black-or-white, all-good vs. all-bad); moreover, the notion of splitting posits that BPD involves a tendency to form unstable evaluations of others characterized by switches between devaluation and idealization (A. T. Beck & Freeman, 1990; A. T. Beck & Freeman, & Davis, 2004; Kernberg, 1986, D. Westen, 1991). Clinical observations of individuals with BPD also suggest that polarized and unstable evaluations of others may possibly lead them to display unstable behaviors within their relationships, such as switches between distance or antagonism (during devaluation) and pursuance or overdependence (during idealization). Such unstable behaviors may potentially elicit feelings of confusion and insecurities in others, in turn fuelling negative behavioral reactions and contributing to a vicious cycle of negative interpersonal dynamics. To gain a better understanding of the mechanisms underlying interpersonal difficulties in borderline personality, the present research examines the relationship between borderline personality and the way one evaluates other people.

“Whether to like or dislike someone is one of the most fundamental and highly consequential interpersonal evaluations that people make about other people” (p. 671 Krause, Back, Egloff, & Schmukle, 2014). Gawronski and Bodenhausen (2011) point that

such evaluations can be the result of thoughtful, deliberate processes that aim at integrating various stimulus characteristics in a coherent evaluative judgment. At the same time, environmental stimuli may elicit evaluative responses instantaneously and in the absence of a conscious intention to evaluate these stimuli (p. 60).

The former kind of evaluations is described as *deliberate*, whereas the later kind is described as *spontaneous*. Spontaneous and deliberate evaluations sometimes conflict. For instance, one may experience positive spontaneous feelings toward a potential partner, while believing that this is not the right person to be in a relationship with. Despite ample evidence supporting the distinction between spontaneous and deliberate evaluations, this distinction has not been taken into account in the BPD literature. Based on these considerations, the primary goal of this thesis is to gain deeper insights into the way individuals with borderline personality (BP) features evaluate other people, focusing on the distinction between *spontaneous* and *deliberate* evaluations of others.

Toward this goal, we utilized self-report tools to measure BP features on a continuum of severity rather than as a discrete diagnostic category. In fact, based on previous research, we recognize that BPD is not a discrete trait, but rather the extreme manifestation of a normally distributed personality disposition (Edens, Marcus, & Ruiz, 2008; Haslam, 2003; Rothschild, Cleland, Haslam, & Zimmerman, 2003). Also, we combined explicit measures (i.e., self-report tools) and implicit measures to capture deliberate and spontaneous evaluations, respectively. Implicit measures are typically performance-based tasks that allow indirectly measuring and inferring spontaneous feelings towards a target stimulus. The implicit association task (IAT; A. G. Greenwald, McGhee, & Schwartz, 1998) and sequential priming task (Fazio, Jackson, Dunton, & Williams, 1995) are among the most commonly used implicit measures; these two tasks measure the strength of an association between a target concept (e.g., snake) and an attribute dimension (e.g., positive) based on the latency with which participants categorize the stimuli. For instance, these tasks can be used to assess spontaneous evaluations of another person based on the strength of the association between faces of that person and positive/negative target words (such as in the first study presented in the present thesis). Other kinds of performance-based measures can be used to measure spontaneous evaluations (for overviews see Gawronski, Deutsch, & Banse, 2011; Teachman, Joormann, Steinman, & Gotlib, 2012). On a fundamental level, the combination of explicit and implicit measures in the present research may help clarifying whether potential evaluation biases of individuals with high BP features are shaped at relatively early (i.e., spontaneous) or late (i.e., deliberate) processing stages. Specifically, if implicit and explicit measures of evaluations converge, it strengthens the obtained findings and suggests that evaluations are shaped at relatively early

(i.e., associative) and late (i.e., propositional) processing stages in borderline personality. In contrast, if implicit and explicit measures diverge, this could fuel new predictions and reveal aspects of evaluations of others that were previously unforeseen in borderline personality (for the example of implicit racism see A. G. Greenwald et al., 1998). Therefore, by examining the interplay between spontaneous and deliberate evaluations, the present research can bring new insights into the dynamic of interpersonal dysfunctions in borderline personality. In addition, in order to better reflect the range of interpersonal situations occurring in real-life, we conducted two studies that focused on two types of social interactions, namely impression formation (Study 1) and partner evaluations (Study 2). Thus, the results obtained in the present research will also help to explore potential differences and similarities between these two interpersonal contexts in relation to borderline personality. We hope that future studies can build on this work to better understand the mechanisms underlying interpersonal dysfunctions in individuals with BP features or BPD, particularly whether difficulties building trusting, secure and stable relationships may stem from evaluation biases that are shaped at relatively early and/or late processing stages.

Spontaneous and deliberate evaluations.

Associative-propositional evaluation (APE) model. Different models have been proposed to explain results obtained with explicit and implicit measures (e.g., De Houwer, in press; Fazio & Olson, 2003; Fazio & Towles-Schwen, 1999; Gawronski & Bodenhausen, 2006; Moors & De Houwer, 2006; Strack & Deutsch, 2004). The associative-propositional evaluation (APE) model is an empirically-based theory that has been developed to delineate the nature of and interplay between spontaneous and deliberate evaluations and to provide explanations of dissociations between implicit and explicit measures in terms of the respective processes guiding the two kinds of evaluations (Gawronski & Bodenhausen, 2006, 2011). According to the APE model, spontaneous and deliberate evaluations result primarily from associative and propositional processes, respectively, but the two kinds of processes and evaluations can also indirectly influence each other. Spontaneous evaluations represent affective gut reactions resulting from the activation of mental associations in memory; in contrast, deliberate evaluations represent the behavioral outcome of propositional processes involving the validation of activated information on the basis of logical consistency. An important implication of the model is that, whereas deliberate evaluations depend on

subjective judgments of truth or falsity, spontaneous evaluations are not directly influenced by subjective truth or falsity. This difference explains why implicit and explicit measures sometimes diverge (for a review on explicit-implicit relation see Nosek, 2007). Also, associative processes depend on pre-existing memory associations. Therefore, while spontaneous evaluations are directly influenced by memory activation, deliberate evaluations are influenced by an additional layer of filtering accomplished by propositional processes and allowing one to consider additional beliefs and information. To exemplify the APE model, take a person with BPD and an historic of traumatic interpersonal experiences. This person will encode these experiences in memory and may also develop firm beliefs of distrust and malevolence, such as “other people should not be trusted” or “other people will try to harm me”. To the extent that these memories and beliefs are very accessible, they are likely to be activated when encountering new people, leading the person to often experience negative spontaneous reactions to new people. Such negative spontaneous evaluations can arise despite no or little actual knowledge of the newly encountered person or stimulus. As for the deliberate evaluations, they can converge or diverge with spontaneous evaluations. In case of convergence, the person may use its negative spontaneous reactions to the newly encountered person as a valid basis to form deliberate judgments on this person. This would lead to corresponding negative deliberate evaluations. In case of divergence, the person may deploy additional propositional processes, such as thinking that there is not enough information to form negative judgments of that person. This would lead the person to disregard its spontaneous negative reactions, resulting in a divergence between spontaneous and deliberate evaluations, assuming that spontaneous evaluations would remain unchanged (i.e., negative spontaneous evaluations, with positive deliberate evaluations). In fact, propositional and associative processes can change without influencing each other, but they can also mutually influence each other. For instance, implicit-explicit correspondence has been shown to increase when participants are prompted to base their deliberate evaluations on their spontaneous feelings (Gawronski & LeBel, 2008) or when participants who have negative spontaneous evaluations are prompted to think of positive counter-attitudinal examples in order to positively influence their spontaneous feelings (Gawronski, Deutsch, Mbirikou, Selbt, & Strack, 2008).

Imagine encountering a new person in a friendly party. Upon this new encounter, which memories are activated depend on the saliency of all potentially accessible memories and on the context, that is, the overall configuration of input stimuli (Gawronski & Bodenhausen, 2011). In a party context, a new encounter may activate memories of positive interpersonal experiences and lead to positive evaluations of that person (e.g., warm, friendly), whereas the same person may activate negative evaluations if encountered in a context that activates negative interpersonal memories (for examples of studies on the effect of context see Barden, Maddux, Petty, & Brewer, 2004; Wittenbrink, Judd, & Park, 2001). Therefore, the same target stimulus can activate different patterns of associations and the resulting evaluation ultimately depends on a combination of (1) contextual cues and (2) pre-existing associative representations of that object (for a discussion and a review of the evidence supporting the model, see Gawronski & Bodenhausen, 2006; Gawronski & Bodenhausen, 2011). Finally, the APE model discuss how the two types of evaluations operate under different conditions: compared to deliberate evaluations, spontaneous evaluations have been proposed to operate under conditions of unawareness (processing of emotional stimuli occurs outside awareness), efficiency (processing emotional meaning uses minimal attentional resources), unintentionality (no goal is needed to engage in processing emotional meaning), and uncontrollability (limited ability to avoid, alter or terminate processing emotional stimuli) (Bargh, 1994; for a critical discussion of these features, see Gawronski & Bodenhausen, 2011).

Other models have been proposed to explain dissociations and relations between implicit and explicit measures as well as their differences in terms of behavior prediction, such as the motivation and opportunity for deliberation (MODE) model (Fazio & Towles-Schwen, 1999; Olson & Fazio, 2008) (for evidence supporting the model see Barrett, Tugade, & Engle, 2004; Baumeister, Bratslavsky, Muraven, & Tice, 1998). Akin to the APE model, the MODE model also distinguishes between two basic processes. However, while the APE model focuses primarily on the nature of and interplay between associative and propositional processes, the MODE model focuses on two factors – motivation and opportunity – influencing whether evaluations guide behavior in a spontaneous or deliberate manner. With respect to motivation, Olson and Fazio (2008, p. 23) argue that, given the effortful reflection required by deliberation,

some motivating force is necessary to induce individuals to engage in the reasoning. The MODE model posits that a variety of motivational factors might push an individual toward a more deliberative scrutiny of behavioral options. Perhaps the most fundamental of these motives is the desire to be accurate: that is, to reach valid conclusions.

As for the notion of opportunity, Olson and Fazio (2008, p. 24) point that

in order for any motivation to overcome the influence of one's attitude, an opportunity for this motive to exert its influence must also be available. And again, the model views «opportunity" broadly; it manifests in a variety of ways. Opportunity can simply amount to a matter of time; careful consideration of information simply cannot be done quickly. But opportunity also comes in psychological forms. Because our cognitive resources are limited, fatigue, distraction, and other factors can interfere with one's ability to process information. (for evidence supporting the model see Barrett et al., 2004; Baumeister et al., 1998).

In sum, the APE model and the MODE model are empirically based frameworks that describe the distinctions between spontaneous and deliberate evaluations. Based on these distinctions, the present study utilized an implicit and an explicit measure to examine the interplay between spontaneous and deliberate evaluations of others in borderline personality.

Theoretical and clinical accounts of evaluations of others in BPD

Evaluations of others pertain to how positively or negatively one evaluates social stimuli and can be investigated using standardized stimuli in the laboratory (e.g., faces or movie characters) or real-life interactions (e.g., friends, family or partner). Both cognitive and psychodynamic theories have discussed evaluations of others in BPD. The cognitive theory posits that *dichotomous thinking* constitutes a central feature of BPD. Dichotomous thinking refers to the tendency of individuals with BPD to interpret their experiences through mutually exclusive categories, including their interpersonal experiences, leading to extreme evaluations of others (A. T. Beck & Freeman, 1990; Aaron T. Beck, Freeman, & Davis, 2004). Moreover, the cognitive theory of BPD hypothesizes that dysfunctions of BPD individuals are at least partly caused by maladaptive cognitive schemas (basic cognitive structures in memory) that bias the evaluation and interpretation of environmental stimuli, including the basic assumption

that the world is dangerous and malevolent and that the self is powerless, vulnerable and unacceptable (A. T. Beck & Freeman, 1990; Aaron T. Beck et al., 2004).

In addition, according to Kernberg (1986), *splitting* constitutes a central mechanism that characterizes and fuels borderline pathology. Individuals with BPD are often unable to integrate complex experiences into nuanced representations of self and others. Splitted representations are thus separated into all-negative or all-positive evaluations and BPD individuals switch between these two polarities (i.e., all-good or all-bad). Accordingly, the psychodynamic concept of *splitting* (Kernberg, 1986) is similar to the notion of dichotomous thinking to the extent that both concepts capture the tendency of individuals with BPD to display extreme evaluations of the self and others. Beyond these similarities, Arntz and colleagues have also pointed to distinctions between dichotomous thinking and splitting (e.g., Veen & Arntz, 2000). They argued that dichotomous thinking can be multidimensional, which means that it may involve extreme evaluations with a mixture of the positive and negative valence (i.e., evaluating someone as being very honest and very rude at the same time); in contrast, they argued that splitting is best defined as unidimensional, that is, splitted evaluations are either all-good or all-bad, with no or little integration of positive and negative valence at a given point in time. Notwithstanding their contribution, Veen and Arntz (2000) did not highlight the fact that splitting contains two distinct predictions (i.e., instability and polarity) and that these should be investigated as such in empirical research. Little research has specifically examined unstable or polarized evaluations of others, but there are a few exceptions. For instance, Coifman, Berenson, Rafaeli, and Downey (2012) showed that BPD involves either all-positive or all-negative evaluations and this finding supports the notion that BPD involves polarity; also, Zeigler-Hill and Abraham (2006) showed that BPD involves switches from negative to positive evaluations and this finding supports the notion that BPD involves instability. In this vein, DSM-5 notes that BPD involves unstable evaluations of others (Criterion 2: “close relationships often viewed in extremes of idealization and devaluation”) and self (Criterion 3: “unstable self-image”; American Psychiatric Association, 2013).

The BPD literature has also discussed another type of evaluation bias, namely the *negativity bias* (or negativity, negative thinking style, negativistic evaluation style, negative evaluation bias). For instance, Arntz and Haaf (2012) describe this phenomenon as a tendency

in individuals with BPD to emphasize and amplify negative qualities and to ignore and to diminish positive qualities. In this vein, the DSM-5 notes that BPD involves negative evaluation of others (“perceptions of others selectively biased toward negative attributes or vulnerabilities”) and self (“feelings of inferior self-worth”) (American Psychiatric Association, 2013, pp. 766-767). Previous research suggests that negative evaluations of others may contribute to emotional dysregulation and problematic interpersonal behaviors among individuals with BPD (Berenson, Downey, Rafaeli, Coifman, & Paquin, 2011; Coifman et al., 2012; Sadikaj, Moskowitz, Russell, Zuroff, & Paris, 2013; Sadikaj, Russell, Moskowitz, & Paris, 2010). Akin to splitting, the notion of negativity bias points to (1) more extreme evaluations, but in contrast to splitting, (2) these extreme evaluations are mostly driven by greater negativity, rather than by a mixture of negativity and positivity (dichotomous thinking) or by polarization or instability (splitting). Table 1 shows the definitions of dichotomous thinking, splitting and negativity bias, as well as their meanings and operationalizations. Whereas previous empirical research has examined the phenomena of negativity biases, polarized and dichotomous thinking in individuals with BPD, no research to the best of our knowledge has specifically examined the hypothesis that BPD is characterized by unstable evaluations of others.

Research on evaluations of others in BPD.

Behavioral and neurological evidence of negativity biases in BPD. Diverse areas of BPD research have found compelling evidence for negativity biases in BPD. For instance, Domes, Schulze, and Herpertz (2009) reviewed behavioral studies of negativity biases in facial emotion recognition in BPD and found evidence of a heightened sensitivity to the detection of negative emotions, along with a negativity or anger bias. In addition, neurobiological findings suggest that the negativity bias in BPD might stem from a hyperreactivity of limbic brain areas and a hyporeactivity of prefrontal areas in response to negative socio-emotional stimuli as well to neutral social stimuli (for a review, see Krause-Utz, Winter, Niedtfeld, & Schmahl, 2014). There is also physiological evidence for a negativity bias in the patterns of physiological reactivity of BPD individuals. For example, Matzke, Herpertz, Berger, Fleischer, and Domes (2014) found that BPD individuals displayed reduced facial responding to positive social signals and increased facial responding to negative social signals. Moreover, several studies have shown that rejection is a common theme to the

negative reactions of individuals with BPD, such that they are more prone to perceive rejection when actually rejected, but also when not rejected (e.g., Renneberg et al., 2012), and to endorse beliefs and expectations that they will be rejected or abandoned (Dreessen & Arntz, 1995a; Staebler, Helbing, Rosenbach, & Renneberg, 2011). There is evidence that the rejection sensitivity in borderline personality can contribute to distrust appraisal of neutral faces (Miano, Fertuck, Arntz, & Stanley, 2013), to stronger reactions of rage (Berenson et al., 2011), to greater negative affect and reduced positive affect (Sadikaj et al., 2010), and to more quarrelsome behavior and negative affect in interpersonal situations (Sadikaj et al., 2013). Hence, there is self-report, behavioral, neurobiological and physiological evidence that BPD individuals interpret social information through a negative lens, as shown by enhanced expectations and perception of rejection, negativity biases in basic socio-cognitive processes (e.g., facial emotion recognition) and in patterns of physiological reactivity, and a failure of frontal areas in the inhibition of limbic hyperreactivity.

Finally, other studies brought further indirect evidence for negativity biases by showing that personality characteristics associated with borderline personality – negative affect flux (instability) (Sadikaj, Moskowitz, & Zuroff, 2015), low self-esteem (Murray, Bellavia, Rose, & Griffin, 2003) and attachment anxiety (Simpson, Ickes, & Grich, 1999) – are related with higher tracking accuracy of partner's negative behavior, which suggests greater vigilance toward a partner's negative behavior and is thus consistent with a negative evaluation bias within romantic relationships. Similarly, participants whose partners or themselves used less effective emotion regulation strategies – which is also a characteristic of borderline personality – perceived more hostile criticisms from their partner (Klein, Renshaw, & Curby, 2016).

Research stemming from psychodynamic theories. Research stemming from clinical psychodynamic paradigms has brought interesting insights into the way BPD may influence evaluations of self and others. Because psychodynamic theories highlight the key role of unconscious representations (e.g., Kernberg, 1986), early studies on splitting used projective methods to investigate unconscious psychological phenomena that are conceptually similar to splitting. Although projective instruments are not traditionally considered as implicit measures and were not originally developed within a framework of spontaneous and deliberate evaluations, they share similarities with implicit measures to the extent that both types of measurement tools assess evaluations of others indirectly, that is, without asking participants

to report on the to-be-measured phenomenon. Moreover, there is evidence that projective tests capture more spontaneous evaluations based on patterns of convergent validity of projective and self-report measures (Bornstein, 2002).

The Social Cognitions and Object Relations Scale (SCORS) is a method that has been used with the goal to reveal potentially unconscious mechanisms underlying personality disorders and has been used for interpreting results on different measurement tools, including the Thematic Aperception Test (TAT) (Westen, 1991). In one study, individuals with BPD showed simplistic representations of others and self, malevolent representations of others, and deficits in the integration of different aspects of the self relative to control participants with other personality disorders (antisocial, narcissistic and cluster C personality disorders) (Ackerman, Clemence, Weatherill, & Hilsenroth, 1999), which supports the splitting and negativity bias hypotheses. Also, individuals with BPD displayed more malevolent representations of others than participants with major depressive disorder or normal participants, in line with a negativity bias (Westen, Lohr, Silk, Gold, & Kerber, 1990), but this study did not include the splitting-relevant scale on identity integration. Using the Rorschach instead of the TAT, Stuart, Westen, Lohr, Benjamin, and et al. (1990) also found evidence for negativity bias in individuals with BPD relative to participants with major depression and normal control participants.

Another way to capture phenomena that participants may not want to or be able to report is to rely on observational data, such as when interpreting the content of interviews through qualitative analysis. For instance, Sheffield et al. (1999) examined idiothetic relational narratives of participants with BPD using semistructured interviews and found evidence of negativity biases; specifically, their qualitative analyses revealed that participants with BPD experienced difficulty in integrating paradoxical experiences, such as the ideas (1) that acceptance is essential but that one cannot be accepted because of its flaws (i.e. negative evaluations of self) and (2) that nurturance is essential but that it should not be accepted when offered because the other will show malevolent intentions (i.e. negative evaluations of others). Another psychodynamically oriented qualitative study investigated evaluations of self and others based on interviews with patients with BPD and patients with major depression (Dammann et al., 2011). Patients with BPD evidenced overall more negative evaluations of self and others than patients with major depression. In sum, psychodynamically oriented

studies that examined evaluations of self and others in BPD utilized projective and interview methods and found evidence that converge with the view that BPD is characterized by negative ways of interpreting relationships' situations. Moreover, results showing simplistic and non-integrated representations on projective tests bring indirect evidence of splitting. However, the projective and interview methods used in these studies were not specifically designed to examine the phenomena under scrutiny here. In fact, these measures rely on participants' reports and thus these are not implicit measures; in fact, participants' verbal responses are not good indicators of spontaneous evaluations (Gawronski & Bodenhausen, 2011). In addition, these studies did not assess evaluations of other people (e.g., rating someone as more or less aggressive or gentle), but rather evaluations of interpersonal situations more generally (e.g., interpreting an interpersonal scenario on the TAT). Finally, they did not examine instability, that is, alternation between negativity (devaluation) and positivity (idealization).

Table 1

Concepts hypothesized to potentially characterize evaluations of others in BPD and their operationalization.

Concept	Meaning	Operationalization
Dichotomous thinking	Extreme evaluation style ("black or white").	The extremity of the evaluations (no matter the valence)
Splitting	Viewing others and self as either completely good or completely bad ("all good or all bad").	1. Polarization: The degree to which all evaluations are polarized to positive or negative qualities. 2. Instability: The degree to which evaluations of one target switch from the positive to the negative across different measurement times.
Negativity bias	Negativistic evaluation style: emphasizing and amplifying negative qualities; ignoring and diminishing positive qualities.	The degree to which evaluations ratings are negative.

Notes: The definitions of concepts included in this table are adapted from Arntz and Haaf (2012)

Cross-sectional studies on evaluations of others in BPD. Other studies have examined the interplay between borderline personality and evaluations of others in the laboratory using cross-sectional or experimental designs. For instance, Oshio (2009) has designed a questionnaire, the *Dichotomous Thinking Inventory*, which divides dichotomous thinking in three dimensions: dichotomous beliefs (e.g. "There are only “winners” and “losers” in this world."), dichotomous preferences (e.g. "All things work out better when likes and dislikes are clear.") as well as profit-and-loss thinking (e.g. "I want to clearly distinguish what is safe and what is dangerous."). While the general dichotomous thinking score did not seem to discriminate between borderline and narcissistic features, the subscale of dichotomous beliefs was more strongly related with borderline features than narcissistic features (Oshio, 2009). Interestingly, the content of this subscale emphasizes splitted representations *of others* (e.g. "There are only “winners” and “losers” in this world.", "People can clearly be distinguished as being “good” or “bad”"). Oshio (2012) conducted another study using the same questionnaire (i.e., the *Dichotomous Thinking Inventory*) and found a relationship between dichotomous beliefs and borderline personality features, replicating its previous finding (Oshio, 2009), with the difference that dichotomous beliefs were more strongly related to features of cluster B personality disorders in general than to borderline personality features specifically.

Another study used a computerized self-report measure that calculated different indices of deliberate self evaluations that reflected aspects of dichotomous thinking and negativity bias in BPD (Evans et al., 2015). Participants were presented with 60 different personality traits (30 positive and 30 negative) and were asked to place each word on a grid with two axis describing self-descriptiveness (x-axis ranging from “Very much like me” to “Not at all like me”) and valence (y-axis ranging from “Very negative” to “Very positive”). A valence score was derived based on the number of positive and negative words endorsed as “like me”. An index of cohesiveness of self-concept was calculated based on the mean Euclidean distance between adjective ratings in the two dimensional grid-space, with separate scores for positive and negative content. A smaller mean interstimulus distance indicates a greater degree of interconnectedness (Dozois & Dobson, 2001). Fragmentation of self-concept reflected

extreme self evaluations, akin to dichotomous thinking, and was computed based on the standard deviation of the distances between each participant's self-endorsed adjectives; again, this was calculated separately for positive and negative content. In line with negativity bias, (1) greater BPD severity was associated with less positive and more negative content of self-concept, consistent with negativity bias; this effect was driven by depression symptoms, and (2) positive content was more diffuse and negative content more interconnected at higher levels of BPD severity, but for positive content, this was most clearly linked to comorbid depression features. In line with dichotomous thinking, BPD severity (over and above depression symptoms) was uniquely associated with greater fragmentation of self-concept for positive and negative content.

Problematically, the scales used by Oshio (2009, 2012) and by Evans et al. (2015) are self-report measures and therefore cannot inform us on spontaneous evaluations in BPD. Moreover, they cannot specifically distinguish between the notions of dichotomous thinking, splitting and negativity bias, which are under scrutiny in the present thesis.

Laboratory studies on evaluations of others in BPD. Arntz and colleagues developed a method that aims at examining whether evaluations of others in BPD are best characterized as splitted (i.e. extreme all-good *or* all-bad evaluations), dichotomous (i.e. extreme thinking comprising *both* positive and negative evaluations), or negative (i.e., all-bad evaluations). The method involves asking participants to evaluate a target person; participants must place this person on a continuum between a negative pole and a positive pole for 22 different qualities (i.e. not jealous-jealous, reliable-unreliable). A first study using this method compared three groups: BPD, cluster C, normal. Participants were presented with different relational film clips depicting either emotional situations assumed to trigger BPD cognitive-emotional disturbances (i.e., BPD-specific clips; e.g. abandonment, rejection), emotional situations not specific to BPD or non-emotional situations. Patients with BPD made more extreme (dichotomous) ratings of characters from the BPD-specific film clips, with no evidence of splitting (Veen & Arntz, 2000). However, there was no index of negativity in this initial study by Veen and Arntz (2000). Another research team used a similar methodology and replicated those results and additionally found greater dichotomous thinking in BPD patients in an emotional situation involving positive emotions (Napolitano & McKay, 2007). Again, this study did not include an index of negativity. A recent study by Arntz and Haaf (2012) used a more natural

methodology involving real-time patient-therapist phone interactions, where the therapist displayed a rejecting, an accepting or a neutral attitude toward the participant. In this study, an index of negativity bias was included, as well as indices of splitting and dichotomous thinking. Again, dichotomous rather than splitted or negative evaluations of the therapist better characterized and discriminated patients with BPD from participants from cluster C and normal controls.

Despite the above findings showing that dichotomous thinking best characterizes and discriminates BPD, three studies applied variations to the methodology originally developed by Veen and Arntz (2000) and found evidence for negativity bias instead (Arntz & Veen, 2001; Sieswerda, Arntz, & Wolfis, 2005; Sieswerda, Barnow, Verheul, & Arntz, 2013). In one study, Sieswerda et al. (2013) utilized movie clips, similarly to Veen and Arntz (2000), and added an index of negativity bias (which was not included in the original study). Accordingly, they investigated whether extreme evaluations in BPD are dichotomous, negativistic, or splitted using BPD-specific and non-specific film clips, with characters appearing in negative, positive, or neutral roles. Interpersonal evaluations of patients with BPD were (1) negativistic; (2) schema related (emerged in BPD-specific situations); and (3) partially related to traumatic childhood experiences. Negative evaluations of caring characters in an intimate context particularly characterized BPD, which again contrasts with the original study by Veen and Arntz (2000). However, Veen and Arntz (2000)'s study was not designed to examine negativity bias, as it included no index to measure this construct. Also, Sieswerda et al. (2013) found no evidence for dichotomous thinking in BPD, which again contrasts with the original study by Veen and Arntz (2000). As a possible explanation of this inconsistency, Sieswerda et al. (2013) concluded that smaller between-group contrasts on the dichotomous thinking scale in this study “resulted mainly from high levels of dichotomous thinking of the control participants. Unlike other studies, BPD and control patients in the current study had more comparable levels of personality psychopathology (p. 48)”. Another study used an open-response format rather than the close-response format used in Veen and Arntz (2000) and found that negative rather than dichotomous evaluations of film characters better discriminated BPD from cluster C and normal participants (Arntz & Veen, 2001). Moreover, when evaluative responses to non-interpersonal situations (i.e., computer games) were examined instead of evaluative responses to movie characters, BPD appeared to be

characterized best by a general negativity bias, more than by dichotomous thinking or splitting (Sieswerda et al., 2005). This suggests that the tendency of individuals with BPD to evaluate stimuli through a negative lens may not be specific to interpersonal situations, but may affect a broader range of stimuli, including non-interpersonal stimuli. In addition, two studies investigating thinking style with the *Personality Disorder Belief Questionnaire* (Dreessen & Arntz, 1995a) also found that a negativity bias discriminates BPD from other personality disorders (Arntz, Dreessen, Schouten, & Weertman, 2004; Arntz, Weertman, & Salet, 2011). Similarly, another study found that negative rather than dichotomous evaluations discriminate individuals with BPD from cluster C and normal participants with an experimental design in which participants evaluated other people on different qualities taken from the Big-five theory of personality (Barnow et al., 2009). In sum, laboratory studies on dichotomous thinking, splitting and negativity have obtained inconsistent findings, with three studies showing evidence for dichotomous thinking (Arntz & Haaf, 2012; Napolitano & McKay, 2007; Veen & Arntz, 2000) and six studies showing evidence for negativity bias in BPD (Arntz et al., 2004; Arntz & Veen, 2001; Arntz et al., 2011; Barnow et al., 2009; Sieswerda et al., 2005; Sieswerda et al., 2013). However, among the former three studies, only one included a measure of negativity bias. The question of whether the inconsistency of those findings is method-dependent or due to type-I errors possibly because of low power remains open and warrants further research (for a discussion on power and type-I errors see LeBel, Campbell, & Loving, in press). In addition, the above studies relied exclusively on self-report methods and on relatively small groups; therefore, before concluding on the interplay between borderline personality and evaluations of others, it is important to consider other studies that included larger samples and more diversified instrument tools (including implicit and explicit measures). Moreover, the fact that none of the studies by Arntz et al. concluded that BPD is characterized by splitting raises a few questions. First, the index of splitting used in the studies by Arntz et al. did not assess temporal instability, as it focused solely on polarity (i.e., all-positive or all-negative evaluations at a given moment in time). Therefore, this index should not be considered as a complete measure of splitting. Moreover, Arntz and Haaf (2012) did find increased splitting in participants with BPD relative to clinical and healthy controls, but they concluded that evidence for splitting was not convincing because post-hoc tests revealed that “extreme ratings by BPD-patients were not completely determined by splitting, or where

evidence for splitting was strong, it was not specific to BPD” (p. 713). Finally, it would be premature to reject the notion of splitting, as other studies have found evidence for splitting using different methods (Beeney, Hallquist, Ellison, & Levy, 2016; Vater, Schroeder-Abe, Weissgerber, Roepke, & Schuetz, 2015).

In fact, Beeney et al. (2016) and Vater et al. (2015) have recently utilized different kinds of measurement techniques to achieve a comprehensive understanding of evaluations of self and others in individuals with BPD. Notably, a recent study combined a card sorting task, a maintenance task, a questionnaire and neural imaging techniques to examine evaluations of self and others in individuals with BPD (Beeney et al., 2016). In the card sorting task, participants were provided a deck of 40 trait cards (20 positive and 20 negative) and asked to sort the cards into different groups, with each group describing a distinct aspect of self. Three separate indices of self evaluations were derived based on how participants sort the cards: (1) the complexity index reflects the number of and overlap between self-aspects, (2) the integration index captures how much the self-concept is organized by valence and (3) the valence index reflects the proportion of all cards sorted that are positive vs. negative. In the maintenance task, participants were presented with 37 different personality traits twice, separated by 3 hours, and were asked to rate how well the traits described their own personality and the personality of a close friend. This task assessed the consistency of evaluations over time, that is, the extent to which participants maintained similar evaluations of self and of their close friend across the 3 hour period. Also, the Differentiation of Self Inventory (DSI; Skowron & Friedlander, 1998) was used as a questionnaire measure assessing four aspects of differentiation in current relationships and fMRI data was collected while participants completed a task designed to elicit evaluations of self and a close friend. Results showed that individuals with BPD sorted self-aspects in a way that suggests more complexity of self evaluation, but also less integration and more negative valence overall. On the maintenance task, individuals with BPD showed less consistency in their evaluations of self and others over the 3-hour period, but only for abstract, personality-based evaluations. Performance on this measure mediated between-groups brain activation in several areas supporting social cognition. Additionally, the BPD group showed main effects in brain areas crucial to social cognition. On the self-report measure, the BPD group evidenced less self-other differentiation and this effect was driven by the subscales on emotional reactivity (e.g.,

“If someone is upset with me, I cannot seem to let it go easily”) and emotional cutoff (e.g., “When one of my relationships becomes very intense, I feel the urge to run away from it”). The pattern of low integration and low consistency support the view that BPD is characterized by splitted evaluations of self and others; there was also evidence for negative self evaluations on the card-sort task. Unfortunately, Beeney et al. (2016) did not include any index to assess negativity in evaluations of others. Vater et al. (2015) also used a card-sort task to examine self evaluations in BPD. Supporting the view that BPD is characterized by a negativity bias and splitting, patients with BPD had the most compartmentalized self-concept structure and a higher proportion of negative self-attributes relative to both a non-clinical and a depressed control group. Moreover, BPD patients rated negative self-aspects as more important than positive ones relative to non-clinical controls. These two studies used different types of measurement techniques to examine evaluations of self and others in BPD; in doing so, they yielded evidence for low integration, low temporal consistency (i.e., instability) and low differentiation in individuals with BPD. These findings are consistent with the view that BPD involves splitted evaluations of self and others. There were also evidence for more negative evaluations of self, but unfortunately there was no index to examine negative evaluations of others. Also, notwithstanding the interest of combining different types of measurement techniques, these two studies did not utilize implicit measures and therefore cannot directly inform us about the nature of spontaneous evaluations in BPD. Another limit of the laboratory studies reviewed thus far is that they may therefore not accurately capture evaluations of self and others as they arise in real-life interactions among individuals with BPD.

Ecological momentary assessment of evaluations of others in BPD. Apart from the studies in the laboratory, one line of research has utilized field-based approaches to assess within- and between-days variations in perceptions, affects and interpersonal disturbances as they occur in real-life interactions of individuals with BPD. They used so-called *Ecological Momentary Assessment* (EMA) methods (sometimes termed *experience sampling*). The basic principle of EMA is to keep track of within-person real-life variations using repeated measurements (for a review see Santangelo, Bohus, & Ebner-Priemer, 2012). For instance Coifman et al. (2012) examined the relationships between polarity of affective and relational experiences, interpersonal stress and impulsive behaviors. Polarity was indexed by the extent to which positive and negative experiences are separated (i.e., strong positive experiences with

low negative experiences or strong negative experiences with low positive experiences). Polarity in reports of affective and relational experiences was greater in BPD ($n = 65$) relative to healthy controls ($n = 61$) and it increased during heightened interpersonal stress; in addition, increased polarity in reports of affective (in low-stress contexts) and relational experiences (in high-stress contexts) predicted increased rate of reports of impulsive behaviors (e.g., self-injury, substance abuse). This study supports the notion that BPD involves a difficulty at integrating positive and negative affects and evaluations of others. Moreover, this effect emerged under social stress and contributed to impulsive behaviors. Other research showed other kinds of dysfunctional responses under social stress in individuals with BPD, such as lower levels of endocrine stress markers (i.e., attenuated cortisol responses), coupled with maladaptive appraisal of the upcoming stressor and higher subjective distress and emotional responses (Nater et al., 2010; Scott, Levy, & Granger, 2013).

In line with a negativity bias, a study on dysfunctions within romantic relationships showed that borderline personality features were associated with reporting greater negative impact and greater emotional loss to both partner-initiated negative and positive experiences (Bhatia, Davila, Eubanks-Carter, & Burckell, 2013). Moreover, these effects remained significant even when controlling for relationship satisfaction, total number of relationship experiences, and depressive symptoms. However, borderline personality features were not associated with reporting greater negative impact and emotional loss in response to self-initiated negative and positive experiences. The authors interpreted these results as evidence that individuals with borderline personality features have a negative interpretation bias to both negative and positive experiences and that the effect is generally specific to partner-initiated experiences. Finally, another experience sampling study examined instability of self-esteem and affect in borderline personality and found that individuals with borderline personality features possess (1) unstable low self-esteem, (2) negative affect that is high and unstable and (3) self-esteem and feelings of rejection that are labile in response to daily interpersonal stress (Zeigler-Hill & Abraham, 2006). In sum, the above studies examined real-life interactions in BPD; in doing so, they shed new light on the interplay between negative evaluations of others, affect and behaviors in borderline personality and brought evidence for negativity biases and splitting (instability of self-esteem and polarity of relationship experiences). Notwithstanding these contributions, one

shortcoming of these studies is the lack of instruments to assess spontaneous evaluations of others.

Stroop studies. Another line of research assessed evaluations of emotional material and self evaluations in borderline personality using performance-based cognitive tasks such as the evaluative priming task (Fazio, Sanbonmatsu, Powell, & Kardes, 1986), the implicit association test (A. G. Greenwald et al., 1998), the Stroop task (e.g., Amir et al., 1996), the dot-probe task (e.g., Asmundson & Stein, 1994) and the go/no-go association task (Nosek & Banaji, 2001). These measurement procedures can be described as indirect because they do not rely on participants' self-assessment of the to-be-measured psychological attribute but on their behavioral responses to the attribute instead. However, the fact that a measure qualifies as indirect does not mean that it necessarily qualifies as implicit: "Specifically, measurement outcomes may be described as implicit if the impact of the to-be-measured psychological attribute on participants' responses is unintentional, resource-independent, unconscious, or uncontrollable" (Gawronski & De Houwer, 2014, p. 284). For example, a measure of partner evaluations may be described as implicit if it reflects participants' partner evaluations even when they do not have the goal to express these evaluations (i.e., unintentional) or despite the goal to conceal these evaluations (i.e., uncontrollable). Given these considerations, although cognitive tasks such as the Stroop task and the the dot-probe task are indirect, these tasks should not be considered as implicit measures without empirical evidence that the impact of the to-be-measured psychological attribute on participants' responses on these tasks is unintentional, resource-independent, unconscious, or uncontrollable. Moreover, these tasks have not been designed to capture such spontaneous responses, but rather to measure cognitive processes such as attention and inhibition (Rullkotter, Markowitsch, & Driessen, 2011). Past research that utilized performance-based cognitive tasks in BPD has mostly relied on the Stroop task; therefore, we will discuss these studies in more details.

The Stroop task measures the latency to name an attribute of a word (or picture), such as ink color. Time to name the ink color is compared when the stimuli are positive, negative or neutral. If participants with BPD are slower to name color of negative words, this suggests that they have more difficulties inhibiting or disengaging their attention from negative information, in line with a negativity bias affecting attention and inhibition processes. The phenomenon of slowing caused by certain contents on the Stroop task is called the *interference effect*.

Arntz et al. conducted different studies to examine Stroop interference in BPD. They compared the interference caused by neutral stimuli, schema-unrelated emotional stimuli and schema-related emotional stimuli (e.g., feeling powerless, unacceptable and feeling like others are malevolent). In one study, BPD patients displayed interference to negative stimuli relative to non-patient controls (Arntz, Appels, & Sieswerda, 2000). Schema-related negative words did not cause greater interference than schema-unrelated negative words, suggesting that the interference effect was driven by a generalized negativity bias rather than by negativity to BPD-related content specifically. Also, BPD and Cluster C patients did not differ, suggesting that the observed pattern of negativity was not specific to BPD. These effects emerged with supraliminal stimuli, whereas subliminal stimuli failed to yield any significant effect. Another study by Sieswerda, Arntz, and Kindt (2007) also showed evidence for a generalized negativity bias in BPD: patients with BPD showed greater interference for both schema-related and schema-unrelated emotional stimuli relative to non-patient controls. Interestingly, this study also examined the effect of therapy on Stroop interference and showed that interference was completely reduced to normalized levels in recovered patients ($n=6$), but not in non-recovered patients ($n=10$) at the end of treatment.

Whereas the two above Stroop studies suggest difficulties disengaging attention from negative information in general (i.e., consistent with a generalized negativity bias in BPD), three other Stroop studies yielded evidence that such difficulties emerge only for certain types of negative material (i.e., consistent with more specific negativity biases in BPD). Wingenfeld, Mensebach, et al. (2009) found greater interference in BPD patients relative to control participants only for words related to personal negative life events that were currently relevant (vs. words related to personal negative life events that were not currently relevant, negative words and neutral words). Further analyses revealed that BPD patients with concurrent PTSD drove this effect. Another Stroop study compared BPD patients to clinical (cluster C and axis I disorder) and normal participants; BPD patients displayed significantly greater interference to schema-related negative stimuli, consistent with the view that BPD is characterized by a specific kind of negativity bias. This effect emerged with supraliminal stimuli and was independently predicted by childhood sexual traumas and BPD anxiety symptoms. When the stimuli were presented subliminally, both BPD and axis I disorder patients showed a trend for a bias for negative schema-related stimuli (Sieswerda, Arntz, Mertens, & Vertommen, 2007).

Witthoft, Borgmann, White, and Dyer (2015) also found evidence for a negativity bias toward specific themes in BPD. They compared patients with PTSD with and without comorbid BPD to healthy controls on an emotional Stroop task examining body perceptions. They found significantly stronger interference to body related words; this effect emerged only among patients with concurrent PTSD and BPD in comparisons with healthy controls.

Whereas the above studies yielded evidence for generalized or specific negativity biases in BPD, other studies brought mixed or disconfirming evidence on this matter. Portella et al. (2011) did not find evidence for negativity biases on Stroop interference scores: although BPD patients displayed an interference effect for borderline-related words and a marginal effect for negative words relative to control participants, they also showed a significant interference effect for neutral words, along with a marginal interference effect for positive words. Similarly, Wingenfeld, Rullkoetter, et al. (2009) observed overall slower reaction times in BPD patients compared to healthy controls and responses to negative words were not associated with increased slowing. In other words, BPD patients showed increased interference in general rather than increased interference to negatively valenced words specifically. Hence, these two studies suggest that BPD involves general inhibition difficulties rather than difficulties inhibiting negative material specifically, contrary to the negativity bias hypothesis. However, Wingenfeld, Rullkoetter, et al. (2009) also measured neural activity during the Stroop and obtained more conclusive results; in fact, patterns of brain activation were indicative of emotion dysregulation in BPD patients compared to control participants during trials with negative words relative to trials with neutral words; such patterns emerged for general as well as specific negative words, in line with a general negativity bias. In another Stroop study, Winter, Krause-Utz, et al. (2015) compared BPD patients who underwent a dissociation induction, BPD patients without dissociation induction and healthy controls. Results were mixed: there was no differences on Stroop interference scores between BPD patients without dissociation induction and healthy controls; however, BPD patients who underwent a dissociation induction showed overall slower and less accurate responses as well as increased reaction times for negative versus neutral words in comparison with BPD patients without dissociation induction. These results do not support the view that BPD is generally characterized by negativity biases, but instead suggest that BPD can lead to such biases during episodes of dissociation. In addition, this study measured neural activity during the Stroop

task: compared to control participants, BPD patients without dissociation induction displayed patterns of brain activation consistent with higher vigilance for emotional stimuli and increased inhibition of positive stimuli, whereas BPD patients with dissociation induction displayed patterns of brain activation consistent with greater effort, but less success, at inhibiting negative stimuli. Domes et al. (2006) also obtained mixed findings. They found no difference between BPD and control participants on interference during the Stroop task; nevertheless, BPD patients showed reduced inhibition of negative material in other performance-based tasks (the directed forgetting task and the negative priming task) and difficulties remembering positive words in the directed forgetting task, which is consistent with a negativity bias. To summarize, based on participants' behavioral responses (i.e., response latencies), Stroop studies can complement self-reports measures and bring useful information on attention and inhibition processes underlying evaluation of emotional information in BPD. While two Stroop studies supported the view that BPD is characterized by difficulties disengaging from negative information in general (in line with a generalized negativity bias) (Arntz et al., 2000; Sieswerda, Arntz, & Kindt, 2007), three studies found such difficulties only for specific material (in line with a specific negativity bias) (Sieswerda, Arntz, Mertens, et al., 2007; Wingenfeld, Mensebach, et al., 2009; Witthoft et al., 2015), and four studies yielded either no or equivocal evidence for greater difficulty disengaging from negative information in BPD based on Stroop interference scores (Domes et al., 2006; Portella et al., 2011; Wingenfeld, Rullkoetter, et al., 2009; Winter, Krause-Utz, et al., 2015). However, among these four studies, one found behavioral evidence for negativity bias with other tasks than the Stroop and two found neuroimaging evidence suggesting emotion dysregulation during the Stroop task in BPD patients compared to controls (Wingenfeld, Rullkoetter, et al., 2009; Winter, Krause-Utz, et al., 2015). Complicating the picture even more, three other studies on emotion processing have used other cognitive tasks and have found that individuals with BPD show longer reaction times during emotional distraction in attention tasks with emotional pictures as distractors (Hagenhoff et al., 2013; Krause-Utz et al., 2012; I.-A. von Ceumern-Lindenstjerna et al., 2010), suggesting a difficulty inhibiting irrelevant emotional material in general (positive and negative) rather than negative material specifically. Another study by Winter, Herbert, et al. (2015) used a priming task to assess evaluations of self and others and found that BPD patients judged positive and neutral words as more negative than

healthy control participants when the words were preceded by a self-referential pronoun or no reference, but not when preceded by an other-referential pronoun. Hence, the findings of this study point to a negative evaluation bias for positive, self-referential information in BPD, but not for information pertaining to others. Interestingly, this task is similar to evaluative priming tasks that are used in implicit research (Herring et al., 2013), but Winter, Herbert, et al. (2015) did not describe the task as implicit as it was not their aim to measure spontaneous evaluations in their study. In fact, participants had no time limit or time pressure to evaluate the positive and negative nouns, which left them the opportunity to intentionally control their response. Such time constraints are crucial in priming tasks that are designed as implicit measures (i.e., typically 250 or 300 ms; Herring et al., 2013) as they serve to increase the likelihood that the impact of the to-be-measured psychological attribute on participants' responses is unintentional, resource-independent, unconscious, or uncontrollable (Gawronski & De Houwer, 2014). In sum, three conclusions can be drawn from the studies that used performance-based indirect measures to investigate cognitive processes underlying evaluations of emotional material in BPD: (1) although these studies used indirect measures, these measures do not qualify as implicit and therefore cannot inform on the nature of spontaneous evaluations in BPD; (2) given inconsistent findings, these studies do not allow concluding on potential inhibition or attention deficits underlying negativity biases in BPD; (3) in addition, no Stroop study has examined evaluations of interpersonal stimuli (e.g., faces) in BPD and therefore the results of these studies might not directly apply to interpersonal situations in BPD.

IAT studies. Apart from the Stroop task, another indirect measure that has been used in several BPD studies is the implicit association test (IAT; A. G. Greenwald et al., 1998). While Stroop studies examined difficulties disengaging from irrelevant emotional material such as negative and positive words, the IAT is one of the most frequently used paradigms to in implicit research and there is ample evidence supporting its validity as an implicit measure (Anthony G. Greenwald, Poehlman, Uhlmann, & Banaji, 2009; Nosek, Greenwald, & Banaji, 2005). IAT studies did not focus on spontaneous evaluations of other people in BPD and their focus was on spontaneous evaluations of self. Nevertheless, although the results of these studies may not apply to spontaneous evaluations of other people in BPD, until such research

exists, these studies may give interesting insights into the way borderline personality influences evaluations at the spontaneous level.

In the IAT, participants are asked to classify stimuli into superordinate categories while the categories are paired in a way that either matches hypothesized automatic associations or that contradicts those associations. For instance, in BPD studies, the IAT has been used to examine shame-self associations. Accordingly, different words reflecting the categories “self”, “other”, “anxiety” and “shame” are presented to participants. On one half of the trials, the strength of self-shame and other-anxiety associations are examined, such that participants are asked to press one key for “self” and “shame” words and the other key for “other” and “anxiety” words. On the other half of the trials, the pairing is reversed; this time, the strength of self-anxiety and other-shame associations is examined, such that participants are asked to press one key for “self” and “anxiety” words and the other key for “other” and “shame” words. In this case, faster reaction times to self-shame relative to self-anxious pairings would typically be interpreted as evidence of stronger self-shame associations, suggesting that spontaneous evaluations of self are tainted by shame. In this vein, Rüsç et al. conducted two studies where they compared four conditions on the IAT to examine: (1) shame-self associations, (2) anxious-self associations, (3) anxious-best friend associations, and (4) shame-best friend associations. In one study, they found evidence for stronger shame-self associations relative to anxious-self associations in BPD individuals relative to social phobic and healthy control individuals (Rusch et al., 2007). In another study, they replicated these effects and found that these emerged independently of diagnosis of post-traumatic stress disorder (Rüsç et al., 2007). Although Rüsç et al. concluded that the results reflected increased shame-self associations in BPD relative to control participants, the results on the IAT may also reflect spontaneous evaluations of other people. In fact, these results may have been influenced by potential associations between “best friend” on the one hand and “anxiety” and “shame” words on the other hand. However, because the IAT relies on relative comparisons, in this case it is impossible to disentangle the contribution of best friend vs. self associations on the IAT. More recently, Ritter et al. (2014) did not replicate the two studies of Rüsç et al.; instead, they found that BPD patients scored lower than NPD patients on spontaneous shame-self associations on the IAT and BPD patients did not differ from healthy controls; these results are inconsistent with the notion that BPD is characterized by negativity

biases (in this case, negative spontaneous self evaluations). Other studies simultaneously examined spontaneous and deliberate self-esteem by combining the IAT and a self-report measure. One study found evidence for a stronger discrepancy between the IAT and questionnaire measure of self-esteem in BPD individuals (I.-A. von Ceumern-Lindenstjerna et al., 2010). In addition, Hedrick and Berlin (2012) found lower spontaneous self-esteem on the IAT (stronger self-negative associations) in BPD patients relative to patients with depersonalization disorder, but BPD patients did not differ from healthy controls, which is contrary to the view that BPD is characterized by negative spontaneous self evaluations. Vater et al. (2013) also combined the IAT with a self-report measure of self-esteem and showed that BPD patients scored lower than NPD patients on spontaneous self-esteem (i.e., self-negative associations) and lower than patients with narcissistic personality disorder (NPD) and healthy controls on self-reported self-esteem. Finally, another IAT study examined spontaneous and deliberate reactions to abuse in BPD patients relative to non-clinical and clinical controls, that is, antisocial personality disorder (ASPD) and Cluster C personality disorder patients. BPD patients reported more negative affect, more maladaptive (and less adaptive) schema modes and displayed greater psychophysiological activation and spontaneous self-abuse associations after the abuse induction (movie clip). The ASPD-group was comparable to the BPD group on abuse-self associations but did not show self-reported and physiological hyper-reactivity to the abuse induction. The pattern of findings in this study suggests that BPD and ASPD-patients are alike in their spontaneous self-abuse associations in reaction to abuse, but can be differentiated in their self-reported and physiological response patterns (Lobbestael & Arntz, 2010). Together, the above IAT studies suggest that BPD may involve negative spontaneous evaluations of self, in line with cognitive (A. T. Beck & Freeman, 1990; Aaron T. Beck et al., 2004) and psychodynamic (Kernberg, 1986) theorizing, but more research is needed before drawing firmer conclusions given that some studies obtained mixed or equivocal findings. Moreover, notwithstanding the contribution of these findings for spontaneous evaluations in BPD, an important open question here, is whether borderline personality may involve negativity biases not only in spontaneous evaluations of self but also in spontaneous evaluations of other people.

Research strategy and hypotheses

Three conclusions can be drawn from our review. First, despite evidence for splitting and dichotomous thinking, such evidence remains rare and findings are inconsistent. Second, although numerous studies showed evidence of negativity biases in BPD (including negative evaluations of others), there are also numerous studies that failed to obtain clear evidence for negativity biases. These mixed findings with respect to negativity biases in BPD can be interpreted in different ways: (a) negativity biases may be present only in certain individuals with BPD, (b) individuals with BPD may display negativity biases only in certain contexts or (c) in certain aspects of their social-cognitive functioning (e.g, evaluations of self vs. evaluations of others, spontaneous evaluations vs. deliberate evaluations) or (d) results failed to replicate because of low-powered studies. Third, and mainly, no study has utilized implicit measure to directly examine spontaneous evaluations of others in borderline personality. The present thesis aims to fill this gap in the BPD literature and therefore combined implicit and explicit measures to examine the interplay between spontaneous and deliberate evaluations of others in borderline personality. In fact, past research has pointed to the importance of studying both spontaneous and deliberate evaluations. Notably, spontaneous and deliberate evaluations result from different processes, they operate under different conditions (Gawronski & Bodenhausen, 2011) and they influence different kinds of behaviors (Perugini, 2005). Also, our research relied on experimental conditions to manipulate spontaneous and deliberate partner evaluations (Study 2) as well as spontaneous and deliberate evaluations of a new target person (i.e., evaluation of a movie character in Study 1). In addition, we focus on potential patterns of convergence or dissociation between implicit and explicit measures. When implicit and explicit measures converge, it strengthens the obtained findings; in the case of evaluations of others, implicit-explicit convergence would also show that such evaluations are shaped at relatively early (i.e., associative) and late (i.e., propositional) processing stages in borderline personality. In contrast, if implicit and explicit measures were to diverge, this could fuel new predictions and reveal aspects of evaluations of others that were previously unforeseen in borderline personality (for the example of implicit racism see A. G. Greenwald et al., 1998). Therefore, by examining the interplay between spontaneous and deliberate evaluations, the present research can bring new insights into the dynamic of interpersonal dysfunctions in borderline personality.

We combined two complementary studies. In Study 1, participants watch two clips from the same movie (*Like Crazy*). One clip depicts a negative situation of rejection and betrayal within the romantic relationship and the other clip depicts a positive romantic interaction. The same two characters are interacting in both clips. All participants watch both clips (order counterbalanced). Spontaneous and deliberate evaluations of the target character are measured twice, that is, once after each clip. Assuming that participants do not know the target character or, at least, have little prior associations to the character, the focus here is on how borderline personality influences impression formation toward an unknown person. In Study 2, participants imagine either (a) that their partner rejected them or (b) sought connection or (c) imagined no scenario (control condition). Accordingly, while Study 1 examines impression formation, Study 2 examines partner evaluations. Specifically, we ask whether the obtained patterns of evaluations generalize between the contexts of impression formation and romantic relationships in borderline personality. Also, Study 2 includes a measure of spontaneous evaluations of self. Moreover, the two studies use different implicit measures: Study 1 uses an evaluative priming task and Study 2 uses a name-letter task (which are described in more details the respective studies). Different measurement tools should bring similar findings to the extent that they measure a similar phenomenon, but we cannot determine a priori whether impression formation and partner evaluations tap into the same aspects of borderline personality. Thus, the results obtained in the present research will also help to explore potential differences and similarities between these two interpersonal contexts in relation to borderline personality. Finally, because BPD is not a discrete trait, but rather the extreme manifestation of a normally distributed personality disposition, both studies used a dimensional approach to borderline personality, where the notion of *borderline personality (BP) features* represent a continuum of severity and is measured with questionnaires (Edens et al., 2008; Haslam, 2003; Rothschild et al., 2003).

Both Studies 1 and 2 examine the hypothesis that borderline personality involves negative spontaneous and deliberate evaluations of others. Apart from this similarity, there are also divergences between Studies 1 and 2: Study 1 measures spontaneous and deliberate evaluations of a movie character at two different times (i.e. after each clip) for every participant; thus, it also allows examining the hypothesis that borderline personality involves unstable evaluations of the movie character (i.e., switching from more positive to more

negative evaluations after the positive and negative clips, respectively). In contrast, Study 2 measures spontaneous and deliberate evaluations at only one occasion. Therefore, it cannot examine instability in terms of intensity of switches between different measurement occasions; instead, Study 2 tests the hypothesis that borderline personality involves polarized spontaneous and deliberate partner evaluations. Specifically, according to the polarity hypothesis, we expect more positive partner evaluations in participants with high BP features assigned to the closeness scenario and more negative partner evaluations in participants with high BP features assigned to the rejection scenario, relative to participants in the control condition (no scenario). In sum, both studies allowed investigating spontaneous and deliberate evaluations of others from two different angles (i.e., within two types of interpersonal contexts). The studies are coherent and complementary in that they aim at dressing a nuanced and more complete picture of two kinds of mechanisms (i.e., spontaneous and deliberate evaluations) potentially involved in interpersonal dysfunctions in borderline personality.

Article 1

Negativity bias and instability in spontaneous and deliberate evaluations of others: The role of
borderline personality disorder features

Mongeon, F., Gawronski, B., Gagnon, J. (2016). Negativity bias and instability in spontaneous
and deliberate evaluations of others: The role of borderline personality disorder features.

Abstract

This study tested the hypotheses that borderline personality (BP) features are characterized by a negativity bias and instability in spontaneous and deliberate evaluations of others. Undergraduate women (N = 204) watched two movie clips depicting either positive or negative conjugal interactions. Spontaneous and deliberate evaluations of the male character were assessed after each clip with an Evaluative Priming Task and a self-report measure, respectively. Participants with high BP features showed unstable spontaneous evaluations. Results revealed a non-significant trend toward more negative spontaneous evaluations after the negative clip and less positive and more negative deliberate evaluations after watching the positive clip first relative to participants with low BP features. These results provide preliminary evidence that impression formation in borderline personality may be characterized by negative and unstable evaluations that are shaped at least in part at earlier processing stages.

Keywords: borderline personality; spontaneous evaluation; deliberate evaluation; negativity bias; instability; social cognition; implicit measure; explicit measure

1. Introduction

Borderline personality disorder (BPD) is a disorder characterized by significant interpersonal dysfunctions. Suicidal and self-damaging behaviors (Soloff, Lis, Kelly, Cornelius, & et al., 1994) and a drastic increase in prevalence in clinical settings exemplify the treatment complexities associated with this disorder: despite a prevalence of 1-2% in the general population (Swartz, Blazer, George, & Winfield, 1990), BPD has been shown to occur in 20% of intern patients in psychiatry (Zanarini, Frankenburg, Khera, & Bleichmar, 2001). Clinical theories of borderline personality disorder (BPD) emphasize the key role of negativity and instability in the identity and interpersonal dysfunctions of individuals with BPD (Beck, Freeman, & Davis, 2004; Kernberg, 1986; Leichsenring, Leibing, Kruse, New, & Leweke, 2011; Lieb, Zanarini, Schmahl, Linehan, & Bohus, 2004). To better understand negativity and instability, BPD research has combined different techniques, including self-report, behavioral, and neurobiological techniques. These research tools have brought ample evidence of negativity biases in BPD. Recently, BPD research has begun to integrate another set of measurement instruments, namely *implicit measures*, to examine negativity at the spontaneous level of responding. So far, no study has combined implicit and explicit measures to examine the interplay between spontaneous and deliberate responses in BPD. In contrast, social cognition research has fruitfully combined explicit and implicit measures to study a plethora of psychological phenomena (for a review, see Gawronski & Payne, 2010), including depression and anxiety (for a review, see Teachman, Joormann, Steinman, & Gotlib, 2012). Hence, the main goal of the present study was to fill this gap in the BPD literature by examining spontaneous and deliberate evaluations in borderline personality using both implicit and explicit measures.

There are several reasons to use implicit measures in BPD research. First, implicit measures capture spontaneous responses that are not necessarily reflected in explicit measures, and the combined use of both measurement instruments can shed light on the interplay between spontaneous and deliberate evaluations. According to the associative–propositional evaluation (APE) model, spontaneous evaluations assessed by implicit measures represent affective gut reactions resulting from the activation of mental associations in memory. In contrast, deliberate evaluations assessed by explicit measures represent more controlled beliefs and are

shaped at later processing stages. Because spontaneous and deliberate evaluations are the outcomes of different underlying processes, responses on implicit and explicit measures often diverge and the APE model provides explanations of such dissociations in terms of the respective processes and operating principles guiding the two kinds of evaluations (for a discussion and a review of the evidence supporting the model, see Gawronski & Bodenhausen, 2006; Gawronski & Bodenhausen, 2011). With this distinction in mind, we focus on two key features that are believed to characterize evaluations of others in borderline personality, namely the notions of *negativity bias* and *instability*.

2. Negativity bias and instability in BPD

Negativity bias and instability are hallmark features of BPD. The negativity bias refers to the tendency to evaluate stimuli through a negative lens. According to DSM-V, BPD individuals are selectively biased toward negative attributes in their evaluations of others (American Psychiatric Association, 2013, p. 766). Similarly, the cognitive theory of BPD hypothesizes that dysfunctions of individuals with BPD are at least partly caused by maladaptive cognitive schemas (basic cognitive structures in memory) that bias the evaluation and interpretation of environmental stimuli, including the basic assumption that the world is dangerous and malevolent (Beck et al., 2004; Beck & Freeman, 1990). Psychodynamic theories also highlight the phenomenon of negative or malevolent evaluations (Kernberg, 1986).

Different studies yielded evidence that individuals with BPD interpret reality through a negative lens. For example, Arntz and colleagues conducted a series of studies in the laboratory and found a stronger tendency in BPD to describe others in a more negative manner (Arntz & Veen, 2001; Sieswerda, Arntz, & Wolfis, 2005; Sieswerda, Barnow, Verheul, & Arntz, 2013; but see also Arntz & Haaf, 2012, for disconfirming evidence). Similarly, Barnow et al. (2009) showed that individuals with BPD formed more negative evaluations of neutral interpersonal stimuli. In addition, there is behavioral evidence of negativity biases in facial emotion recognition in BPD, such as a heightened sensitivity to the detection of negative emotions, along with a negativity or anger bias (for a review, see Domes, Schulze, & Herpertz, 2009). Neurobiological findings suggest that the negativity bias in BPD might stem from a

hyperreactivity of limbic brain areas and a hyporeactivity of prefrontal areas in response to negative socio-emotional stimuli as well to neutral social stimuli (for a review, see Krause-Utz, Winter, Niedtfeld, & Schmahl, 2014). There is also physiological evidence for a negativity bias in the patterns of physiological reactivity of individuals with BPD. For example, Matzke, Herpertz, Berger, Fleischer, and Domes (2014) found that individuals with BPD displayed reduced facial responding to positive social signals and increased facial responding to negative social signals. Moreover, there is ample evidence that rejection is a common theme to the negative reactions of individuals with BPD. For instance, they are more prone to perceive rejection when actually rejected, but also when not rejected (Renneberg et al., 2012), and to endorse beliefs and expectations that they will be rejected or abandoned (Dreessen & Arntz, 1995; Staebler, Helbing, Rosenbach, & Renneberg, 2011). Together, these findings obtained with diverse methods that individuals with BPD interpret social information through a negative lens and evaluate other people negatively.

Along with the negativity bias, another hallmark feature of BPD is instability, including instability in affect, relationships, and evaluations of the self and others. According to DSM-V, individuals with BPD view close relationships in extremes of idealization and devaluation and alternate between overinvolvement and withdrawal (American Psychiatric Association, 2013, p. 766). In its psychodynamic theory, Kernberg (1986) also described a tendency in individuals with BPD to switch between all-positive and all-negative experiences of self and others. Kernberg (1986) further hypothesized that instability in BPD stems from the mechanism of splitting; specifically, he argued that the extreme and polarized affective and interpersonal experiences of individuals with BPD, coupled with their difficulty integrating such experiences into more nuanced representations, leads them to form all-negative or all-positive evaluations of self and others and to switch between these two polarities. The cognitive notion of dichotomous thinking similarly captures the tendency of individuals with BPD to evaluate their experience through mutually exclusive categories (black or white) instead of more nuanced shades of grey (Beck et al., 2004; Beck & Freeman, 1990). Past research has obtained empirical support for the view that BPD involves splitted evaluations of self and others, as evidenced by polarized relationship experiences (Coifman, Berenson, Rafaeli, & Downey, 2012), less integration in self evaluations, less stability in evaluations of self and others over a 3-hour period (Beeney, Hallquist, Ellison, & Levy, 2016),

compartmentalized self-concept structure (Vater, Schroeder-Abe, Weissgerber, Roepke, & Schuetz, 2015) and greater diffusion of positive self evaluations and greater interconnection of negative self evaluations (Evans et al., 2015), as well as more extreme evaluations of film characters (Veen & Arntz, 2000) and alleged mental health worker trainees (Arntz & Haaf, 2012). Although these studies support the view that BPD involves polarized or extreme evaluations of others and difficulties integrating such evaluations, only two studies (Beeney, Hallquist, Ellison, & Levy, 2016; Coifman et al., 2012), to the best of our knowledge, have used temporal measurements and supported the notion that BPD indeed involves unstable evaluations of others, defined in terms of switches between positive and negative evaluations of others.

3. The present study

The present study was designed to further investigate evaluations of other people in borderline personality, expanding previous research in three ways. First, this is the first study to examine spontaneous evaluations of other people in borderline personality using an implicit measure. Previous studies have examined spontaneous evaluations of self in patients with BPD (e.g., Hedrick & Berlin, 2012; Rusch et al., 2007), but not spontaneous evaluations of other people.

Second, it investigated whether borderline personality involves not only more negativity in evaluations of other people, but also more instability. In fact, dichotomous thinking and splitting have been hypothesized to manifest not only as more negative evaluations, but also as switches between positive and negative evaluations. Moreover, instability is a DSM-V criterion of BPD (American Psychiatric Association, 2013). Despite ample evidence that BPD involves negativity biases, polarization (extreme evaluations) and difficulties integrating evaluations, only two studies, to the best of our knowledge, have found evidence that individuals with BPD display more instability in the way they evaluate others (Beeney et al., 2016; Coifman et al., 2012). These two studies focused on participants' real-life relationships. Complementing this approach, the present study involved a standardized procedure in the laboratory where all participants evaluated the same target person. In order to examine negativity and instability more thoroughly, the method also allowed distinguishing between

the positive and negative evaluative dimensions as well as between positive and negative interpersonal contexts.

Third, our study used a dimensional approach to borderline personality and assessed borderline personality by means of a questionnaire instead of a diagnostic interview. This was done to reflect the continuum of severity in the distribution of features of BPD. Previous research has brought empirical evidence that BPD is not a discrete trait, but rather the extreme manifestation of a normally distributed personality disposition (e.g., Hedrick & Berlin, 2012; Rusch et al., 2007). Furthermore, individuals who endorse significant levels of *borderline personality (BP) features* display traits and impairments that are qualitatively similar (though less intense) to clinical forms of the disorder (Edens, Marcus, & Ruiz, 2008; Haslam, 2003; Rothschild, Cleland, Haslam, & Zimmerman, 2003).

In sum, the present study aimed to test the following two hypotheses: BP features are related to making relatively *more negative* (Hypothesis 1) and *more unstable* (Hypothesis 2) evaluations of the evaluated person in all or certain interpersonal situations. These two hypotheses are not necessarily mutually exclusive as it is possible to observe overall more negativity, coupled with instability (i.e. unstable evaluations that vary within overall more negative ranges). For instance, there is evidence that individuals with BP features have unstable low self-esteem and negative affect that is high and unstable (Zeigler-Hill & Abraham, 2006). Based on the distinction between spontaneous and deliberate evaluations, the present study explored whether potential negative and unstable evaluations in borderline personality arise from relatively early or late processing stages.

4. Method

4.1. Sample and procedure

Participants were 204 undergraduate women recruited in psychology classes. The mean age was 22.18 ($SD = 3.561$, range = 18-41). Participants were told that the study examined personality and relationships. The Personality Assessment Inventory—Borderline Features Scale (PAI-BOR; Morey, 1991b) along with demographic questions was first sent online to participants who accepted to be contacted for the study. After completion of the first part, participants were invited to complete all other measures individually in the laboratory. The

time between completion of the PAI-BOR questionnaire and the other measures ranged from 0 day to 17 weeks ($M = 18.9$ days, $SD = 25.1$ days). During the laboratory session, all participants completed the implicit and explicit measures twice, once after the first clip and once after the second clip. The Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) and Self-assessment Manikin were both used as emotion variables that were assessed once after each of the two clips. The PANAS was administered together with the explicit measure. Half of the participants completed the PANAS after the implicit measure, following explicit-implicit counterbalancing. The Self-assessment Manikin served as a manipulation check, because it was administered immediately after participants watched the movie clips and before any other measure. Participants were fully debriefed following the completion of the experiment.

The PAI-BOR mean ($M = 24.79$, $SD = 11.37$) was larger than for the standardisation sample ($M = 18.03$, $SD = 10$; Morey, 1991b), but similar to the mean obtained by Trull (1995) using a nonclinical student population ($M = 27.23$, $SD = 10.87$). 32 participants (15.69% of the present sample) scored above threshold on the PAI-BOR ($\geq 70T$ or raw score ≥ 38) and the remaining of the sample (172/204, 84.31%) scored below threshold ($< 70T$ or raw score < 38). Previous research has shown that participants who score above this threshold display traits and impairments that are qualitatively similar (though less intense) to clinical forms of the disorder (Bagge et al., 2004; Jackson & Trull, 2001; Trull, Ueda, Conforti, & Doan, 1997). Therefore, this threshold provides a well-established benchmark point of reference. The distribution of BP features was normal (skew = .529; kurtosis = -.242), that is, absolute values of skew < 3 and kurtosis < 10 per guidelines provided by Kline (2011).

4.2. Design

Hypotheses were tested with an experimental design. The interpersonal situations consisted of two 5-minute film clips depicting (1) a negative conjugal interaction (negative clip) and (2) a positive conjugal interaction (positive clip) between two characters, a young adult male and a young adult female. Both clips were taken from the film *Like Crazy*. During the first half of the negative clip, the couple seems withdrawn. Then, the woman secretly looks at her boyfriend's cellphone and discovers an affectionate message from another woman. She confronts him about the message and the conversation escalates into a conflict involving

themes of reciprocal jealousy, betrayal and rejection. The positive clip depicts the first romantic date between the male and female protagonists. Both characters seem reciprocally joyful, charmed and absorbed in their interaction. Because gender could influence the way participants evaluate female or male characters in romantic interactions, we decided to measure evaluations of the male character among female participants to avoid potential confounds and to better capture the variance in the evaluations of the film character accounted for by BP features. All participants watched both clips in a within-participants design with the order of the two film clips counterbalanced across participants. The order of the two evaluation measures was also counterbalanced across participants. The main authors selected the film clips based on the two following assumptions: (1) these clips appear to depict realist interactions in a young couple and (2) the undergraduate participants in this study may potentially identify more easily to the characters because of their age similarity. These assumptions were not tested and the clips were not validated or used in prior studies.

4.3. Measures

4.3.1. BP features

The French-validated version of the Personality Assessment Inventory—Borderline Features Scale (PAI-BOR; Morey, 1991a) is a 24-item self-report measure used to assess BP features. Principal component analyses in a non-clinical sample ($n = 4682$) revealed a six-factor solution: 1) Control/impulsive behavior; 2) Mood instability; 3) Chronic emptiness/loneliness/boredom; 4) Separation and abandonment concerns; 5) Negative relationships; 6) Reckless spending (Jackson & Trull, 2001). The measure has shown good reliability (e.g., $\alpha = .84$) and the scores on the full scale have been established as a valid measure of BP features in non-clinical samples (Trull, 1995). The reliability was also good in the present study ($\alpha = .89$). We use the total score on the PAI-BOR as our measure of BP features.

4.3.2. Explicit measure

Participants rated the film character on 7-point scales describing various interpersonal qualities. The qualities were single words that reflected either positive or negative qualities. These words were selected by the present authors based on their judgments of positive and

negative personality traits that may possibly be associated with the male character. Table 1, below, shows the words. The words were not validated prior to the present study. 10 positive qualities described the character as a warm, affectionate and accepting person and 10 negative qualities described the character as a cold, rejecting and hostile person. Participants rated the film character and two baseline characters not seen in the film after each film clip (2 clips \times 20 qualities \times 3 characters = 120 scores). Figure 1, below, shows the characters' faces. Difference scores were calculated by subtracting the averaged ratings of the two baseline characters from the averaged ratings of the film character for each quality. Two positivity scores, one for each clip, were obtained by averaging the difference scores of the positive qualities. Higher scores reflect a more *positive* deliberate evaluation of the character relative to the baseline characters. Likewise, two negativity scores, one for each clip, were obtained by averaging the difference scores of the negative qualities; accordingly, higher scores reflect a more *negative* deliberate evaluation of the character relative to the baselines. These indices do not reflect evaluations of faces in general, but rather evaluations specifically related with the film character because subtracting evaluative judgments to the baseline faces allow controlling for potential biases related with evaluations of faces in general. In addition, the use of differences scores helps to increase the conceptual correspondence between the explicit and the implicit measure and to avoid method-related confounds in the comparison of measurement scores (see Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Payne, Burkley, & Stokes, 2008). Cronbach's α for the positive difference scores were .938 after the positive clip, .949 after the negative clip and .951 for both clips together; cronbach's α for the negative difference scores were .934 after the positive clip, .932 after the negative clip and .944 for both clips together.

4.3.3. *Implicit measure: Evaluative Priming Task*

Each trial of the Evaluative Priming Task (EPT) included the following sequence: (i) a fixation cross was presented for 500 ms, (ii) a prime (i.e., a picture of the film character or of a baseline character) was presented for 200 ms, and then (iii) a positive or negative target word (e.g., warm, cold) was presented and remained on-screen until the participant indicated whether the word was positive or negative using one of two response keys on the keyboard. If the response was incorrect, "ERROR!" was displayed for 1,500 ms. The intertrial interval was 500 ms. Each of the eight faces served as a prime on 20 trials, split between 10 trials with

negative and 10 trials with positive target words, summing up to a total of 160 trials presented in a computer-randomized order (same randomized order for all participants).

Mirroring the procedure for the explicit measure, priming indices were based on difference scores that aimed at isolating evaluations associated with the film character, while controlling for potential biases associated with evaluations of faces in general. Mean reaction times on trials where a given target word was preceded by faces of the film character were subtracted from the mean reaction times on trials where the same target word was preceded by faces of neutral baseline characters. Thus, higher scores indicate facilitated responses (i.e., shorter reaction times) to the target word when the target word followed the face of the film character relative to the faces of the baselines, suggesting a stronger association to the film character relative to the baselines. To obtain indices of positive and negative evaluations, separate difference scores were derived for trials with positive target words and for trials with negative target words. EPT trials with incorrect responses (3.3%) were excluded. To control for anticipations and outliers, response cutoffs were employed to exclude trials with reaction times longer than 1,500 ms (1.5% of valid trials). The percentage of excluded trials per participant and per EPT ranged from 0.00% to 20.00%. Although reliability estimates of the EPT tend to be lower compared to various other implicit measures, the EPT has been used extensively in social cognition research and a vast literature supports its validity as a measure of spontaneous evaluation (Fazio, Sanbonmatsu, Powell, & Kardes, 1986).

Table 1
List of positive and negative words used in the explicit and implicit measures

Positive words	Negative words
Accepting	Distant
Agreeable	Indifferent
Gentle	Blaming
Affectionate	Cold
Caring	Angry
Responsive	Dishonest
Kind	Rejecting
Lovely	Aggressive
Open	Disagreeable
Warm	Insensitive

Note: The original words used in the study are in French and were translated for this table.

Figure 1

Primes used in the implicit measure. Movie character primes (A) and baseline primes (B).

A



B



4.3.4. Emotional responses

To assess positive (10 items) and negative (10 items) affect, participants completed the French-validated version (Gaudreau, Sanchez, & Blondin, 2006) of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) using 7-point scales. The PANAS can be used to measure affective state or affective trait depending on the research question. In the present study, the state measure was used, that is, participants were asked to rate their affective state at the moment of completing the measure. Also, using the 9-point scales of the Self-assessment Manikin, participants rated the level of valence (1 = negative; 9 = positive) and arousal (1 = calm; 9 = aroused) of their affective state immediately after each clip (Bradley & Lang, 1994). In the present study, Cronbach's coefficients of the PANAS were as follows: .88 (positive affect scale for both clip) and .92 (negative affect scale for both clips). In the French validation study (Gaudreau, Sanchez, & Blondin, 2006), Cronbach's coefficients were as follows: Sample 1 = .90; Sample 2 = .91 (positive affect scale); Sample 1 = .80; Sample 2 = .84 (negative affect scale). It is not possible to compute indices of internal

consistency for the Self-assessment Manikin because this scale consists of two single items (i.e., valence and arousal).

4.4. Statistical analyses

The analyses collapsed across the order of the implicit and explicit measures. To test the effects of BP features on spontaneous and deliberate evaluations, indices of spontaneous and deliberate evaluation were separately submitted to a 2 (Evaluation Dimension: positivity vs. negativity) \times 2 (Clip Valence: positive vs. negative) repeated measures general linear model (GLM), with BP Features (centered) as a covariate and Clip Order (positive clip first vs. negative clip first) as a between participants factor. Table 3 presents the results of these analyses. To investigate the effects of the film clips on emotions (manipulation check), we ran a separate GLM on each of the four emotion variables: (1) positive emotions index of the PANAS; (2) negative emotions index of the PANAS; (3) emotional valence item of the Self-assessment Manikin and (4) emotional arousal item of the Self-assessment Manikin. The predictors were those described above and BP features were entered as a covariate. Table 2 presents the results of these analyses. Finally, to test more thoroughly the hypothesis that BP features are characterized by unstable evaluations of others, we computed separate instability indices for (a) spontaneous evaluations and (b) deliberate evaluations. First, for each clip, the negative evaluation score was subtracted from the positive evaluation score, with higher scores reflecting relatively greater positivity than negativity. Then, the score obtained on the negative clip was subtracted from the score obtained on the positive clip, reflecting discrepancies between evaluations following the positive and negative clips, with higher scores reflecting greater discrepancies. A score above zero indicates that the individual had relatively more positive evaluations following the positive clip compared to the negative clip. A score below zero indicates that the individual had relatively more positive evaluations following the negative clip compared to the positive clip.

5. Results

5.1. Emotions and manipulation check.

As a manipulation check, a GLM was performed on the Self-assessment Manikin. The analysis revealed a main effect of Clip Valence on the score of emotional valence: Participants reported more negative emotional reactions after the negative clip ($M = 3.927$, $SD = 1.640$) than after the positive clip ($M = 6.730$, $SD = 1.753$). A main effect of Clip Valence was also observed on the score of emotional arousal: participants reported stronger emotional arousal after the negative clip ($M = 5.029$, $SD = 1.973$) than after the positive clip ($M = 4.059$, $SD = 1.964$).

The GLM on the PANAS revealed a main effect of Evaluative Dimension, that is, participants reported more positive emotions, $M = 2.646$, $SD = .506$, and less negative emotions on the PANAS, $M = 1.684$, $SD = .550$. In addition, a significant two-way interaction was obtained between Clip Valence and Evaluative Dimension. Specifically, participants reported more positive emotions after the positive clip, $M = 2.768$, $SD = .5843$, relative to the negative clip, $M = 1.450$, $SD = .501$. The GLM also revealed a significant two-way interaction between Evaluative Dimension and BP Features. Specifically, BP features were positively correlated with reports of negative emotions, $r(202) = .391$, $p < .001$, but not with ratings of positive emotions, $r(202) = -.035$, $p = .624$. Simple slopes analyses of participants with high (1SD above) vs. low (1SD below) BP features showed higher reports of positive than negative emotions in both groups, but participants with high BP features reported relatively more negative emotions than participants with low BP features (Figure 2D shows the results). Finally, main effects of Clip Valence and BP features were observed, as well as significant two-way interactions between Clip Valence and BP features, and Clip Valence and Clip Order. No other effects in the GLM analyses on the Self-assessment Manikin and on the PANAS reached statistical significance. Table 2 presents the results.

5.2. Spontaneous evaluations

The GLM predicting spontaneous evaluations revealed a significant main effect of Evaluation Dimension, with higher scores on the index reflecting positive spontaneous evaluations, $M = 7.933$, $SD = 49.113$, relative to the index reflecting negative spontaneous evaluations, $M = -2.900$, $SD = 51.880$. The GLM also showed a main effect of BP features. The main effect of Clip Valence was not statistically significant. The main effect of Clip Order revealed a non-significant trend ($p = .076$). In addition, a significant two-way interaction was

observed between Clip Valence and Clip Order. Non-significant trends were observed for the two-way interaction between Clip Valence and BP features ($p = .063$) and for the three-way interaction between Evaluation Dimension, Clip Valence, and BP Features ($p = .051$). Other interactions were not significant, $F_s(1, 199) < 2.700$, $p_s > .100$. Table 3 presents the results.

Given that three-way interaction between Evaluation Dimension, Clip Valence, and BP Features was close to reach statistical significance ($p = .051$), this interaction was further inspected. Correlations were performed among BP features and the four scores of spontaneous evaluations, that is, 2 Evaluative Dimension (positive vs. negative) \times 2 Clip Valence (positive vs. negative). BP features showed a significant positive correlation with negative spontaneous evaluations after the negative clip, $r(203) = .239$, $p = .001$, whereas the correlations between BP features and the other three indices were not significant, $-.003 \leq r_s \leq .043$, $.543 \leq p_s \leq .964$.

As for the instability index, BP features showed a significant positive correlation with the score reflecting instability of spontaneous evaluations, $r(203) = .142$, $p = .044$, reflecting greater discrepancies between spontaneous evaluations following the positive and negative clips as a function of BP features. Partial correlations showed that this correlation remained significant when controlling for the order of the film clips, $r(199) = .138$, $p = .050$.

5.3. *Deliberate evaluations*

With respect to deliberate evaluations, the GLM revealed a significant main effect of Evaluation Dimension, with higher scores on the index reflecting positive deliberate evaluations, $M = 10.591$, $SD = 24.504$, relative to the index reflecting negative deliberate evaluations, $M = -4.438$, $SD = 25.875$. The main effects of Clip Valence and Clip Order were not significant (Table 3 shows the results). In addition, significant two-way interactions were observed between Evaluative Dimension and Clip Valence and between Evaluative Dimension and Clip Order, qualified by a significant three-way interaction. Specifically, participants reported more positive and less negative deliberate evaluations after the positive clip and this effect was strongest among participants who watched the positive clip first; in contrast, participants reported relatively more negative and less positive deliberate evaluations after the negative clip and this effect was similar across both clip orders (Table 4 shows the results). With respect to BP features, the three-way interaction between Evaluation Dimension, Clip

Valence and BP Features showed a non-significant trend ($p = .059$), qualified by a significant four-way interaction with Clip Order.

To specify the nature of the obtained four-way interaction, the four scores of deliberate evaluations (2 Clip Valence X 2 Evaluative Dimension) were separately submitted to four ANCOVAs, with BP features (centered covariate) and Clip Order (dummy-coded: 0 = negative clip first, 1 = positive clip first) and their centered interaction term. With respect to the evaluations reported after the negative clip, the ANCOVAs revealed that the interaction between Clip Order and BP Features was not significant on the positive index, $F(1, 199) = .056, p = .814, sr^2 = .000$, and on the negative index, $F(1, 199) = .135, p = .714, sr^2 = .001$; the main effects of BP features and Clip Order were not significant either, $ts(1, 199) < 1.000, p > .400$. Moreover, there was no significant difference between positive and negative evaluations after the negative clip, $F(1, 203) = 1.542, p = .125$. With respect to the evaluations reported after the positive clip, the ANCOVAs revealed significant two-way interactions between Clip Order and BP Features on the positive index, $F(1, 199) = 10.238, p = .002, sr^2 = .049$, and on the negative index, $F(1, 199) = 5.946, p = .016, sr^2 = .029$. Simple slopes analyses were performed to specify these interactions. BP features were associated with reporting relatively less positive (Figure 2B) and relatively more negative (Figure 2C) deliberate evaluations of the character after the positive clip, but only among participants who watched the positive clip first (positive evaluations: $b = -.308, t(3, 199) = -3.182, p = .002$; negative evaluations: $b = .293, t(3, 199) = 2.712, p = .007$) and not for those who watched the negative clip first (positive evaluations: $b = .114, t(3, 199) = 1.271, p = .205$; negative evaluations: $b = -.066, t(3, 199) = -.658, p = .512$). Comparisons of participants with low and high BP features (i.e., 1SD below and 1SD above the mean, respectively) confirmed that this pattern was driven by participants with high BP features reporting relatively more negative evaluations, $b = .264, t(3, 199) = 2.230, p = .027$, and relatively less positive evaluations, $b = -.308, t(3, 199) = -2.832, p = .005$, than participants with low BP features after watching the positive clip first. Hence, consistent with the negativity bias hypothesis, BP features involved relatively more negative (and less positive) deliberate evaluations of the film character and this effect emerged only after watching the positive clip first. Further inspection of this interaction suggests that these effects were primarily driven by changes in participants with low BP features, who showed less negativity and more positivity after watching the positive clip first relative to after

watching the negative clip first, whereas participants with high BP features showed little change (see Figures 2B-2C).

As for the instability index, there was a non-significant trend toward greater discrepancies between deliberate evaluations following the positive and negative clips as a function of BP features, $r(203) = -.127$, $p = .071$. Partial correlations showed that this negative correlation remained a non-significant trend when controlling for the order of the film clips, $r(199) = -.119$, $p = .092$.

Given the amount of time between the initial screening of borderline features and the experiment, the main analyses were performed again while controlling for the time interval. The effects of BP features remained unchanged for both measures (i.e., implicit and explicit) and for all three indices of evaluations (i.e., positive, negative and unstable evaluations). There was one exception: For the instability indices, the correlations between BP features and the score reflecting instability of spontaneous evaluations were no longer significant. Instead, non-significant trends were observed when controlling for the effect of time interval, $r(199) = .124$, $p = .079$, and when simultaneously controlling for the effects of clip order and time interval, $r(198) = .123$, $p = .082$.

Table 2

Summary of GLMs on emotion variables: PANAS and Self-Assessment Manikin.

Predictor	PANAS ^a		Self-Assessment Manikin - Valence ^b		Self-Assessment Manikin - Arousal ^b	
	<i>F</i>	<i>sr</i> ²	<i>F</i>	<i>sr</i> ²	<i>F</i>	<i>sr</i> ²
Clip Valence	25.412 ***	.113	318.957 ***	.615	38.458 ***	.161
Evaluation Dimension	439.260 ***	.688	-	-	-	-
Clip Order	.242	.001	.336	.002	.297	.001
BPF	12.075 **	.057	.493	.002	1.973	.010
Evaluation Dimension × BPF	24.906 ***	.111	-	-	-	-
Evaluation Dimension × Clip Order	.091	.000	-	-	-	-
Clip Valence × BPF	7.720**	.037	.082	.000	.272	.001
Clip Valence × Clip Order	12.250 **	.058	.010	.000	3.381	.017 [†]
Clip Valence × Evaluation Dimension	131.067 ***	.397	-	-	-	-
Evaluation Dimension × Clip Order × BPF	.599	.003	-	-	-	-
Clip Valence × Clip Order × BPF	.580	.003	.469	.002	.572	.003
Clip Valence × Evaluation Dimension × BPF	.528	.003	-	-	-	-
Clip Valence × Evaluation Dimension × Clip Order	2.416	.012	-	-	-	-
Clip Valence × Evaluation Dimension × Clip Order × BPF	2.000	.010	-	-	-	-

Note: There are no results regarding Evaluation Dimension for the Self-Assessment Manikin because this measure does not include separate items for the positive and negative valence.

Degrees of freedom: ^a 199 ^b 200

[†] *p* < .10. * *p* < .05. ** *p* < .01. *** *p* < .001.

Table 3

Summary of GLMs on spontaneous and deliberate evaluations.

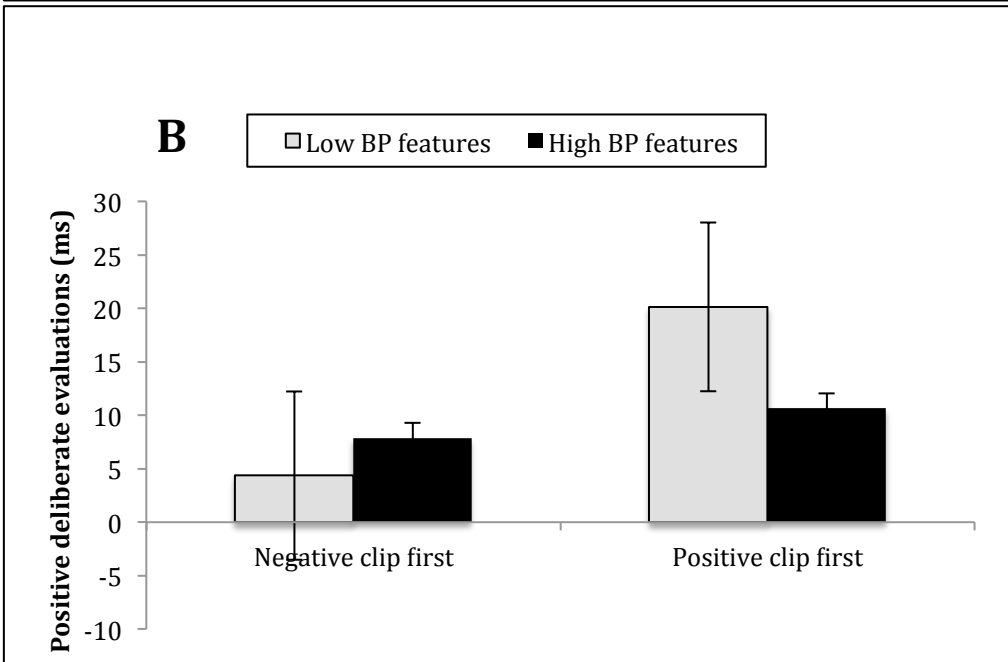
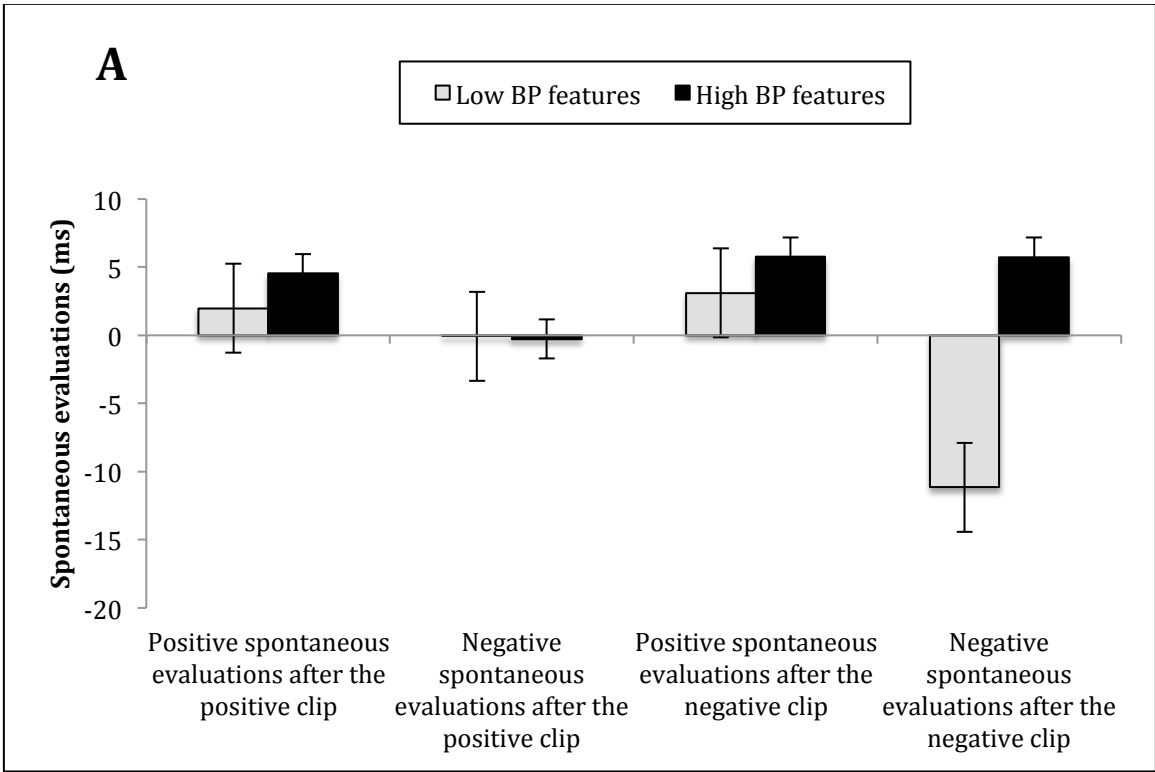
Predictor	Spontaneous evaluations		Deliberate evaluations	
	<i>F</i>	<i>sr</i> ²	<i>F</i>	<i>sr</i> ²
Clip Valence	.123	.001	1.269	.006
Evaluation Dimension	4.151*	.020	18.596***	.085
Clip Order	3.190 [†]	.016	.939	.005
BPF	4.671*	.023	.649	.003
Evaluation Dimension × BPF	1.492	.007	.141	.001
Evaluation Dimension × Clip Order	.040	.000	9.133**	.044
Clip Valence × BPF	3.487 [†]	.017	.001	.000
Clip Valence × Clip Order	5.290*	.026	1.928	.010
Clip Valence × Evaluation Dimension	.814	.004	62.776***	.240
Evaluation Dimension × Clip Order × BPF	.465	.002	2.391	.012
Clip Valence × Clip Order × BPF	2.607	.013	.422	.002
Clip Valence × Evaluation Dimension × BPF	3.847 [†]	.019	3.611 [†]	.018
Clip Valence × Evaluation Dimension × Clip Order	.189	.001	12.574***	.059
Clip Valence × Evaluation Dimension × Clip Order × BPF	.102	.001	5.977*	.029

Degrees of freedom: 199

[†] *p* < .10. * *p* < .05. ** *p* < .01. *** *p* < .001.**Table 4**

Contrasts between positive and negative indices of deliberate evaluations

Contrasts	Positive deliberate evaluations				Negative deliberate evaluations			
	<i>M</i> _{Group} (<i>SD</i>)	<i>M</i> _{ContrastGroup} (<i>SD</i>)	<i>F</i> or <i>t</i>	<i>p</i>	<i>M</i> _{Group} (<i>SD</i>)	<i>M</i> _{ContrastGroup} (<i>SD</i>)	<i>F</i> or <i>t</i>	<i>p</i>
Positive clip first vs. positive clip second	15.623 (15.425)	6.198 (13.835)	F = 20.987	.000	-12.922 (17.843)	-2.084 (14.441)	F = 22.600	.000
Positive clip first vs. negative clip second	15.623 (15.425)	-.177 (16.959)	<i>t</i> = 7.679	.000	-12.922 (17.843)	3.010 (16.197)	<i>t</i> = - 7.584	.000
Positive clip first vs. negative clip first	15.623 (15.425)	-.5099 (16.062)	N/A	N/A	-12.922 (17.843)	3.173 (16.201)	N/A	N/A



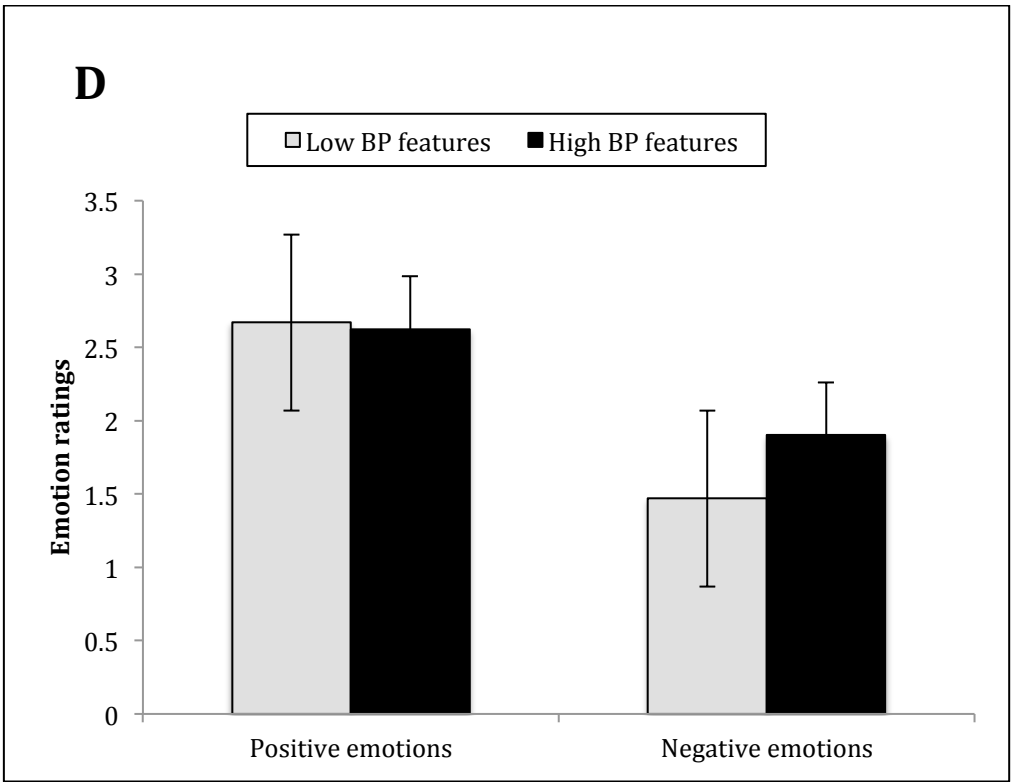
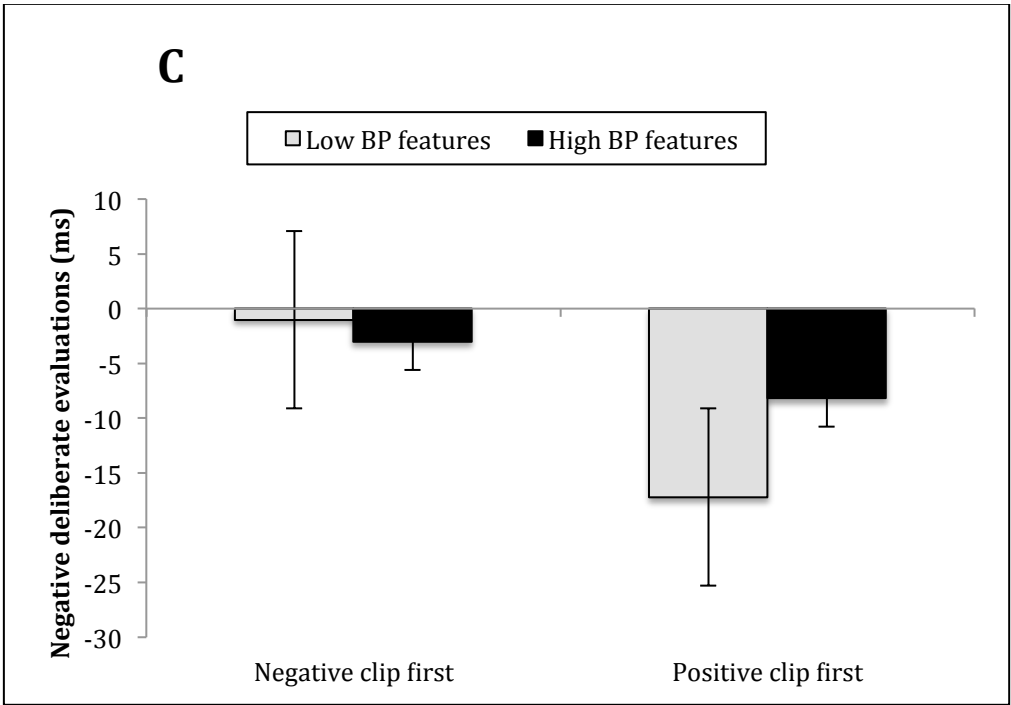


Figure 2. Figure 2A: Negative spontaneous evaluations after the negative clip as a function of BP features (continuous); Figure 2B: Positive deliberate evaluations after the positive clip as a function of BP features (continuous) and Evaluative Dimension (positive vs. negative); Figure 2C:

Negative deliberate evaluations after the positive clip as a function of BP features (continuous) and Evaluative Dimension (positive vs. negative) Figure 2D: PANAS reports of positive and negative emotions aggregated across Clip Valence (positive vs. negative) as a function of BP features (continuous).

Note: Error bars represent standard errors.

6. Discussion

6.1. Summary of findings

Clinical observations describe evaluations of others as negative and unstable in BPD (e.g., Beck et al., 2004; Kernberg, 1986) and empirical research has obtained self-report, behavioral, neurobiological and physiological evidence of negativity biases in different aspects of social cognition in BPD. However, no study, to the best of our knowledge, had yet utilized an implicit measure or combined explicit and implicit measures to examine the interplay between spontaneous and deliberate evaluations of others in relation to borderline personality (BP) features. Based on these clinical and empirical observations, this study combined implicit and explicit measures to test the hypotheses that BP features are characterized by a negativity bias and instability in spontaneous and deliberate evaluations of another person (i.e., a film character).

Results showed that participants had overall positive evaluations of the film character both at the spontaneous and deliberate levels of evaluations. The results partially supported the instability hypothesis: BP features involved greater switches from one valence to another between the two film clips. However, this finding was not replicated across analyses. We also found evidence for negativity biases for both spontaneous and deliberate evaluations. For spontaneous evaluations, BP features were related with relatively more negative spontaneous evaluations of the film character and this effect emerged only after participants watched the negative clip (i.e., conjugal dispute). After the positive clip, participants showed a pattern of positive spontaneous evaluations and BP features did not moderate this effect. These effects appeared to be driven primarily by evaluative changes in participants with low BP features, whereas participants with high BP features appeared to show less changes in their spontaneous

evaluations of the film character in response to film clips. Results obtained on the explicit measure were also consistent with a negativity bias, but contrary to the implicit measure, the negativity bias emerged after the positive clip rather than after the negative clip. Specifically, among participants who watched the positive clip first, participants with low BP features displayed a pattern of increased positivity and reduced negativity, which was not observed in participants with high BP features. After the negative clip, both participants with high and low BP features displayed a pattern of ambivalence involving equally low levels of positivity and negativity in their deliberate evaluations. Together, the results obtained on the explicit and implicit measures partially support the view that BP features involves more negative and unstable evaluations of others (e.g., Beck et al., 2004; Kernberg, 1986) and converge with empirical evidence that BPD involves negative (e.g., Domes et al., 2009; Sieswerda et al., 2013) and unstable evaluations of others (Beeney et al., 2016; Coifman et al., 2012). Moreover, the obtained pattern of negativity biases is accompanied by the finding that participants with high BP features reported more negative emotions in the PANAS relative to participants with low BP features.

6.2. Implications

According to the APE model, spontaneous evaluations tend to operate at earlier processing stages and under more automatic conditions, whereas deliberate evaluations tend to operate at later processing stages and under conditions that allow for more control to be exerted (Gawronski & Bodenhausen, 2011). With this in mind, the results suggest that unstable evaluations of others in individuals with BP features may be rooted at least in part in affective gut reactions. With respect to negativity biases, the reduced capacity of individuals with high BP features to modulate their spontaneous and deliberate evaluations positively (or less negatively) appears not only to be rooted in explicit and controlled beliefs about others but also in affective gut reactions to others resulting from the activation of mental associations in memory. Specifically, considering the different pattern of results for the explicit and the implicit measures, the reduced positivity of individuals with high BP features after the negative clip appeared to be shaped at relatively early processing stages that do not require intention and effort, whereas their relative negativity (reduced positivity) after the positive clip appeared to be shaped at relatively late processing stages that are more largely guided by

effortful and intentional thinking processes. In contrast, individuals with low BP features seemed to modulate their spontaneous evaluations more positively after the negative clip. This increased positivity after a negative situation may be a healthy mechanism serving to protect interpersonal interactions in the face of threat. In this vein, positive spontaneous evaluations have been shown to foster feelings and behavioral expressions of trust (i.e., approaching the partner) (Murray et al., 2011). Our results suggest that such relationship-protective processes may be compromised in individuals with BP features, but more research is needed to verify potential impacts of reduced positivity on interpersonal processes and behaviors of individuals with BP features and BPD.

6.3. Limits and conclusion

The amount of time between the initial screening of borderline features and the experiment varied between participants. Statistical analyses were performed to control for the time interval and the results remained largely unchanged. However, it remains possible that the time interval influenced our results in a non-systematic (i.e., increasing the error in some participants more than in others). The statistical control unfortunately cannot eliminate error when it is non-systematic. In addition, it is important to nuance the conclusions pertaining to spontaneous evaluations considering that it was based on a non-significant trend in the three-way interaction between Evaluation Dimension, Clip Valence, and BP Features. The findings regarding instability should also be interpreted as tentative given that they did not replicate across analyses. This is the first study to specifically examine spontaneous evaluations of other people in relation to BP features. Therefore, more research is needed to assess the replicability of our findings and to achieve accurate estimates of the obtained effects across multiple experiments. Some of the effects obtained in this study were small and/or non-significant trends. This may be due to (1) a genuinely small relationship between BP features and evaluative responses, (2) the potential influence of other factors and traits that could moderate the relationship between BP features and evaluative responses, (3) the use of a student sample (in opposition to a clinical sample) or to (3) ineffective manipulations. In addition, film clips may not adequately capture relationship instability that usually unfolds over a period of time (for a review of affective instability in BPD, see Santangelo, Bohus, & Ebner-Priemer, 2014). However, there is also evidence that individuals with BPD can experience relatively rapid

fluctuations from a mood state with positive valence to a mood state with negative valence at intervals varying between 10 to 20 minutes (Ebner-Priemer et al., 2007). Hence, it is possible that instability in BPD occurs both at shorter and longer intervals.

To address the aforementioned limits, future studies may assess patterns of spontaneous evaluations and emotions in the natural environment of individuals with BPD. One interesting avenue of research would be to examine how spontaneous evaluations influence behavior in borderline personality using a longitudinal design and to analyze the prospective relationships between spontaneous/deliberate evaluations and spontaneous/deliberate behaviors. In fact, there is evidence that spontaneous evaluations tend to be superior in predicting behaviors that are difficult to control (e.g., non-verbal responses) and behaviors under conditions of limited cognitive resources (e.g., ego depletion), whereas deliberate evaluations tend to be superior in predicting controllable behaviors (e.g., verbal behavior) and behaviors under conditions of sufficient cognitive resources (Perugini, 2005).

In summary, this is the first study to examine the interplay between spontaneous and deliberate evaluations of others in relation to BP features. By doing so, the present study yielded evidence that impression formation in individuals with BP features may be characterized by instability and by negativity biases and that processes operating at relatively early stages may contribute to these evaluations. However, this preliminary investigation has several limits that should be noted when interpreting these results and future research should examine the replicability of the findings. Nevertheless, our provisional findings suggest that future studies may gain deeper insights into the mechanisms underlying interpersonal difficulties in borderline personality by examining the interplay between spontaneous and deliberate evaluations in borderline personality.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Arntz, A., & Haaf, J. T. (2012). Social cognition in borderline personality disorder: Evidence for dichotomous thinking but no evidence for less complex attributions. *Behaviour research and therapy*, *50*(11), 707-718. doi: 10.1016/j.brat.2012.07.002
- Arntz, A., & Veen, G. (2001). Evaluations of others by borderline patients. *Journal of Nervous and Mental Disease*, *189*(8), 513-521. doi: 10.1097/00005053-200108000-00004

- Bagge, C., Nickell, A., Stepp, S., Durrett, C., Jackson, K., & Trull, T. J. (2004). Borderline personality disorder features predict negative outcomes 2 years later. *Journal of Abnormal Psychology, 113*(2), 279-288. doi: 10.1037/0021-843x.113.2.279
- Barnow, S., Stopsack, M., Grabe, H. J., Meinke, C., Spitzer, C., Kronmueller, K., & Sieswerda, S. (2009). Interpersonal evaluation bias in borderline personality disorder. *Behaviour Research and Therapy, 47*(5), 359-365. doi: 10.1016/j.brat.2009.02.003
- Beck, A. T., Freeman, A., & Davis, D. D. (2004). *Cognitive therapy of personality disorders (2nd ed.)*. New York, NY: Guilford Press; US.
- Beck, A. T., & Freeman, A. M. (1990). *Cognitive therapy of personality disorders*. New York, NY: Guilford Press; US.
- Beeney, J. E., Hallquist, M. N., Ellison, W. D., & Levy, K. N. (2016). Self-other disturbance in borderline personality disorder: Neural, self-report, and performance-based evidence. *Personality disorders, 7*(1), 28-39. doi: 10.1037/per0000127
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59. doi: 10.1016/0005-7916(94)90063-9
- Coifman, K. G., Berenson, K. R., Rafaeli, E., & Downey, G. (2012). From negative to positive and back again: polarized affective and relational experience in borderline personality disorder. *Journal of Abnormal Psychology, 121*(3), 668-679. doi: 10.1037/a0028502
- Domes, G., Schulze, L., & Herpertz, S. C. (2009). Emotion recognition in borderline personality disorder - a review of the literature. *Journal of Personality Disorders, 23*(1), 6-19.
- Dreessen, L., & Arntz, A. (1995). *The personality disorder beliefs questionnaire*. Maastricht.
- Ebner-Priemer, U. W., Kuo, J., Kleindienst, N., Welch, S. S., Reisch, T., Reinhard, I., . . . Bohus, M. (2007). State affective instability in borderline personality disorder assessed by ambulatory monitoring. *Psychological Medicine, 37*(7), 961-970. doi: 10.1017/s0033291706009706
- Edens, J. F., Marcus, D. K., & Ruiz, M. A. (2008). Taxometric analyses of borderline personality features in a large-scale male and female offender sample. *Journal of Abnormal Psychology, 117*(3), 705-711. doi: 10.1037/0021-843x.117.3.70.5
- Evans, D., Dalgleish, T., Dudas, R. B., Denman, C., Howard, M., & Dunn, B. D. (2015). Examining the shared and unique features of self-concept content and structure in borderline personality disorder and depression. *Cognitive Therapy and Research, 39*(5), 613-626. doi: 10.1007/s10608-015-9695-3
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*(2), 229-238. doi: 10.1037//0022-3514.50.2.229
- Gaudreau, P., Sanchez, X., & Blondin, J. P. (2006). Positive and negative affective states in a performance-related setting - Testing the factorial structure of the PANAS across two samples of French-Canadian participants. *European Journal of Psychological Assessment, 22*(4), 240-249. doi: 10.1027/1015-5759.22.4.240
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692-731. doi: <http://dx.doi.org/10.1037/0033-2909.132.5.692>

- Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. In K. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology, Vol 44* (Vol. 44, pp. 59-127).
- Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of implicit social cognition: measurement, theory, and applications*. New York, NY: Guilford Press.
- Haslam, N. (2003). Categorical versus dimensional models of mental disorder: the taxometric evidence. *Australian and New Zealand Journal of Psychiatry, 37*(6), 696-704. doi: 10.1080/j.1440-1614.2003.01258.x
- Hedrick, A. N., & Berlin, H. A. (2012). Implicit self-esteem in borderline personality and depersonalization disorder. *Frontiers in Psychology, 3*. doi: 10.3389/fpsyg.2012.00091
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*(10), 1369-1385. doi: 10.1177/0146167205275613
- Jackson, K. M., & Trull, T. J. (2001). The factor structure of the personality assessment inventory-borderline features (PAI-BOR) scale in a nonclinical sample. *Journal of Personality Disorders, 15*(6), 536-545. doi: 10.1521/pedi.15.6.536.19187
- Kernberg, O. F. (1986). *Severe personality disorders: Psychotherapeutic strategies*. New Haven, CT: Yale University Press; US.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling (3rd ed.)*. New York, NY: Guilford Press.
- Krause-Utz, A., Winter, D., Niedtfeld, I., & Schmahl, C. (2014). The latest neuroimaging findings in borderline personality disorder. *Current Psychiatry Reports, 16*(3). doi: 10.1007/s11920-014-0438-z
- Leichsenring, F., Leibling, E., Kruse, J., New, A. S., & Leweke, F. (2011). Borderline personality disorder. *Lancet, 377*(9759), 74-84. doi: 10.1016/s0140-6736(10)61422-5
- Lieb, K., Zanarini, M. C., Schmahl, C., Linehan, M. M., & Bohus, M. (2004). Borderline personality disorder. *Lancet, 364*(9432), 453-461. doi: 10.1016/s0140-6736(04)16770-6
- Matzke, B., Herpertz, S. C., Berger, C., Fleischer, M., & Domes, G. (2014). Facial reactions during emotion recognition in borderline personality disorder: a facial electromyography study. *Psychopathology, 47*(2), 101-110. doi: 10.1159/000351122
- Morey, L. C. (1991a). Inventaire de la personnalité (C. Routhier, Trans.), *Personality Assessment Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Morey, L. C. (1991b). *The personality assessment inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Murray, S. L., Pinkus, R. T., Holmes, J. G., Harris, B., Gomillion, S., Aloni, M., . . . Leder, S. (2011). Signaling when (and when not) to be cautious and self-protective: impulsive and reflective trust in close relationships. *Journal of Personality and Social Psychology, 101*(3), 485-502. doi: 10.1037/a0023233
- Payne, B. K., Burkley, M., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology, 94*(1), 16-31.
- Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology, 44*(1), 29-45.

- Renneberg, B., Herm, K., Hahn, A., Staebler, K., Lammers, C.-H., & Roepke, S. (2012). Perception of social participation in borderline personality disorder. *Clinical Psychology & Psychotherapy*, *19*(6), 473-480. doi: 10.1002/cpp.772
- Rothschild, L., Cleland, C., Haslam, N., & Zimmerman, M. (2003). A taxometric study of borderline personality disorder. *Journal of Abnormal Psychology*, *112*(4), 657-666. doi: 10.1037/0021-843x.112.4.657
- Rusch, N., Lieb, K., Gottler, I., Hermann, C., Schramm, E., Richter, H., . . . Bohus, M. (2007). Shame and implicit self-concept in women with borderline personality disorder. *American Journal of Psychiatry*, *164*(3), 500-508. doi: 10.1176/appi.ajp.164.3.500
- Santangelo, P., Bohus, M., & Ebner-Priemer, U. W. (2014). Ecological momentary assessment in borderline personality disorder: A review of recent findings and methodological challenges. *Journal of Personality Disorders*, *28*(4), 555-576.
- Sieswerda, S., Arntz, A., & Wolfis, M. (2005). Evaluations of emotional noninterpersonal situations by patients with borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *36*(3), 209-225. doi: 10.1016/j.jbtep.2005.05.004
- Sieswerda, S., Barnow, S., Verheul, R., & Arntz, A. (2013). Neither dichotomous nor split, but schema-related negative interpersonal evaluations characterize borderline patients. *Journal of Personality Disorders*, *27*(1), 36-52.
- Staebler, K., Helbing, E., Rosenbach, C., & Renneberg, B. (2011). Rejection sensitivity and borderline personality disorder. *Clinical Psychology & Psychotherapy*, *18*(4), 275-283. doi: 10.1002/cpp.705
- Teachman, B. A., Joormann, J., Steinman, S. A., & Gotlib, I. H. (2012). Automaticity in anxiety disorders and major depressive disorder. *Clinical Psychology Review*, *32*(6), 575-603. doi: 10.1016/j.cpr.2012.06.004
- Trull, T. J. (1995). Borderline personality disorder features in nonclinical young adults I: Identification and validation. *Psychological Assessment*, *7*, 33-41.
- Trull, T. J., Useda, J. D., Conforti, K., & Doan, B. T. (1997). Borderline personality disorder features in nonclinical young adults: 2. Two-year outcome. *Journal of Abnormal Psychology*, *106*(2), 307-314. doi: 10.1037/0021-843x.106.2.307
- Vater, A., Schroeder-Abe, M., Weissgerber, S., Roepke, S., & Schuetz, A. (2015). Self-concept structure and borderline personality disorder: evidence for negative compartmentalization. *Journal of Behavior Therapy and Experimental Psychiatry*, *46*, 50-58. doi: 10.1016/j.jbtep.2014.08.003
- Veen, G., & Arntz, A. (2000). Multidimensional dichotomous thinking characterizes borderline personality disorder. *Cognitive Therapy and Research*, *24*(1), 23-45. doi: 10.1023/a:1005498824175
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063-1070. doi: 10.1037/0022-3514.54.6.1063
- Zeigler-Hill, V., & Abraham, J. (2006). Borderline personality features: Instability of self-esteem and affect. *Journal of Social and Clinical Psychology*, *25*(6), 668-687. doi: 10.1521/jscp.2006.25.6.668

Article 2

Negativity bias and polarization in spontaneous and deliberate partner evaluations: The role of
borderline personality disorder features

Mongeon, F., Gawronski, B., Gagnon, J. (2016). Negativity bias and polarization in spontaneous
and deliberate partner evaluations: The role of borderline personality disorder features.

Abstract

This study examined whether borderline personality (BP) features are characterized by negative or polarized evaluations of self and partner. Participants ($N = 292$) imagined either (a) that their partner rejected them or (b) sought connection or (c) imagined no scenario (control condition). Spontaneous and deliberate evaluations were measured with a name-letter task and a self-report measure, respectively. Overall, participants showed positive evaluations of self and partner, and participants with high BP features showed negativity biases, as evidenced by relatively less positive spontaneous and deliberate partner evaluations independent of the conditions in the study, and less positive spontaneous self evaluations in the control condition only. However, the expected main effects of experimental conditions were not observed. Together, the overall pattern of results suggests two possible interpretations: (a) borderline personality may be characterized by negative evaluation biases that influence both relatively early and late processing stages within romantic relationships; (b) alternatively, given the absence of main effects of experimental conditions, the obtained pattern of negativity across experimental conditions may have artificially resulted from the use of potentially ineffective manipulations in this study.

Keywords: borderline personality; spontaneous evaluation; deliberate evaluation; negativity bias; splitting; social cognition; implicit measure; explicit measure

1. Introduction

Interpersonal dysfunctions are central to borderline personality disorder (BPD) and romantic relationships are particularly affected (Hill et al., 2008; 2011, but see Daley et al., 2000). In fact, individuals with BPD experience romantic dysfunctions characterized by intimate violence (Ross & Babcock, 2009), higher rates of breakups and lower marital satisfaction (Bouchard, Sabourin, Lussier, & Villeneuve, 2009), dysfunctional sexual behaviors (Tragesser, Lippman, Trull, & Barrett, 2008), negative sexual experiences (Bouchard, Godbout, & Sabourin, 2009; Hurlbert, Apt, & White, 1992), increased emotional reactivity and quarrelsomeness (Sadikaj et al., 2013), rejection sensitivity and attachment insecurity (Boldero et al., 2009) and greater uncertainty about one's identity within the relationship (Valentiner, Hiraoka, & Skowronski, 2014). Although more research has examined clinical samples, nonclinical individuals with borderline personality (BP) features report levels of romantic and interpersonal dysfunctions similar to those of clinical samples (Daley et al., 2000; Tragesser & Benfield, 2012; Trull et al., 1997; Whisman & Schonbrun, 2009). In addition to the research examining romantic dysfunctions, there is also evidence pointing to the protective role of satisfying romantic relationships in borderline personality (Gunderson et al., 2003; Gunderson et al., 2006; Kuhlken, Robertson, Benson, & Nelson-Gray, 2014). For example, Kuhlken et al. (2014) showed that individuals with elevated BP features reported less anger and more improvement of negative affect when in a satisfying romantic relationship. Together, these findings highlight the importance of addressing romantic dysfunctions among both individuals with BPD and nonclinical individuals with BP features. Although previous research has highlighted different mechanisms that can contribute to interpersonal difficulties within different types of interpersonal situations in borderline personality, little research has examined the mechanisms specifically associated with difficulties within romantic relationships in borderline personality – rather than within interpersonal situations in general. The way individuals with BP features evaluate their partner is one mechanism that may influence their romantic functioning and that mechanism has received little attention in past research. Hence, the main goal of the present study was to fill this gap in the literature on borderline personality, focusing on the mechanism of *partner evaluations* and distinguishing between two kinds of evaluations, namely *spontaneous* and

deliberate partner evaluations. In doing so, we used a dimensional approach to borderline personality, where BP features represent a continuum of severity (Edens et al., 2008; Haslam, 2003; Rothschild et al., 2003)

2. Spontaneous and deliberate evaluations

According to the associative–propositional evaluation (APE) model, spontaneous evaluations represent affective gut reactions resulting from the activation of mental associations in memory. In contrast, deliberate evaluations represent the behavioral outcome of propositional processes involving the validation of activated information on the basis of logical consistency (for a discussion and a review of the evidence supporting the model, see Gawronski & Bodenhausen, 2006, 2011). The main implication of the theory is the following: whereas deliberate evaluations depend on subjective judgments of truth or falsity, spontaneous evaluations are not directly influenced by subjective truth or falsity. In addition, the two types of evaluations operate under different conditions: compared to deliberate evaluations, spontaneous evaluations have been proposed to operate under conditions of unintentionality, efficiency, and uncontrollability (Bargh, 1994; for a critical discussion of these features, see Gawronski & Bodenhausen, 2011). Based on these differences, different measurements tools are used to measure each type of evaluations. Specifically, implicit measures are usually performance-based instruments and are designed to capture spontaneous evaluations; in implicit measures, “the impact of the to-be-measured psychological attribute on participants’ responses is unintentional, resource-independent, unconscious, or uncontrollable” (Gawronski & De Houwer, 2014, p. 284). In contrast, explicit measures are self-report instruments and are used to capture deliberate evaluations. Implicit and explicit measures often diverge and the APE model provides explanations of such dissociations in terms of the respective processes and operating principles guiding the two kinds of evaluations. With respect to partner evaluations, past research has illustrated the utility of combining implicit and explicit measures (e.g., LeBel & Campbell, 2009, 2013; Murray, Gomillion, Holmes, Harris, & Lamarche, 2013; Murray et al., 2011); for instance, LeBel and Campbell (2013) showed that deliberate and spontaneous partner evaluations both contribute to daily relationship quality and positive relationship behaviors enacted toward one’s partner. However, no study, to the

best of our knowledge, has used an implicit measure to examine spontaneous partner evaluations in relation to borderline personality. Based on these considerations, the present study utilized an implicit and an explicit measure to examine the interplay between spontaneous and deliberate partner evaluations in borderline personality.

3. Negative and splitted evaluations in borderline personality

Past clinical observations and empirical research suggest that evaluations of BPD individuals may be best characterized as negative and polarized. The term *negativity bias* refers to the tendency to evaluate stimuli through a negative lens and *polarity* refers to a difficulty integrating disparate feelings, leading one to compartmentalize experiences and feelings in all-good or all-bad categories and to alternate between these polarities. In line with these notions, DSM-5 notes that BPD involves negative evaluation of others (“perceptions of others selectively biased toward negative attributes or vulnerabilities”) and self (“feelings of inferior self-worth”), as well as polarized evaluations of others (criterion 2: “alternating between extremes of idealization and devaluation”) and self (criterion 3: “unstable self-image”) (American Psychiatric Association, 2013, pp. 663, 766-767). Similarly, the cognitive theory of BPD hypothesizes that dysfunctions of BPD individuals are at least partly caused by maladaptive cognitive schemas (basic cognitive structures in memory) that bias the evaluation and interpretation of environmental stimuli, including the basic assumption that the world is dangerous and malevolent and that the self is powerless, vulnerable and unacceptable (A. T. Beck & Freeman, 1990; Aaron T. Beck et al., 2004). Psychodynamic theories also highlight the phenomenon of negative or malevolent evaluations (Kernberg, 1986; Westen, 1991). Moreover, the cognitive notion of *dichotomous thinking* (A. T. Beck & Freeman, 1990; Aaron T. Beck et al., 2004) and the psychodynamic concept of *splitting* (Kernberg, 1986) both capture the tendency of individuals with BPD to experience extreme polarity or to alternate between extreme experiences (e.g., viewing the self or partner as all good or all bad).

Little research has been able to support the claim that borderline personality involves negativity biases within romantic relationships. One study showed that BP features were related with reporting more negative reactions to partner-initiated negative and positive experiences (Bhatia et al., 2013) and another study found that women with BPD involved in

romantic relationships reported more negative self-reported sexual experiences than controls (Bouchard, Godbout, et al., 2009). In contrast, a vast amount of research has examined negativity biases in BPD within different social domains – rather than within romantic relationships specifically. Studies that utilized performance-based cognitive tasks brought evidence supporting the view that individuals with BPD have difficulties disengaging their attention from negatively valenced stimuli (Arntz, Appels, & Sieswerda, 2000; Sieswerda, Arntz, & Kindt, 2007; Sieswerda, Arntz, Mertens, et al., 2007; Wingenfeld, Mensebach, et al., 2009; Witthoft, Borgmann, White, & Dyer, 2015); however, there are also numerous other studies that used similar performance-based cognitive tasks and found equivocal findings or more general inhibition deficits rather than inhibition deficits for negative stimuli specifically (Domes et al., 2006; Hagenhoff et al., 2013; Krause-Utz et al., 2012; Portella et al., 2011; von Ceumern-Lindenstjerna et al., 2010; Wingenfeld, Rullkoetter, et al., 2009; Winter, Herbert, et al., 2015; Winter, Krause-Utz, et al., 2015). Previous BPD research also examined negativity biases with self-report tools. Arntz and colleagues conducted a series of studies in the laboratory and found a stronger tendency in BPD to describe others in a more negative manner (Arntz & Veen, 2001; Sieswerda, Arntz, & Wolfis, 2005; Sieswerda, Barnow, Verheul, & Arntz, 2013) but they also found disconfirming evidence (Arntz & Haaf, 2012; Veen & Arntz, 2000). Barnow et al. (2009) showed that BPD individuals formed more negative evaluations of neutral interpersonal stimuli. In addition, previous work found that perceptions of rejection contribute to quarrelsomeness, negative affect (Sadikaj et al., 2013) and anger (Berenson et al., 2011) in individuals with BPD relative to control participants. However, given that these studies examined evaluations of others without specifically measuring partner evaluations (i.e., how positively or negatively one perceives its partner), they constitute only indirect evidence that borderline personality is related with negative partner evaluations. Nevertheless, it is possible that the patterns of negativity biases found in different social domains also affect how those with BP features appraise their partners. Other indirect evidence of negative partner evaluations in borderline personality comes from studies showing that features of borderline personality, such as low self-esteem, insecure attachment, negative affect intensity and variability and deficits in emotion regulation strategies, can contribute to negative or polarized partner evaluations (Cavallo et al., 2012; Graham & Clark, 2006; Klein et al., 2016; Murray et al., 2002; Sadikaj et al., 2015). There is also evidence for polarized partner evaluations in

BPD. Notably, Coifman et al. (2012) examined polarity of affective and relational experiences, defined as strong positive experiences with low negative experiences or strong negative experiences with low positive experiences. Polarity in reports of affective and relational experiences was greater in BPD participants relative to healthy controls and it increased during heightened interpersonal stress. In addition, increased polarity in reports of affective reactions (in low-stress contexts) and relational experiences (in high-stress contexts) predicted increased reports of impulsive behaviors (e.g., self-injury, substance abuse). Again, this study did not focus specifically on partner evaluations. Together, despite indirect evidence suggesting that borderline personality may involve negative and polarized partner evaluations, there is no direct evidence supporting these claims.

To summarize, notwithstanding the contributions of past research on borderline personality and romantic dysfunctions, there is no study directly examining how borderline personality influences partner evaluations. Hence, the main goal of the present study was to fill this gap in the BPD literature by examining the role of borderline personality (BP) features on the way individuals evaluate their partner; in addition, because spontaneous and deliberate partner evaluations both play a role in relationships' functioning, we focus particularly on the interplay between spontaneous and deliberate partner evaluations using implicit and explicit measures. A secondary aim was to replicate past research showing negative self evaluations in borderline personality (e.g., Abdul-Hamid, Denman, & Dudas, 2014; Beeney et al., 2016; Evans et al., 2015; Hawes, Helyer, Herlianto, & Willing, 2013; Hedrick & Berlin, 2012; Kopala-Sibley, Zuroff, Russell, Moskowitz, & Paris, 2012; Rüscher et al., 2007; Rusch et al., 2007; Rusch et al., 2011; Vater et al., 2015).

4. The present study

The name-letter task (NLT, Nuttin, 1985) was used to measure spontaneous partner and self evaluations. The NLT indicates the relative preference for partner's and self's initials. Also, a self-report measure tapping positive and negative reactions to the partner was used to measure deliberate partner evaluations. Participants were randomly assigned to one of three conditions: One third of the participants were asked to imagine a hypothetical situation where their partner rejects them (negative condition); another third were asked to imagine a hypothetical situation

where their partner seeks connection (positive condition); the last third completed the measures without being asked to imagine a hypothetical situation (control condition).

Based on our theoretical considerations and on the evidence available so far, we expected that (a) on average, people would possess positive evaluations of themselves and of their partner, that (b) this positivity effect would be increased by the closeness manipulation and reduced by the rejection manipulation relative to the control condition and that (c) BP features would be related with less positive evaluations (negativity bias), such that participants with high (vs. low) BP features would show less preference for self and partner initials on the implicit measure and less positive self-reported reactions to their partner on the explicit measure. However, because previous BPD research led to inconsistent findings as to the situations or stimuli that may potentially trigger negativity biases, we cannot go so far as to specify in which conditions of the present study (rejection, closeness, control) the negativity bias should emerge. In contrast with the notion of negativity bias, the notion of polarity assumes that BPD may not only involve more negative evaluations, but also more positive evaluations in some situations. This idea is in line with two of the current diagnostic criteria for BPD (Criterion 2: “unstable and intense interpersonal relationships characterized by alternating between extremes of idealization and devaluation”; Criterion 3: “unstable self-image or sense of self”; American Psychiatric Association, 2013, p. 663). Based on these considerations, we contrasted the negativity bias hypothesis with the polarity hypothesis. (d) According to this hypothesis, we expected participants with high BP features would display relatively greater negativity in the rejection condition (devaluation) and increased positivity in the closeness condition (idealization) relative to the control condition on all three kinds of evaluations (i.e., spontaneous partner and self evaluations and deliberate partner evaluations). Based on the dimensional approach to borderline personality described earlier, we tested our hypotheses in a non-clinical sample of college students with varying degrees of borderline personality (BP) features.

5. Method

5.1. Sample and procedure

Two hundred ninety-two adults (155 women, 136 men and 1 participant who preferred not to say) who were in a romantic relationship for at least 6 months participated in exchange for monetary compensation. Participants were recruited through Mechanical Turk and completed all measures online (MTurk; see Buhrmester, Kwang, & Gosling, 2011). Participants were told that the study examined personality and relationships. The mean age was 32.37 (SD = 10.279, range = 18-68). The distribution of BP features on the Personality Diagnostic Questionnaire Fourth Edition-BPD Scale (PDQ4-BPD; Hyler, 1994) was normal (skew = .831; kurtosis = .063), that is, absolute values of skew < 3 and kurtosis < 10 per guidelines provided by Kline (2011). The PDQ4-BPD mean ($M = 2.24$, $SD = 2.03$) was smaller than in previous studies (e.g., Fossati et al., 2004; Gardner & Qualter, 2009; Klonsky, 2008). The reliability was acceptable ($\alpha = .70$).

Participants were directed to a secure website to first complete demographics and the questionnaire on BP features. Table 1 shows the socio-demographic information. They were then randomly assigned to either three conditions (rejection scenario, closeness scenario or control condition). Participants assigned to the rejection or closeness scenarios were informed that they would see a hypothetical scenario that might happen in a romantic relationship and were asked to vividly imagine the scenario as if it was real. The rejection scenario was made ambiguous based on the assumption that BP features might have a greater impact on participants' partner evaluations following an ambiguous rejection rather than an explicit one. Also, there is evidence that individuals with BPD assign more negative traits than controls to a person seen in an ambiguous context (Barnow et al., 2009). In the rejection condition ($n = 93$), participants read the following scenario: "You have breakfast with your partner and he/she tells you that he/she wants to be alone this evening and also the entire weekend." In the closeness condition ($n = 99$), participants read the following scenario: "You have breakfast with your partner and he/she tells you that he/she wants to spend this evening with you and also the entire weekend." Participants were then instructed to write about the scenario, "what exactly your partner says when he/she tells you that he/she wants to be with you this evening and the entire weekend. In addition, please describe his/her voice tone, gesture, and facial expression." Participants were also instructed to write about their emotional response, "Please describe what you are feeling when your partner tells you that he/she wants to be with you this evening and the entire weekend". Then, participants completed the explicit and implicit

measures of partner evaluations. As for participants assigned to the control condition ($n = 100$), they completed the measures without being asked to imagine a hypothetical situation. The order of implicit-explicit measures was randomly counterbalanced between participants.

Table 1
Socio-demographic statistics.

	<i>M</i>	<i>SD</i>
Gender (% female)	53.1	
Age	32.37	10.28
Education (%)		
High school	37.30	
Bachelor	45.20	
Post-grad.	17.5	
Relationship duration (in years, range .17 to 69.67)	6.81	8.01
Cohabiting with partner (%)	72.60	

5.2. Measures

5.2.1. BP features

The PDQ4-BPD (Hyer, 1994) contains nine dichotomous items addressing the DSM-5 BPD criteria. Internal consistency is adequate ($K-R-20 = .71$) and there is convergence with clinical interviews and other screening questionnaires ($k = .57$ and $r > .70$; e.g., Blackburn, Donnelly, Logan, & Renwick, 2004; Hyler, Skodol, Kellman, Oldham, & Rosnick, 1990). Discriminant validity regarding antisocial features and psychopathy is supported (Salekin, Trobst, & Krioukova, 2001). The convergence of the PDQ4-BPD with three other measures of BP features also supports its validity (Gardner & Qualter, 2009; Hyler et al., 1990). The internal consistency obtained in the present sample was adequate ($K-R-20 = .70$) and similar to that observed in previous research (e.g., Hyler et al., 1990).

5.2.3. Explicit measure

As an indicator of deliberate partner evaluations, participants responded to single word items reflecting either positive or negative emotional reactions to their partners on a 9-point scale (1 = *not at all*, 9 = *very much*). 7 words described negative reactions to the partner (e.g., “my partner makes me feel...” distressed, nervous, ...) and 7 words described positive

reactions to the partner (e.g., “my partner makes me feel...” delighted, happy, ...). Table 1 shows the full list of words. We computed an index of deliberate partner evaluations by subtracting summed responses to the items reflecting negative emotions from summed responses to the items reflecting positive emotional reactions ($\alpha = .944$). Accordingly, higher scores indicate greater levels of positivity of deliberate partner evaluations.

Table 2
List of positive and negative words used in the explicit measure

Negative emotional reactions	Positive emotional reactions
Distressed	Relaxed
Nervous	Content
Upset	Happy
Irritable	Delighted
Afraid	Agreeable
Angry	Calm
Gloomy	Excited

5.2.4. *Implicit measure: Partner-NLT*

As an indicator of spontaneous partner evaluations, participants completed the NLT (Nuttin, 1985) which involves rating how esthetically pleasing each letter of the alphabet is on a 9-point scale, anchored by 1 = not esthetically pleasing at all and 9 = very esthetically pleasing (letters presented one at a time in a fixed random order). Previous research has shown that people tend to evaluate their own and partner's initials more favorably compared to baseline evaluations of those letters, which is commonly interpreted as an indicator of positive spontaneous evaluations of self and partner (LeBel & Campbell, 2009). As in past research, participants were instructed to make their judgments quickly and focus on their intuitive feelings toward each letter. To score letter ratings, we used a special instantiation of a NLT scoring algorithm that controls for baseline letter favorability and individual differences in response tendencies of letter ratings (Baccus, Baldwin, & Packer, 2004; LeBel & Gawronski, 2009). The algorithm extracts indices of relative preference for partner initials using an ipsatized scores (for more information on the algorithm, see LeBel & Gawronski, 2009). After completing the NLT, participants were asked to provide their own and partner's initials. Preference scores for partner's first and second initials were averaged to form an index of

spontaneous partner evaluations ($\alpha = .518$). Higher scores indicate greater levels of positive spontaneous partner evaluations. Using this algorithm, as originally applied to the self, we also calculated an index of spontaneous self evaluations ($\alpha = .542$; Baccus et al., 2004). Some participants might have the same initials as their partner; however, the steps to compute the index of spontaneous partner evaluations remains unchanged for those participants based on LeBel & Gawronski's guidelines (2009).

6. Results

The analyses collapsed across the order of the implicit and explicit measures. Positive overall scores were obtained on all three evaluations' indices, that is, values were statistically superior to zero for self-reported reactions to the partner, $M = 33.003$, $SD = 20.077$, $t(291) = 28.091$, $p < .001$, $d = 2.594^1$, preference for partner's initials, $M = .875$, $SD = 1.635$, $t(291) = 9.146$, $p < .001$, $d = .535$, and preference for own's initials, $M = 1.619$, $SD = 1.540$, $t(291) = 17.960$, $p < .001$, $d = 1.051$. Hence, participants showed a significant overall positive bias on all three kinds of evaluations. These results replicate prior research on partner and self evaluation (LeBel & Campbell, 2009, 2013).

To examine the main hypotheses and test the effects of BP features on spontaneous and deliberate evaluations, indices of spontaneous and deliberate evaluations were separately submitted to two linear regressions. Specifically, we entered (a) two dummy coded contrasts to capture the evaluative conditioning conditions (one compared the closeness condition, coded 1, against the control and rejection conditions, coded 0; the other compared the rejection condition, coded 1, against the control and closeness conditions, coded 0); (b) BP features (centered); and (c) terms to capture the two-way interactions between experimental conditions and BP features. Table 2 presents the results.

With respect to spontaneous partner evaluations, no main effects were observed for the closeness condition, $b = .007$, $t(5, 291) = .100$, $p = .921$, and the rejection condition, $b =$

¹ Because the data for deliberate partner evaluations were skewed, we performed a square-root transformation so that the data approximated a normal distribution. Accordingly, the reported statistics are based on this square-root transformed variable, except for the means and standard deviations, which are based on the original untransformed variable to facilitate interpretability and comprehension of the data. No other variable was transformed.

.003, $t(5, 291) = .041$, $p = .967$. Therefore, the experimental conditions did not have the expected effect on spontaneous partner evaluations. However, the expected main effect of BP features emerged. Consistent with the negativity bias hypothesis, BP features were associated with more negative spontaneous partner evaluations, $b = -.257$, $t(5, 291) = -2.604$, $p = .010$. No interaction terms were significant. Therefore, although inspection of means (see Figure 2A) suggests that the effect of BP features on spontaneous partner evaluations appeared to be mostly driven by the control condition, it is impossible to draw such a conclusion because the interaction terms were not significant. Therefore, BP features were generally associated with more negative spontaneous partner evaluations across experimental conditions, with no evidence of polarity.

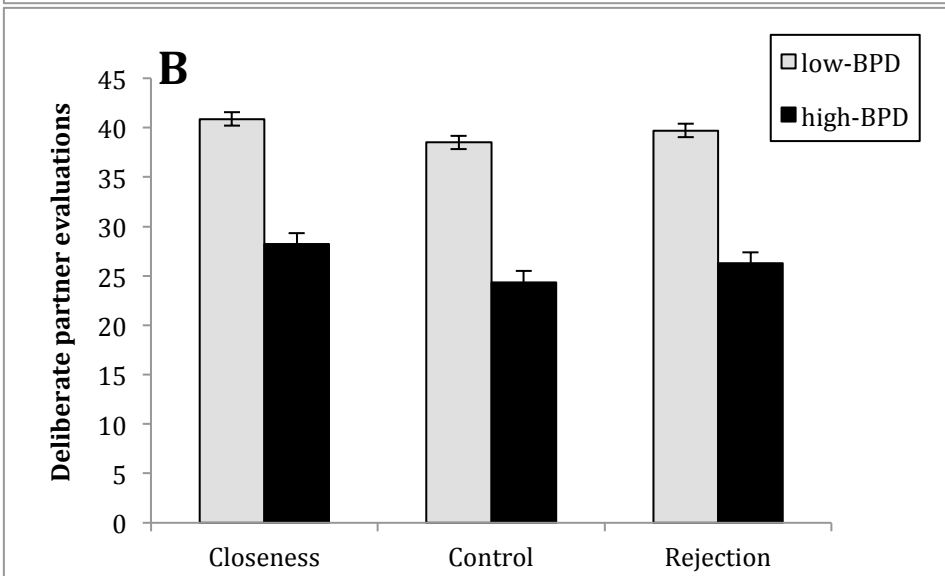
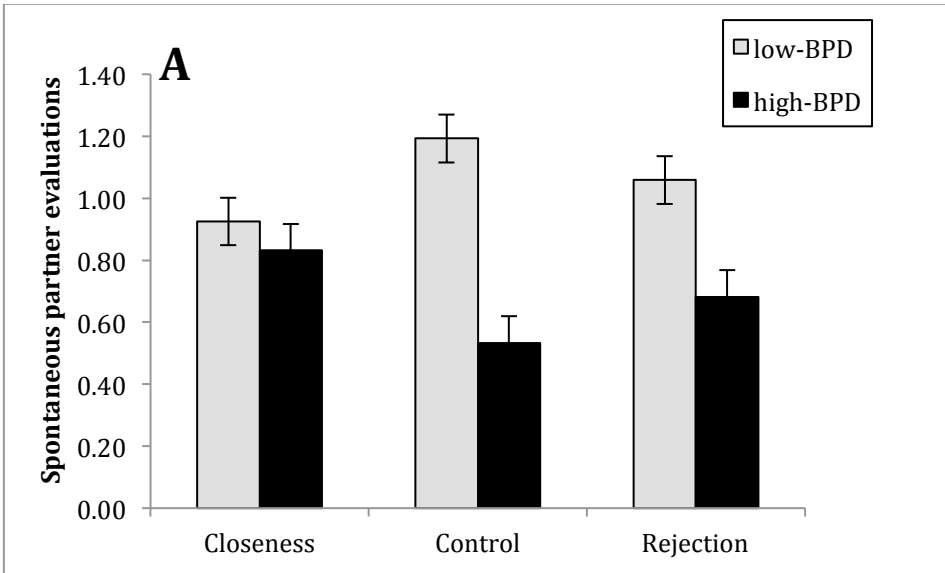
With respect to deliberate partner evaluations, a marginal main effect of closeness condition was observed, $b = .108$, $t(5, 291) = 1.735$, $p = .084$. The main effect of rejection condition was not significant, $b = -.084$, $t(5, 291) = -1.346$, $p = .179$. Consistent with the negativity bias hypothesis, BP features were associated with more negative deliberate partner evaluations, $b = -.427$, $t(5, 291) = -4.661$, $p < .001$. No interaction terms were significant. The pattern of results for deliberate evaluations was similar to that obtained for spontaneous evaluations; that is, on both types of evaluations BP features were generally associated with more negative partner evaluations across experimental conditions, with no evidence of polarity. Figure 1 illustrates the results.

With respect to spontaneous self evaluations, no main effects were observed for the closeness condition, $b = -.006$, $t(5, 291) = -.088$, $p = .930$, and the rejection condition, $b = .048$, $t(5, 291) = .722$, $p = .471$. Therefore, mirroring spontaneous partner evaluations, the experimental conditions did not have the expected effect on spontaneous self evaluations. Again, the expected main effect of BP features emerged. Consistent with the negativity bias hypothesis, BP features were associated with more negative spontaneous self evaluations, $b = -.303$, $t(5, 291) = -3.083$, $p = .002$. Contrary to partner evaluations, results showed a two-way interaction between closeness condition and BP features, $b = .176$, $t(5, 291) = 2.101$, $p = .037$. Simple slopes analyses (Aiken & West, 1991) indicated that this interaction was driven by the control condition, such that BP features were negatively related to spontaneous self evaluations after the control condition, $b = -.185$, $t(3, 288) = -3.084$, $p = .002$, but not after the closeness condition, $b = -.009$, $t(3, 288) = -.159$, $p = .874$. There was also a marginally

significant interaction between rejection condition and BP features, $b = .140$, $t(5, 291) = 1.797$, $p = .073$, but we refrained from interpreting this marginal interaction.

Table 3
Summary of regressions on spontaneous and deliberate partner evaluations

Predictor	Spontaneous evaluations			Deliberate evaluations		
	β	t	p	β	t	p
Closeness condition	.007	.100	.921	.108	1.735	.084
Rejection condition	.003	.041	.967	-.084	-1.346	.179
BP features	-.257	-2.604	.010	-.427	-4.661	.000
Closeness X BPF	.129	1.537	.125	.035	.450	.653
Rejection X BPF	.125	1.602	.110	.049	.670	.503



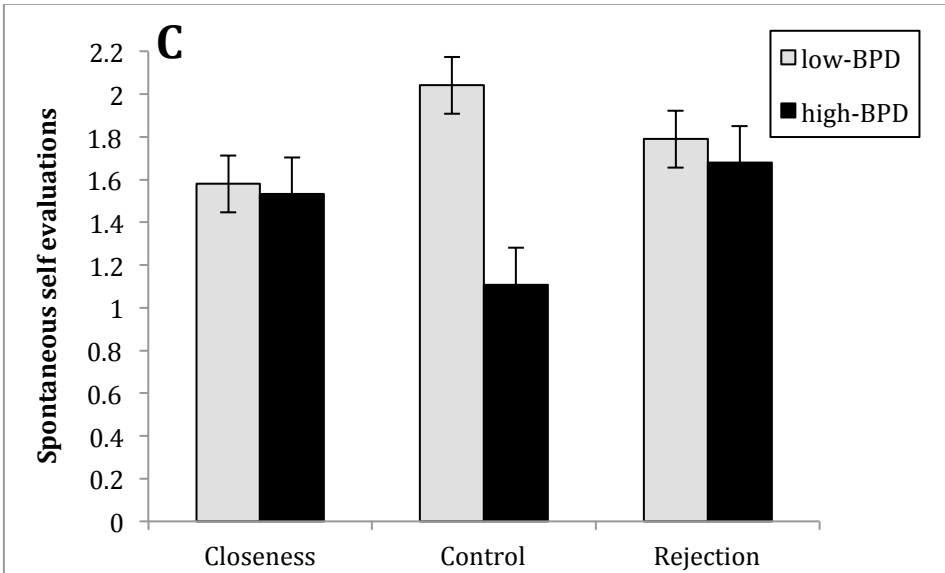


Figure 2. Scores above zero are positive, with higher scores reflecting higher positivity. Figure 2A: Spontaneous partner evaluations as a function of BP features (continuous) and study condition (closeness, control, rejection); Figure 2B: Deliberate partner evaluations as a function of BP features (continuous) and study condition (closeness, control, rejection); Figure 2C: Spontaneous self evaluations as a function of BP features (continuous) and study condition (closeness, control, rejection)

Note: Error bars represent standard errors.

7. Discussion

7.1. Summary of findings

Clinical observations describe evaluations of others as negative and polarized in BPD (e.g., Beck et al., 2004; Kernberg, 1986) and empirical research has obtained self-report, behavioral, neurobiological and physiological evidence of negativity biases in different aspects of social cognition in BPD. However, no study had examined the hypotheses that partner evaluations are negative and polarized in individuals with borderline personality (BP) features; nor had the interplay between spontaneous and deliberate partner evaluations been assessed directly in relation to BP features. Based on these considerations, this study tested, for the first time, the hypotheses that BP features are characterized by a negativity bias and polarity in spontaneous and deliberate partner evaluations and examined spontaneous self evaluations.

First, replicating prior research (LeBel & Campbell, 2009, 2013), we found that individuals had predominantly positive evaluations of themselves and of their partner. Second, contrary to expectations, the rejection and closeness manipulations did not influence spontaneous evaluations and their influence on deliberate partner evaluations was marginal and small. Third, based on the polarity hypothesis, we expected that individuals with high BP features would display greater positivity after the closeness condition and greater negativity after the rejection condition. However, the results did not confirm this pattern. Instead, we found evidence for a generalized negativity bias and the same pattern qualified both spontaneous and deliberate partner evaluations. Specifically, individuals with high BP features showed a reduced preference for their partner's initials and reported less positive reactions to their partner independent of study conditions. In addition, the obtained pattern of negative partner evaluations is accompanied by the finding that participants with high BP features also showed a reduced preference for their own initials relative to participants with low BP features, reflecting a negativity bias in spontaneous self evaluations. This effect emerged only in the control condition, contrary to partner evaluations.

The current study carries several implications for the understanding of romantic dysfunctions in individuals with BP features. The obtained pattern of reduced positivity for both spontaneous and deliberate evaluations (negativity bias) suggests that BP features can compromise the capacity to form positive partner evaluations and that this effect is not only rooted in explicit beliefs about one's partner but also depends on affective gut reactions to the partner resulting from the activation of mental associations in memory. Hence, the negativity bias in individuals with BP features seems to affect partner evaluations both at early and later processing stages (Gawronski & Bodenhausen, 2006, 2011). This may contribute to relationship dysfunctions or decrease the protective role of positive partner feelings; in fact, both spontaneous and deliberate positive partner evaluations are predictors of positive relationship outcomes (LeBel & Campbell, 2013) and healthy relationships have been shown to benefit to both individuals with BP features (Kuhlken et al., 2014) and individuals with BPD (Gunderson et al., 2003; Gunderson et al., 2006). Based on these considerations, our findings point to the importance of addressing both kinds of evaluations in order to help people with BP features achieving more satisfying relationships. Supporting this conclusion, previous research has shown that changes in deliberate evaluations do not necessarily lead to

corresponding changes in spontaneous evaluations (e.g., Gregg, Seibt, & Banaji, 2006), and vice-versa (e.g., Gawronski & LeBel, 2008).

It is important to nuance the conclusion that BP features may compromise the capacity to form positive partner evaluations. In fact, there is an alternative account to the present findings. Considering that there were no significant main effects of experimental manipulations on spontaneous and deliberate partner evaluations, it is plausible that the absence of any interaction between BP features and experimental manipulations was due to inappropriate manipulations. Consistent with this claim, previous research has successfully found variations in patterns of evaluations and affects in response to events as they occur in real-life interactions of individuals with BPD. Therefore, the obtained pattern of generalized negativity bias in individuals with high BP features may be an artificial finding. Problematically, given that the manipulations seemed to be ineffective, it remains possible that the use of effective manipulations would have led to alternative findings, such as (1) greater polarity in individuals with high BP features (more positive partner evaluations in one condition and more negative partner evaluations in the other condition among individuals with high BP features relative to individuals with low BP features); (2) or a specific negativity bias in individual with high BP features (e.g., more negative partner evaluations after the rejection scenario only among individuals with high BP features relative to individuals with low BP features). Future research should use different types of manipulations to determine what manipulations are best suited to reveal the potential relationships between BP features and evaluations of others at the spontaneous and deliberate levels of evaluations. Future research could also examine factors and traits that may account for or modulate the existence and strength of potential negativity biases and polarized evaluations in relation to BP features. In fact, previous research shows that traits associated with BPD, such as low self-esteem, insecure attachment, negative affect intensity and variability and deficits in emotion regulation strategies, can contribute to negative or polarized partner evaluations (Cavallo et al., 2012; Graham & Clark, 2006; Klein et al., 2016; Murray et al., 2002; Sadikaj et al., 2015). Hence, negative and/or polarized evaluations of others may be more strongly associated with some of these traits than with BP features. Future research could examine this question.

7.2. Limits and conclusion

This study has some limits and raises several questions that could be tested in future research. First, because this is the first study to examine spontaneous and deliberate partner evaluations in relation to BP features, the results should be replicated in future studies. In addition, it would be interesting to examine whether or not BPD patients would show the same pattern of results as the one obtained here with a non-clinical sample. Also, future research could test whether the negativity of spontaneous evaluations is a shared characteristic of different personality traits or a specific characteristic of individuals with BP features or BPD. An intriguing finding is that the reduced positivity of partner evaluations among participants with high (vs. low) BP features was not greater in the rejection condition relative to the control and closeness conditions. It is possible that BP features involve more negative partner evaluations in general, but it is also possible that the manipulations were not potent enough to reveal different patterns of partner evaluations following interpersonal situations of rejection and closeness. This latter interpretation is coherent with the absence of significant main effect of experimental conditions on our dependent variables (i.e., spontaneous and deliberate partner evaluations). Unfortunately, no manipulation check was included; therefore, it is impossible to determine whether the manipulations had any observable effect on the participants. The power may also have been too low to detect more subtle interaction effects between BP features and study conditions, as suggested by the inspection of means on the index reflecting spontaneous partner evaluations, but this remains an open empirical question. Future research could address these issues with larger sample sizes and by examining whether different manipulations would lead to different patterns of results in individuals with BP features or BPD. In this vein, an interesting avenue for future research is to examine how partner evaluations influence real-life romantic interactions in individuals with BP features using an ecological longitudinal design and analyzing the prospective relationships between spontaneous/deliberate partner evaluations and spontaneous/deliberate romantic behaviors in the natural environment of participants. In fact, there is evidence that spontaneous evaluations tend to be superior in predicting behaviors that are difficult to control (e.g., non-verbal responses) and behaviors under conditions of limited cognitive resources (e.g., ego depletion), whereas deliberate evaluations tend to be superior in predicting controllable behaviors (e.g., verbal behavior) and behaviors under conditions of sufficient cognitive resources (Perugini, 2005).

In summary, the present study suggests that BP features may be associated with relatively more negative evaluations at least in certain situations and that such negativity may not only be present at the deliberate level of evaluations, but also at the spontaneous level of evaluations. However, given that the experimental manipulations used in the present study did not show main effects on either kind of evaluations, this study does not allow concluding on the two main hypotheses that were contrasted: whether individuals with BP features are best characterized by a negativity bias (i.e., increased negativity, with no situations of increased positivity) or by polarity (i.e., increased negativity in some situations and increased positivity in other situations). Accordingly, this study cannot inform us on how individuals with BP features react to positive and negative relationship situations. Hence, more research is needed to clarify whether BP features are best characterized by negativity biases (e.g., Domes et al., 2009; Sieswerda et al., 2013) or by polarity (e.g., Beeney et al., 2016, Vater et al., 2015) at the spontaneous and deliberate level of partner evaluations. Nevertheless, the present study extends previous work and yields the first direct evidence that BP features may involve negative partner evaluations influencing both earlier and later processing stages, pointing to the importance of addressing both types of evaluations in order to help individuals with BP features achieving more satisfying relationships.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Abdul-Hamid, S., Denman, C., & Dudas, R. B. (2014). Self-relevant disgust and self-harm urges in patients with borderline personality disorder and depression: a pilot study with a newly designed psychological challenge. *Plos One*, *9*(6). doi: 10.1371/journal.pone.0099696
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Arntz, A., Appels, C., & Sieswerda, S. (2000). Hypervigilance in borderline disorder: A test with the emotional Stroop paradigm. *Journal of Personality Disorders*, *14*(4), 366-373.
- Arntz, A., & Haaf, J. T. (2012). Social cognition in borderline personality disorder: evidence for dichotomous thinking but no evidence for less complex attributions. *Behaviour research and therapy*, *50*(11), 707-718. doi: 10.1016/j.brat.2012.07.002
- Arntz, A., & Veen, G. (2001). Evaluations of others by borderline patients. *Journal of Nervous and Mental Disease*, *189*(8), 513-521. doi: 10.1097/00005053-200108000-00004
- Baccus, J. R., Baldwin, M. W., & Packer, D. J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science*, *15*, 498-502.

- Bargh, J. A. (1994). The four horsemen of automaticity: awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1-40). Hillsdale, NJ: Erlbaum.
- Barnow, S., Stopsack, M., Grabe, H. J., Meinke, C., Spitzer, C., Kronmueller, K., & Sieswerda, S. (2009). Interpersonal evaluation bias in borderline personality disorder. *Behaviour Research and Therapy, 47*(5), 359-365. doi: 10.1016/j.brat.2009.02.003
- Beck, A. T., Freeman, A., & Davis, D. D. (2004). *Cognitive therapy of personality disorders (2nd ed.)*. New York, NY: Guilford Press.
- Beck, A. T., & Freeman, A. M. (1990). *Cognitive therapy of personality disorders*. New York, NY: Guilford Press.
- Beeney, Joseph E., Hallquist, Michael N., Ellison, William D., & Levy, Kenneth N. (2016). Self-other disturbance in borderline personality disorder: Neural, self-report, and performance-based evidence. *Personality disorders, 7*(1), 28-39. doi: 10.1037/per0000127
- Berenson, K. R., Downey, G., Rafaeli, E., Coifman, K. G., & Paquin, N. L. (2011). The rejection-rage contingency in borderline personality disorder. *Journal of Abnormal Psychology, 120*(3), 681-690. doi: 10.1037/a0023335
- Bhatia, V., Davila, J., Eubanks-Carter, C., & Burckell, L. A. (2013). Appraisals of Daily Romantic Relationship Experiences in Individuals With Borderline Personality Disorder Features. *Journal of Family Psychology, 27*(3), 518-524. doi: 10.1037/a0032870
- Blackburn, R., Donnelly, J. P., Logan, C., & Renwick, S. J. D. (2004). Convergent and discriminative validity of interview and questionnaire measures of personality disorder in mentally disordered offenders: A multitrait-multimethod analysis using confirmatory factor analysis. *Journal of Personality Disorders, 18*(2), 129-150. doi: 10.1521/pedi.18.2.129.32779
- Boldero, J. M., Hulbert, C. A., Bloom, L., Cooper, J., Gilbert, F., Mooney, J. L., & Salinger, J. (2009). Rejection sensitivity and negative self-beliefs as mediators of associations between the number of borderline personality disorder features and self-reported adult attachment. *Personality and Mental Health, 3*, 248-262.
- Bouchard, S., Godbout, N., & Sabourin, S. (2009). Sexual attitudes and activities in women with borderline personality disorder involved in romantic relationships. *Journal of Sex & Marital Therapy, 35*(2), 106-121. doi: 10.1080/00926230802712301
- Bouchard, S., Sabourin, S., Lussier, Y., & Villeneuve, E. (2009). Relationship quality and stability in couples when one partner suffers from borderline personality disorder. *Journal of Marital and Family Therapy, 35*(4), 446-455. doi: 10.1111/j.1752-0606.2009.00151.x
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3-5.
- Cavallo, J. V., Holmes, J. G., Fitzsimons, G. M., Murray, S. L., & Wood, J. V. (2012). Managing motivational conflict: how self-esteem and executive resources influence self-regulatory responses to risk. *Journal of Personality and Social Psychology, 103*(3), 430-451. doi: 10.1037/a0028821

- Coifman, K. G., Berenson, K. R., Rafaeli, E., & Downey, G. (2012). From negative to positive and back again: polarized affective and relational experience in borderline personality disorder. *Journal of Abnormal Psychology, 121*(3), 668-679. doi: 10.1037/a0028502
- Daley, S. E., Burge, D., & Hammen, C. (2000). Borderline personality disorder symptoms as predictors of 4-year romantic relationship dysfunction in young women: addressing issues of specificity. *Journal of Abnormal Psychology, 109*(3), 451-460. doi: 10.1037//0021-843x.109.3.451
- Domes, G., Czeschnek, D., Weidler, F., Berger, C., Fast, K., & Herpertz, S. C. (2008). Recognition of facial affect in borderline personality disorder. *Journal of Personality Disorders, 22*(2), 135-147. doi: 10.1521/pedi.2008.22.2.135
- Domes, G., Schulze, L., & Herpertz, S. C. (2009). Emotion recognition in borderline personality disorder-a review of the literature. *Journal of Personality Disorders, 23*(1), 6-19.
- Domes, G., Winter, B., Schnell, K., Vohs, K., Fast, K., & Herpertz, S. C. (2006). The influence of emotions on inhibitory functioning in borderline personality disorder. *Psychological Medicine, 36*(8), 1163-1172. doi: 10.1017/s0033291706007756
- Dyck, M., Habel, U., Slodczyk, J., Schlummer, J., Backes, V., Schneider, F., & Reske, M. (2009). Negative bias in fast emotion discrimination in borderline personality disorder. *Psychological Medicine, 39*(5), 855-864. doi: 10.1017/s0033291708004273
- Edens, J. F., Marcus, D. K., & Ruiz, M. A. (2008). Taxometric analyses of borderline personality features in a large-scale male and female offender sample. *Journal of Abnormal Psychology, 117*(3), 705-711. doi: 10.1037/0021-843x.117.3.70.5
- Evans, D., Dalgleish, T., Dudas, R. B., Denman, C., Howard, M., & Dunn, B. D. (2015). Examining the shared and unique features of self-concept content and structure in borderline personality disorder and depression. *Cognitive Therapy and Research, 39*(5), 613-626. doi: 10.1007/s10608-015-9695-3
- Gardner, K., & Qualter, P. (2009). Reliability and validity of three screening measures of borderline personality disorder in a nonclinical population. *Personality and Individual Differences, 46*(5-6), 636-641.
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692-731. doi: <http://dx.doi.org/10.1037/0033-2909.132.5.692>
- Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: theory, evidence, and open questions. *Advances in Experimental Social Psychology, 44*, 59-127. doi: 10.1016/b978-0-12-385522-0.00002-0
- Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology, 44*(5), 1355-1361. doi: 10.1016/j.jesp.2008.04.005
- Graham, S. M., & Clark, M. S. (2006). Self-esteem and organization of valenced information about others: The "Jekyll and Hyde"-ing of relationship partners. *Journal of Personality and Social Psychology, 90*(4), 652-665. doi: 10.1037/0022-3514.90.4.652
- Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology, 90*(1), 1-20. doi: 10.1037/0022-3514.90.1.1
- Gunderson, J. G., Bender, D., Sanislow, C., Yen, S., Rettew, J. B., Dolan-Sewell, R., . . . Skodol, A. E. (2003). Plausibility and possible determinants of sudden "remissions" in

- borderline patients. *Psychiatry-Interpersonal and Biological Processes*, 66(2), 111-119. doi: 10.1521/psyc.66.2.111.20614
- Gunderson, J. G., Daversa, M. T., Grilo, C. M., McGlashan, T. H., Zanarini, M. C., Shea, M. T., . . . Stout, R. L. (2006). Predictors of 2-year outcome for patients with borderline personality disorder. *American Journal of Psychiatry*, 163(5), 822-826. doi: 10.1176/appi.ajp.163.5.822
- Hagenhoff, M., Franzen, N., Gerstner, L., Koppe, G., Sarnmer, G., Netter, P., . . . Lis, S. (2013). Reduced sensitivity to emotional facial expressions in borderline personality disorder: effects of emotional valence and intensity. *Journal of Personality Disorders*, 27(1), 19-35.
- Haslam, N. (2003). Categorical versus dimensional models of mental disorder: the taxometric evidence. *Australian and New Zealand Journal of Psychiatry*, 37(6), 696-704. doi: 10.1080/j.1440-1614.2003.01258.x
- Hawes, D. J., Helyer, R., Herlianto, E. C., & Willing, J. (2013). Borderline personality features and implicit shame-prone self-concept in middle childhood and early adolescence. *Journal of Clinical Child and Adolescent Psychology*, 42(3), 302-308. doi: 10.1080/15374416.2012.723264
- Hedrick, A. N., & Berlin, H. A. (2012). Implicit self-esteem in borderline personality and depersonalization disorder. *Frontiers in Psychology*, 3. doi: 10.3389/fpsyg.2012.00091
- Hill, J., Pilkonis, P., Morse, J., Feske, U., Reynolds, S., Hope, H., . . . Broyden, N. (2008). Social domain dysfunction and disorganization in borderline personality disorder. *Psychological Medicine*, 38(1), 135-146. doi: 10.1017/s0033291707001626
- Hill, J., Stepp, S. D., Wan, M. W., Hope, H., Morse, J. Q., Steele, M., . . . Pilkonis, P. A. (2011). Attachment, borderline personality, and romantic relationship dysfunction. *Journal of Personality Disorders*, 25(6), 789-805.
- Hurlbert, D. F., Apt, C., & White, L. C. (1992). An empirical-examination into the sexuality of women with borderline personality-disorder. *Journal of Sex & Marital Therapy*, 18(3), 231-242.
- Hyler, S. E. (1994). *Personality Diagnostic Questionnaire-4*. New York: New York State Psychiatric Institute.
- Hyler, S. E., Skodol, A. E., Kellman, H. D., Oldham, J. M., & Rosnick, L. (1990). Validity of the personality diagnostic questionnaire revised - comparison with 2 structured interviews. *American Journal of Psychiatry*, 147(8), 1043-1048.
- Kernberg, O. F. (1986). *Severe personality disorders: psychotherapeutic strategies*. New Haven, CT: Yale University Press.
- Klein, S. R., Renshaw, K. D., & Curby, T. W. (2016). Emotion regulation and perceptions of hostile and constructive criticism in romantic relationships. *Behavior Therapy*, 47(2), 143-154.
- Kopala-Sibley, D. C., Zuroff, D. C., Russell, J. J., Moskowitz, D. S., & Paris, J. (2012). Understanding heterogeneity in borderline personality disorder: differences in affective reactivity explained by the traits of dependency and self-criticism. *Journal of Abnormal Psychology*, 121(3), 680-691. doi: 10.1037/a0028513
- Krause-Utz, A., Oei, N. Y. L., Niedtfeld, I., Bohus, M., Spinoven, P., Schmahl, C., & Elzinga, B. M. (2012). Influence of emotional distraction on working memory performance in borderline personality disorder. *Psychological Medicine*, 42(10), 2181-2192. doi: 10.1017/s0033291712000153

- Kuhlken, K., Robertson, C., Benson, J., & Nelson-Gray, R. (2014). The interaction of borderline personality disorder symptoms and relationship satisfaction in predicting affect. *Personality Disorders-Theory Research and Treatment*, *5*(1), 20-25. doi: 10.1037/per0000013
- LeBel, E. P., & Campbell, L. (2009). Implicit partner affect, relationship satisfaction, and the prediction of romantic breakup. *Journal of Experimental Social Psychology*, *45*(6), 1291-1294. doi: 10.1016/j.jesp.2009.07.003
- LeBel, E. P., & Campbell, L. (2013). The interactive role of implicit and explicit partner evaluations on ongoing affective and behavioral romantic realities. *Social Psychological and Personality Science*, *4*(2), 167-174. doi: 10.1177/1948550612448196
- LeBel, E. P., & Gawronski, B. (2009). How to find what's in a name: Scrutinizing the optimality of five scoring algorithms for the name-letter task. *European Journal of Personality*, *23*, 85-106.
- Matzke, B., Herpertz, S. C., Berger, C., Fleischer, M., & Domes, G. (2014). Facial reactions during emotion recognition in borderline personality disorder: a facial electromyography study. *Psychopathology*, *47*(2), 101-110. doi: 10.1159/000351122
- Mongeon, F., Gawronski, B., & Gagnon, J. (2016). *Negativity bias and instability in spontaneous and deliberate evaluations of others: the role of borderline personality disorder features*. University of Montreal.
- Murray, S. L., Gomillion, S., Holmes, J. G., Harris, B., & Lamarche, V. (2013). The Dynamics of Relationship Promotion: Controlling the Automatic Inclination to Trust. *Journal of Personality and Social Psychology*, *104*(2), 305-334. doi: 10.1037/a0030513
- Murray, S. L., Pinkus, R. T., Holmes, J. G., Harris, B., Gomillion, S., Aloni, M., . . . Leder, S. (2011). Signaling When (and When Not) to Be Cautious and Self-Protective: Impulsive and Reflective Trust in Close Relationships. *Journal of Personality and Social Psychology*, *101*(3), 485-502. doi: 10.1037/a0023233
- Murray, S. L., Rose, P., Bellavia, G. M., Holmes, J. G., & Kusche, A. G. (2002). When rejection stings: How self-esteem constrains relationship-enhancement processes. *Journal of Personality and Social Psychology*, *83*(3), 556-573. doi: 10.1037//0022-3514.83.3.556
- Napolitano, L. A., & McKay, D. (2007). Dichotomous thinking in borderline personality disorder. *Cognitive Therapy and Research*, *31*(6), 717-726. doi: 10.1007/s10608-007-9123-4
- Nuttin, M. J. Jr. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology*, *64*, 723-739.
- Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*, *44*(1), 29-45.
- Portella, M. J., Soler, J., Tiana, T., Barrachina, J., Pascual, J. C., Tejero, A., . . . Perez, V. (2011). Slow processing in borderline personality disorder: the emotional Stroop paradigm. *Actas Espanolas De Psiquiatria*, *39*(6), 356-362.
- Renneberg, B., Herm, K., Hahn, A., Staebler, K., Lammers, C.-H., & Roepke, S. (2012). Perception of social participation in borderline personality disorder. *Clinical Psychology & Psychotherapy*, *19*(6), 473-480. doi: 10.1002/cpp.772

- Ross, J. M., & Babcock, J. C. (2009). Proactive and reactive violence among intimate partner violent men diagnosed with antisocial and borderline personality disorder. *Journal of Family Violence, 24*(8), 607-617. doi: 10.1007/s10896-009-9259-y
- Rothschild, L., Cleland, C., Haslam, N., & Zimmerman, M. (2003). A taxometric study of borderline personality disorder. *Journal of Abnormal Psychology, 112*(4), 657-666. doi: 10.1037/0021-843x.112.4.657
- Rüsch, N., Corrigan, P. W., Bohus, M., Kuehler, T., Jacob, G. A., & Lieb, K. (2007). The impact of posttraumatic stress disorder on dysfunctional implicit and explicit emotions among women with borderline personality disorder. *Journal of Nervous and Mental Disease, 195*(6), 537-539. doi: 10.1097/NMD.0b013e318064e7fc
- Rusch, N., Lieb, K., Gottler, I., Hermann, C., Schramm, E., Richter, H., . . . Bohus, M. (2007). Shame and implicit self-concept in women with borderline personality disorder. *American Journal of Psychiatry, 164*(3), 500-508. doi: 10.1176/appi.ajp.164.3.500
- Rusch, N., Schulz, D., Valerius, G., Steil, R., Bohus, M., & Schmahl, C. (2011). Disgust and implicit self-concept in women with borderline personality disorder and posttraumatic stress disorder. *European Archives of Psychiatry and Clinical Neuroscience, 261*(5), 369-376. doi: 10.1007/s00406-010-0174-2
- Sadikaj, G., Moskowitz, D. S., Russell, J. J., Zuroff, D. C., & Paris, J. (2013). Quarrelsome behavior in borderline personality disorder: influence of behavioral and affective reactivity to perceptions of others. *Journal of Abnormal Psychology, 122*(1), 195-207. doi: 10.1037/a0030871
- Sadikaj, G., Moskowitz, D. S., & Zuroff, D. C. (2015). Intrapersonal variability in negative affect as a moderator of accuracy and bias in interpersonal perception. *Journal of Personality Disorders, 29*(4), 468-485.
- Sadikaj, G., Russell, J. J., Moskowitz, D. S., & Paris, J. (2010). Affect dysregulation in individuals with borderline personality disorder: persistence and interpersonal triggers. *Journal of Personality Assessment, 92*(6), 490-500. doi: 10.1080/00223891.2010.513287
- Salekin, R. T., Trobst, K. K., & Krioukova, M. (2001). Construct validity of psychopathy in a community sample: A nomological net approach. *Journal of Personality Disorders, 15*(5), 425-441. doi: 10.1521/pedi.15.5.425.19196
- Sieswerda, S., Arntz, A., & Kindt, M. (2007). Successful psychotherapy reduces hypervigilance in borderline personality disorder. *Behavioural and Cognitive Psychotherapy, 35*(4), 387-402. doi: 10.1017/s1352465807003694
- Sieswerda, S., Arntz, A., Mertens, I., & Vertommen, S. (2007). Hypervigilance in patients with borderline personality disorder: Specificity, automaticity, and predictors. *Behaviour Research and Therapy, 45*(5), 1011-1024. doi: 10.1016/j.brat.2006.07.012
- Sieswerda, S., Arntz, A., & Wolfis, M. (2005). Evaluations of emotional noninterpersonal situations by patients with borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 36*(3), 209-225. doi: 10.1016/j.jbtep.2005.05.004
- Sieswerda, S., Barnow, S., Verheul, R., & Arntz, A. (2013). Neither dichotomous nor split, but schema-related negative interpersonal evaluations characterize borderline patients. *Journal of Personality Disorders, 27*(1), 36-52.
- Tragesser, S. L., & Benfield, J. (2012). Borderline personality disorder features and mate retention tactics. *Journal of Personality Disorders, 26*(3), 334-344.

- Tragesser, S. L., Lippman, L. G., Trull, T. J., & Barrett, K. C. (2008). Borderline personality disorder features and cognitive, emotional, and predicted behavioral reactions to teasing. *Journal of Research in Personality, 42*(6), 1512-1523. doi: 10.1016/j.jrp.2008.07.003
- Trull, T. J., Useda, J. D., Conforti, K., & Doan, B. T. (1997). Borderline personality disorder features in nonclinical young adults: 2. Two-year outcome. *Journal of Abnormal Psychology, 106*(2), 307-314. doi: 10.1037//0021-843x.106.2.307
- Valentiner, D. P., Hiraoka, R., & Skowronski, J. J. (2014). Borderline personality disorder features, self-verification, and committed relationships. *Journal of Social and Clinical Psychology, 33*(5), 463-480.
- Vater, A., Schroeder-Abe, M., Weissgerber, S., Roepke, S., & Schuetz, A. (2015). Self-concept structure and borderline personality disorder: evidence for negative compartmentalization. *Journal of Behavior Therapy and Experimental Psychiatry, 46*, 50-58. doi: 10.1016/j.jbtep.2014.08.003
- von Ceumern-Lindenstjerna, I.-A., Brunner, R., Parzer, P., Mundt, C., Fiedler, P., & Resch, F. (2010). Attentional bias in later stages of emotional information processing in female adolescents with borderline personality disorder. *Psychopathology, 43*(1), 25-32. doi: 10.1159/000255960
- Westen, D. (1991). Social cognition and object relations. *Psychological Bulletin, 109*(3), 429-455. doi: 10.1037/0033-2909.109.3.429
- Whisman, M. A., & Schonbrun, Y. C. (2009). Social consequences of borderline personality disorder symptoms in a population-based survey: marital distress, marital violence, and marital disruption. *Journal of Personality Disorders, 23*(4), 410-415.
- Wingenfeld, K., Mensebach, C., Rullkoetter, N., Schlosser, N., Schaffrath, C., Woermann, F. G., . . . Beblo, T. (2009). Attentional bias to personally relevant words in borderline personality disorder is strongly related to comorbid posttraumatic stress disorder. *Journal of Personality Disorders, 23*(2), 141-155.
- Wingenfeld, K., Rullkoetter, N., Mensebach, C., Beblo, T., Mertens, M., Kreisel, S., . . . Woermann, F. G. (2009). Neural correlates of the individual emotional Stroop in borderline personality disorder. *Psychoneuroendocrinology, 34*(4), 571-586. doi: 10.1016/j.psyneuen.2008.10.024
- Winter, D., Herbert, C., Koplín, K., Schmahl, C., Bohus, M., & Lis, S. (2015). Negative evaluation bias for positive self-referential information in borderline personality disorder. *Plos One, 10*(1). doi: 10.1371/journal.pone.0117083
- Winter, D., Krause-Utz, A., Lis, S., Chiu, C. D., Lanius, R. A., Schriener, F., . . . Schmahl, C. (2015). Dissociation in borderline personality disorder: Disturbed cognitive and emotional inhibition and its neural correlates. *Psychiatry Research-Neuroimaging, 233*(3), 339-351. doi: 10.1016/j.psychresns.2015.05.018
- Witthoft, M., Borgmann, E., White, A., & Dyer, A. (2015). Body-related attentional biases in patients with posttraumatic stress disorder resulting from childhood sexual abuse with and without co-occurring borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 46*, 72-77. doi: 10.1016/j.jbtep.2014.07.010

Conclusion

Summary of objectives. The main aim of the thesis and of the two empirical papers was to examine the interplay between spontaneous and deliberate evaluations of others in relation to BP features. Evaluations of others are ubiquitous and influence social interactions from the moment we meet someone and form first impressions (e.g., Gawronski, Geschke, & Banse, 2003) to the most intimate interactions that we share with a loved one (e.g., LeBel & Campbell, 2009). Accordingly, in order to reflect the range of interpersonal situations occurring in real-life, we conducted two studies that focused on two types of social interactions, namely impression formation (Study 1) and partner evaluations (Study 2). Specifically, we hypothesized that individuals with high BP features would display relatively more negativity and splitting in their evaluations of a movie character (Study 1) and of their partner (Study 2), in line with clinical observations (e.g., Beck et al., 2004; Kernberg, 1986). The present thesis extends previous work on evaluations of others in borderline personality in several ways: (1) it examined an aspect of borderline personality that has received no empirical attention in previous BPD research, namely spontaneous evaluations of others, (2) it investigated the interplay between spontaneous and deliberate evaluations and focused on potential patterns of convergence and dissociation between the two kinds of evaluations, (3) it utilized experimental manipulations to examine whether positive or negative interpersonal situations can potentially moderate the relationship between spontaneous/deliberate evaluations and BP features, (4) it simultaneously tested two hypotheses (i.e., negativity bias and splitting), which have received significant attention in clinical theories of BPD (A. T. Beck & Freeman, 1990; Aaron T. Beck et al., 2004; Kernberg, 1986) and (5) combined two studies to examine the interplay between spontaneous and deliberate evaluations in two different interpersonal contexts, that is, impression formation (Study 1) and romantic interactions (Study 2). Because the splitting hypothesis can be further decomposed in two more specific predictions (i.e., instability and polarization), each study examined one of these two predictions. Specifically, the first paper examined instability and the second paper examined polarization.

Study 1. The first paper investigated spontaneous and deliberate evaluations of a movie character. For each participant, spontaneous and deliberate evaluations of the target character

were measured twice, that is, (1) once after a positive movie clip depicting a situation of romantic connection and (2) once after a negative clip depicting a situation of betrayal and verbal fighting. We expected that individuals with high BP features would display relatively more negativity in both their spontaneous and deliberate evaluations of the movie character. We also tested whether individuals with high BP features would display relatively greater instability in their spontaneous and deliberate evaluations of the movie character. Specifically, we expected that BP features would be related with greater switches from positive to negative evaluations or from negative to positive evaluations between the two movie clips, in line with Criterion 2 of BPD in DSM-5 (“close relationships often viewed in extremes of idealization and devaluation”; American Psychiatric Association, 2013).

Results showed that participants displayed overall positive evaluations of the movie character, replicating previous research showing that evaluations of movie characters are overall more positive than negative (Sieswerda et al., 2013). Contrary to the splitting hypothesis, the results of Study 1 did not show the expected pattern of increased instability in participants with high BP features. Specifically, on both the implicit and explicit measures, BP features were not related with greater switches from positive to negative evaluations or from negative to positive evaluations. With respect to the negativity bias hypothesis, we expected BP features to be related with relatively more negative (less positive) evaluations of the movie character, that is, individuals with high BP features would form *relatively* more negative (less positive) evaluations than individuals with low BP features. Results on the implicit and explicit measures brought support for this hypothesis. Specifically, BP features involved a reduced capacity to positively modulate evaluations of the movie character. For spontaneous evaluations, this effect emerged after the negative clip (i.e., conjugal dispute) and not after the positive clip, whereas for deliberate evaluations, this effect emerged after participants watched the positive clip first and there was no effect after the negative clip. The results obtained on the explicit and implicit measures converge with empirical evidence of negativity biases in BPD (e.g., Domes et al., 2009; Sieswerda et al., 2013) and with the view that BPD involves more negative evaluations of the world and others (e.g., Beck et al., 2004; Kernberg, 1986). Results also suggest that while individuals with high BP features may not have negative evaluations in absolute terms (i.e., below an absolute value of zero), they nevertheless showed a reduced capacity to modulate their evaluations positively relative to individuals with low BP

features. Moreover, this pattern of reduced positivity is accompanied by the finding that after both clips participants with high BP features reported more negative emotions in the PANAS relative to participants with low BP features. The results do not converge with the splitting hypothesis and therefore diverge with clinical observations of instability stemming from cognitive theories (A. T. Beck & Freeman, 1990; Aaron T. Beck et al., 2004) and psychodynamic theories (Kernberg, 1986).

However, the results and conclusions pertaining to the negativity bias and splitting hypotheses should be considered provisional for the following reasons: The results on the implicit measure relied on a marginal three-way interaction, which raises some questions about the validity or the potential reliability of these findings. Moreover, this is the first study to examine the relationship between spontaneous evaluations of others and BP features; therefore, the present claims about this relationship remain preliminary because these are based on only one effect, and this is a marginally significant effect. Given these limits, the asymmetry of the findings obtained with the implicit and explicit measures raises additional questions. The obtained asymmetry is that, for spontaneous evaluations the negativity bias emerged only after the negative clip, whereas for deliberate evaluations the negativity bias emerged only after watching the positive clip first. This asymmetry can be interpreted in two ways: (1) It may be artificially caused by the low reliability of the implicit measure and by potential measurement errors or (2) it may reflect genuine differences in the relationships between BP features and the two types of evaluations. In the former, low reliability of measurement tools can hinder real effects; in fact, past research has shown that lower levels of reliability in dependent variables are associated with lower levels of replicability, which can make it difficult to capture potentially real effects (Lebel & Paunonen, 2011). In other words, the fact that the evaluative priming task did not reveal any effect in one condition of the study (i.e., after the positive clip) could be due to its low reliability; in fact, because of its low reliability, this measurement tool sometimes fail to capture effects even when they are present. It is also possible that measurement error reduced the power of the measures. For instance, the stimuli that were used in the tasks were not previously validated. Hence, more research with different measurement tools will be needed to address these questions. Alternatively, if the asymmetry of the findings obtained with the implicit and explicit measures reflects genuine effects, this would suggest that the reduced positivity of individuals with high BP features was

shaped at early processing stages that do not require intention and effort for the negative clip and that it was shaped at later processing stages that are more largely guided by effortful and intentional thinking processes for participants who watched the positive clip first (Gawronski & Bodenhausen, 2011). These considerations are discussed in more details below (in the implications). In sum, this is the first study to examine spontaneous evaluations of others in individuals with BP features. In doing so, the present results suggest that this may be a promising avenue of research in order to shed new light on the evaluative mechanisms underlying interpersonal difficulties in individuals with BP features. However, given the discussed limits, the results and conclusions regarding potentially less positive spontaneous evaluations in individuals with BP features should be considered provisional until future research can test the replicability of the findings.

Study 2. The second paper complemented the first paper in different ways. First, in order to investigate evaluations of others in a different interpersonal context, the second paper examined evaluations within romantic relationships; accordingly, we measured partner evaluations instead of evaluations of a movie character. In addition, the second paper examined spontaneous evaluations with a name-letter task instead of an evaluative priming task. Third, the two papers examined different aspects of splitting, namely instability (Study 1) and polarization (Study 2). Finally, the second paper examined spontaneous self evaluations. In the second study, spontaneous and deliberate evaluations were measured after participants imagined either (a) that their partner rejected them or (b) sought connection or (c) imagined no scenario (control condition).

Results of Study 2 showed that individuals had predominantly positive evaluations of themselves and of their partner, which replicates prior research (LeBel & Campbell, 2009, 2013). This result is also in line with the overall greater positivity found in Study 1. However, the analyses did not reveal the expected main effects of the rejection and closeness manipulations; specifically, participants who imagined that their partner rejected them or sought connection did not differ from control participants on their spontaneous partner evaluations; moreover, the effect of experimental conditions was marginal and small for deliberate partner evaluations. In addition, the notion of polarity derived from splitting led us to expect that individuals with high BP features would display greater positivity after the closeness condition and greater negativity after the rejection condition. However, the results

did not confirm this pattern. Therefore, in line with Study 1, Study 2 did not find evidence of splitting. Instead, results of study 2 showed that participants with high BP features had a reduced preference for their partner's initials and reported less positive reactions to their partner independent of study conditions. In addition, the obtained pattern of negative partner evaluations is accompanied by the finding that participants with high BP features also showed a reduced preference for their own initials relative to participants with low BP features, reflecting a negativity bias in spontaneous self evaluations, in line with previous research (Zeigler-Hill & Abraham, 2006).

Taken together, the results of Study 2 could be interpreted in two different ways: (1) They may be interpreted as evidence that BP features involves negativity biases that are shaped at relatively early and late processing stages. (2) Alternatively, given that the experimental manipulations did not produce any main effect of partner evaluations, these results may not accurately reflect partner evaluations in individuals with BP features. Failures to experimentally manipulate partner evaluations may be attributed to ineffective manipulations or to the use of a non-significant moderator (in this case BP features). In this vein, Murray et al. (2011) conducted a series of experimental studies where they attempted to manipulate partner evaluations and other relationship variables and the experimental manipulations did not systematically succeed to produce main effects. For instance, participants subliminally conditioned to evaluate their partner positively did not differ from controls in their deliberate partner evaluations; however, they did report greater trust, as expected (Murray et al., 2011, see Study 1). Similarly, participants who were asked to vividly describe a time when their partner had seriously hurt them did not show more negative spontaneous partner evaluations than controls regarding the main effect of experimental conditions (Murray et al., 2011, see Study 5). However, participants who were led to believe that their partner had provided a laundry list of their faults did show more negative spontaneous partner evaluations than controls (Murray et al., 2011, see Study 4). Moreover, across the different studies of Murray et al. (2011), effects of experimental conditions were more reliably detected when additional moderators were considered (e.g., working memory capacity). Based on these results, different reasons may explain why experimental conditions may not reliably produce main effects on partner evaluations: (1) changes in partner evaluations may be easier to detect when variables reflecting personality differences are considered as moderators and it is possible that BP

features is not a significant moderator, (2) partner evaluations might be more difficult to manipulate than other relationship variables; in this case, if the true effects of experimental manipulations on partner evaluations are low, these effects may be more difficult to obtain without sufficient power and they may not systematically emerge across studies, (3) if effects are low, it may be necessary to use manipulations that create larger effects to observe changes in partner evaluations, (4) spontaneous and deliberate partner evaluations may react to different manipulations (this issue is discussed in more details later). More research is needed to examine these different possibilities; for instance, future research may attempt to determine what manipulations are best suited to influence spontaneous and deliberate evaluations and to identify moderators. With respect to moderators, there is growing evidence showing that self-esteem and working memory capacity can moderate the relationship between relationship threats and dependent measures such as partner evaluations (Murray, Gomillion, Holmes, & Harris, 2015; Murray et al., 2013; Murray, Holmes, MacDonald, & Ellsworth, 1998; Murray et al., 2011; Murray et al., 2002).

We hope that future research will help clarify whether BP features (and BPD) are indeed characterized by more negative (less positive) partner evaluations or splitting. In fact, other empirical research suggests that relatively negative partner evaluations may contribute to relationship dysfunctions and decrease the protective role of positive partner feelings. For instance, spontaneous and deliberate positive partner evaluations both are predictors of positive relationship outcomes (LeBel & Campbell, 2013) and healthy relationships have been shown to benefit to both individuals with BP features (Kuhlken et al., 2014) and individuals with BPD (Gunderson et al., 2003; Gunderson et al., 2006). Based on these considerations, it seems important to address both kinds of evaluations in order to help people with BP features achieving more satisfying relationships. Supporting this claim, previous research has shown that changes in deliberate evaluations do not necessarily lead to corresponding changes in spontaneous evaluations (e.g., Gregg et al., 2006), and vice-versa (e.g., Gawronski & LeBel, 2008).

Integration of the results

In brief, although the results of Studies 1-2 provide preliminary evidence that BP features may be characterized by negativity biases in spontaneous and deliberate evaluations of others,

the overall pattern of findings suggest that this conclusion remains provisional until future research is able to dismiss alternative accounts. Moreover, although there was no evidence of splitting in either study, this finding could also reflect some of the issues that were raised in both studies.

The main contribution of this thesis is to examine spontaneous evaluations of other people in relation to BP features for the first time and to do so in two studies that used different and complementary methodologies. Notwithstanding the contribution and novelty of the present thesis, we also have identified several limits. These are discussed in more details in the respective papers, but two important shortcomings deserve more consideration in order to appropriately nuance the conclusions and interpretations that we have derived from the findings of Studies 1-2. First, with respect to Study 1, we did find a pattern of negativity biases among individuals with high BP features; however, because this pattern relied on a marginally significant three-way interaction for the implicit measure, these results should be considered with prudence. Future studies are necessary to test the replicability of the findings and achieve accurate estimates of the obtained effects across multiple experiments. Second, with respect to Study 2, the absence of main effects of experimental manipulations suggest that the obtained pattern of generalized negativity in individuals with high BP features may artificially result from the use of ineffective manipulations. Future research should use different types of manipulations to determine what manipulations are best suited to capture the interplay between BP features and evaluations of others. As discussed in the introduction of this thesis, researchers have reported mixed findings with respect to the existence and nature of negativity biases in BPD. These mixed findings can be interpreted in different ways: (a) negativity biases may be present only in certain individuals with BPD, (b) these biases may be easier to detect in individuals with BPD than in individuals with BP features, (c) negativity biases may emerge only in certain contexts or (d) in certain aspects of social-cognitive functioning among individuals with BPD (e.g., evaluations of self vs. evaluations of others, spontaneous evaluations vs. deliberate evaluations). Future research could attempt to disentangle these issues (a) by using multidimensional assessment of BPD to take into consideration its heterogeneity (e.g., by using questionnaires with multiple dimensions), (b) by examining BPD in different types of experimental conditions in the laboratory and in different types of real-life situations in the natural environment of participants, (c) and by distinguishing between

different types of negativity biases (e.g., negativity biases at the spontaneous vs. deliberate levels of evaluations).

Differential implications for spontaneous and deliberate evaluations². There is evidence that both spontaneous and deliberate evaluations contribute to interpersonal functioning and behaviors. In fact, past research suggests that spontaneous evaluations are best in predicting spontaneous behavior (e.g., eye gaze in interracial interactions predicted by implicit measures of racial prejudice), whereas deliberate evaluations are best in predicting deliberate behavior (e.g., content of verbal responses in interracial interactions predicted by explicit measures of racial prejudice) (e.g., Asendorpf, Banse, & Mucke, 2002; Fazio et al., 1995, cited in Gawronski & De Houwer, 2014, p. 294). Past research has also shown that negative spontaneous evaluations of others lead to more negative behaviors, which in turn lead to more negative interactions (Dovidio, Kawakami, & Gaertner, 2002); hence, the potential propensity of individuals with BP features to evaluate others negatively may contribute to dysfunctional interpersonal behaviors. Moreover, the literature on romantic relationships brought ample evidence that spontaneous partner evaluations predict a range of outcomes pertaining to romantic functioning, such as relationship satisfaction and relationship stability (LeBel & Campbell, 2009), reduced risk of breakup (Lee, Rogge, & Reis, 2010), more attachment security (Zayas & Shoda, 2005), daily relationship quality and positive relationship behaviors (LeBel & Campbell, 2013), lower rejection and selfish behavior toward and from the partner (Murray, Gomillion, et al., 2015), and more positive relationships expectations (Murray, Lupien, & Seery, 2012). Applied to the present question, these findings suggest different implications. Notably, the preliminary evidence of negativity biases obtained not only in deliberate evaluations of others, but also in spontaneous evaluations, suggests that individuals with high BP features may have difficulties modulating their evaluations positively and that this may contribute to their difficulties in a range of interpersonal outcomes that are associated with deliberate as well as spontaneous evaluations. This pattern of findings also implies that behavioral and interpersonal dysfunctions in individuals with BP features may be much engrained because they may be not only rooted in deliberate knowledge and beliefs, but

² Because of the problems highlighted above, it is important to keep in mind that the implications described in the next sections remain tentative until future research is able to clarify the nature of negativity biases in BPD and the existence of negativity biases at the spontaneous level of evaluations.

also in spontaneous affective gut reactions arising from less intentional processes. This would suggest that negativity biases in individuals with BP features may be ubiquitous, in line with clinical observation and empirical evidence suggesting that BPD is hard to treat in that it creates difficult reactions in the therapist, might complicate prognosis and necessitate a longer treatment and implies higher risks of hospitalisation (Gunderson et al., 2006; Zanarini et al., 1998; Zanarini, Frankenburg, Khera, & Bleichmar, 2001). A new direction for future research would be to investigate how spontaneous evaluations of others may influence relationship outcomes in individuals with BP traits or disorder. For instance, it would be interesting to examine whether negativity biases that operate at the spontaneous level of evaluations may explain the previously observed tendency of individuals with BPD to fail to create effective trusting interactions and to repair ruptures in trust (King-Casas et al., 2008). However, more research is needed to establish the existence and nature of negativity biases at the spontaneous and deliberate levels of evaluations in individuals with BP features. Nevertheless, the present thesis brings preliminary evidence that point to the importance of examining whether negativity biases may operate not only at the deliberate level, but also at the spontaneous level of evaluations. Future research could examine how to address both types of evaluations and how changes in spontaneous and deliberate evaluations may influence romantic functioning of individuals with BP features. In fact, given that spontaneous evaluations have a key influence on behavior beyond deliberate evaluations, it is important for research on BP features and BPD to not only examine deliberate evaluations of others, but to also examine spontaneous evaluations.

Changing spontaneous evaluations. To the extent that spontaneous and deliberate evaluations are guided by different processes and operate in different conditions (Gawronski et al., 2008; Gawronski & LeBel, 2008), it would make sense that the interventions required to change evaluations differ for the two kinds of evaluations. Although there is no research testing interventions to reduce negativity biases at the spontaneous level of evaluations in individuals with BP features or BPD, research on prejudice has yielded compelling evidence on this matter. This line of research has brought accumulating evidence that Black people face continuing discrimination and that discriminatory outcomes result in part from negativity biases at the spontaneous level of evaluations, a phenomenon termed implicit race bias. Devine, Forscher, Austin, and Cox (2012) reviewed evidence supporting these claims and

examined the role of five intervention strategies to reduce implicit race bias, including counter-stereotypical imaging (i.e., imagining in detail counterstereotypic examples), perspective taking (i.e., taking the perspective in the first person of the target of negative evaluations) and increasing contacts (i.e., seeking opportunities to encounter and engage in positive interactions with the person). Such strategies were most effective when people were aware of their tendency to discriminate and reported using the intervention strategies. Applied to the present question, these findings suggest that it may be possible to reduce negativity in spontaneous evaluations of others in individuals with BP features in several ways, such as (1) raising awareness of negative automatic gut reactions to others and of negative beliefs that are related to such reactions. In therapy, this could be done by first exploring the client's interpersonal experiences and dynamics in the present and then exploring past interpersonal experiences in such a way to identify the experiences and basic beliefs that underlie the person's dysfunctional reactions to others in the present (e.g., distrust, blame, distance, quarrelsomeness). The therapist may balance validation of the client's present experiences and reformulations of the client's experiences into more complex units of understanding in order to raise awareness while maintaining a good working alliance and being careful not to elicit too much defensiveness. The therapist may also attempt to (2) replace automatic negative thoughts by thoughts of previous positive interactions with the other person or by thoughts of positive experiences that contradict the automatically activated negative evaluation. For instance, negativity biases may lead individuals with BPD to selectively remember negative experiences and the therapist may highlight positive interpersonal experiences when they happen and question why such experiences are minimized. However, it may be difficult for a client with BPD to question its thoughts and experiences and the therapist must be careful not to repeat experiences of invalidation. With that respect, an important tool of therapy is to build a positive and repairing relationship that will make it easier for the client to form more positive and trusting beliefs. In that respect, it is important also to help the client experience positive relationships outside of therapy, as therapeutic progress does not happen in a vacuum. Hence, the therapy could help the client to (3) initiate positive contacts with the person who is evaluated negatively. For instance, in a romantic relationship, the therapist could help the client with BPD to identify interpersonally rewarding activities and gestures that the client could try with its partner and reinforce those behaviors. This places an active role on the client

not only to explore, accept and change psychological factors but also to take action into making its relationship more rewarding. Different interventions will be necessary to increase the willingness and capacity of the client with BPD to approach relationships in a positive and stable manner, such as increasing the capacity for empathy and perspective taking, which will in turn be helped by increasing emotion regulation strategies. Finally, the findings of Devine et al. (2012) point to the importance of (4) using the intervention strategies often enough in order to change spontaneous evaluations. However, given that these recommendations are based on research on prejudice, future research is needed to determine their effectiveness in reducing negative spontaneous evaluations of others in the context of BP features and BPD. In fact, borderline personality presents specific challenges. For instance, our results suggest that one challenge in individuals with BP features may be the tendency to show relatively more negative reactions even in neutral or positive situations.

Spontaneous evaluations, inhibition deficits and BPD. In future studies, the integration of research on spontaneous evaluations, associative-propositional processes and inhibition deficits in BPD may shed new light on the mechanisms of interpersonal dysfunctions in BPD. Different clinical models of BPD highlight the links between emotions, inhibition and BPD. Notably, Linehan's biosocial theory has thoroughly articulated BPD as arising from deficits in emotion regulation. According to Linehan's biosocial theory of BPD (1993), "BPD is primarily a disorder of emotion dysregulation and emerges from transactions between individuals with biological vulnerabilities and specific environmental influences" (cited in Crowell, Beauchaine, & Linehan, 2009, p. 496). In this model, emotion dysregulation refers to (a) heightened emotional sensitivity, (b) inability to regulate intense emotional responses, and (c) slow return to emotional baseline. The theory further suggests that BPD develops within and invalidating environment characterized by intolerance toward the expression of private emotional experiences. This early dysfunctional dynamic prevents the child from learning how to understand, label, regulate, or tolerate emotional responses; instead, the child learns to oscillate between emotional inhibition and extreme emotional lability. Crowell et al. (2009) extended Linehan's initial model (1993) and proposed that "early impulsivity is a predisposing vulnerability for a substantial subset of those who eventually meet criteria for BPD" (p. 496). In a similar vein, Kernberg et al. have conducted empirical research on neurobiological features of BPD (Lenzenweger, Clarkin, Fertuck, & Kernberg, 2004; Silbersweig et al., 2007)

and they updated their psychodynamic model by integrating in more depth the interactive role of biological features of BPD. The model posits a "dynamic interaction of temperament (individual differences in motor and emotional reactivity and self-regulation), a preponderance of negative affect, low effortful control, and an absence of a coherent sense of self and others" (Clarkin & Posner, 2005; cited in Silbersweig et al., 2007). With this in mind, the interaction of high negative affect and lack of control (which is similar to the notion of emotion regulation deficits) may potentially underlie the hypothesized tendency of individuals with BPD to form negative or split-off evaluations of others (i.e., switch between devaluation and idealization). Ultimately, the existence, nature and potential underlying factors of negativity biases and splitting are empirical questions that have to be investigated as such. The APE model may help extending these questions and the aforementioned clinical models of BPD even further by integrating the notion of spontaneous evaluations and by bringing new tools to conduce more thorough examinations of these questions empirically.

According to the APE model, spontaneous and deliberate evaluations result primarily from associative and propositional processes, respectively. Spontaneous evaluations represent affective gut reactions resulting from the activation of mental associations in memory; in contrast, deliberate evaluations represent the behavioral outcome of propositional processes involving the validation of activated information on the basis of logical consistency. With the APE, biosocial and psychodynamic models in mind, the combination of trait impulsivity and reactivity on the one hand and interpersonal adversity (i.e., invalidating environment) on the other hand may lead the child to encode negative interpersonal experiences in memory and develop or maintain difficulties in emotional regulation. As a result, upon new encounters, one may form negative spontaneous evaluations of other people based on few cues. Inhibition deficits may then be a predisposing trait and mechanism in BPD that may contribute to maintaining strong emotional responses and impulsive behaviors throughout the individual's development. In fact, inhibition deficits may make it difficult to counteract negative affective gut reactions toward other people (i.e., spontaneous evaluations) and to positively modulate negative deliberate evaluations reflecting consciously endorsed negative beliefs. These negative evaluations may lead the person to display negative behaviors (e.g., distance, hostility), fueling more negative behaviors from others. The negative responses of other people may in turn trigger negative interpersonal memories and negative emotional responses

in the individual with BPD, fueling negative interpersonal expectations and beliefs and making it more likely to form negative evaluations of other people in future encounters. With accumulating evidence of negative interactions and negative interpretation bias, the person with BPD may not only form negative spontaneous evaluations of others, but may also consciously endorse beliefs that other people should not be trusted (negative deliberate evaluations). Inhibition deficits may not only make it difficult for individuals with BPD to modulate their negative evaluations, but also to alter the influence of negative evaluations on their behavior with other people. In fact, Barrett et al. (2004) review evidence suggesting that low working memory capacity (which strongly overlaps with low inhibition) may compromise the capacity to introspect, regulate emotions, take other people's perspective and suppress emotions/cognitions/reactions. These factors may possibly mediate the relationship between inhibition deficits and negative evaluations in BPD. These propositions are theoretical as this is the first research examining spontaneous evaluations of others in relation to BP features. We believe that the integration of spontaneous evaluations, inhibition deficits and emotion processes in borderline personality may be a promising avenue for future inquiry.

Treatment implications. Based on the preliminary evidence that there may be negativity biases at the spontaneous and deliberate levels of evaluations in individuals with BP features, the treatment of these individuals may benefit from the therapist paying equal attention to what the patient deliberately says and thinks on the one hand and to how he spontaneously feels on the other hand. In doing so, the therapist may keep in mind that interventions that reduce negativity of deliberate thoughts might not reduce negativity of spontaneous feelings, and vice-versa. Based on evidence that spontaneous and deliberate partner evaluations both predict relationship quality and general well-being (LeBel & Campbell, 2013), it seems important that the therapy impacts both types of evaluations to increase relationship functioning of individuals with BP features or disorder. Toward this goal, the therapist may track how the interventions impact spontaneous and deliberate evaluations and make sure to address both kinds of evaluations. The therapist can also keep in mind factors that influence both types of evaluations and try to achieve change through these factors. For instance, say the therapist is able to help the client identify the reasons he/she stays with the partner, but that this does not lead to more positive and less negative spontaneous gut reactions toward the partner. In this case, the therapist could work with the client on building the conditions to

gradually experience positive emotional interactions with the partner. In fact, there is evidence that positive interactions (Devine et al., 2012) and lesser conflict early in the relationship (Murray, Holmes, & Pinkus, 2010) contribute to forming more positive spontaneous evaluations of others. Alternatively, say the client has positive feelings and experiences positive moments with its partner, but has negative thoughts and uncertainties about the partner. In this case, it is possible to help the client rely more on the positive gut feelings that shape its spontaneous evaluations in order to build more positive interactions and have more quality time in its relationship (for the protective role of positive spontaneous evaluations see Murray et al., 2012). Mindfulness may be useful in this case in order to short-circuit the influence of negative working thoughts or ruminations. That being said, these examples of interventions suggest that the partner is committed enough to the relationship to work towards being happier with its partner. However, this is not always the case and the therapy may also benefit from taking into account the ambivalence of the client. Nevertheless, given that the quality of current relationships of patients with BPD is a positive indicator for prognosis (K. L. Gratz, Tull, & Gunderson, 2008), the therapist can keep in mind that helping the client to increase relationship quality may help reducing other symptoms (e.g., anger; Hedrick & Berlin, 2012). In paying particular attention to helping individuals with BPD developing positive and pleasing interactions, the therapist also moves beyond interventions focusing solely on individual changes, which may possibly reduce feelings of inadequacy and increase perceptions of control in the client. This can potentially be done in couple therapy where it is possible to focus on dyadic change and relationship dynamic (Johnson, 2003), by identifying interactions where one feels safe and able to have quality time and by making time for positive social interactions. However, because BPD involves difficulties regulating emotions and strong feelings of insecurities, the therapist should dose psychoeducative interventions carefully in order to stay attuned to the more complex socio-affective processes that arise within the client's relationships both in and outside of therapy. Nevertheless, psychoeducation and mindfulness techniques are, amongst other concrete techniques, helpful in enhancing emotion regulation and reducing ineffective action tendencies linked with dysregulated emotions (e.g., dialectical behavior therapy, Linehan, 1987; Lynch, Chapman, Rosenthal, Kuo, & Linehan, 2006; McMMain et al., 2009). Finally, therapies that work on building positive interpersonal experiences outside and within therapy (e.g., Transference focused

psychotherapy, Kernberg, Yeomans, Clarkin, & Levy, 2008) may be particularly effective in building more positive spontaneous evaluations, given that spontaneous evaluations are shaped by previous emotional experiences rather than by conscious beliefs (Murray et al., 2010).

Implications of the divergent patterns of negativity biases. Patterns of negativity were found in certain conditions of Studies 1-2 and not in other conditions. Specifically, preliminary evidence of negativity biases was observed after the positive clip on the explicit measure in Study 1 and independent of experimental conditions on both the implicit and explicit measures in Study 2 (i.e., positive, negative and control conditions). As discussed earlier, the present thesis does not allow concluding whether this pattern of results reflects genuine evaluative processes in individuals with BP features or whether it is an artefact of the limits that we highlighted (e.g., low reliability of implicit measures). Nevertheless, if we explore the former scenario, one intriguing question to ask is why would individuals with high BP features display negativity biases after a positive situation rather than specifically after a negative situation? This result is surprising given the growing amount of evidence showing that signs of rejection are a central trigger of negative reactions in BPD in the affective, cognitive and behavioral domains (Ayduk et al., 2008; Berenson et al., 2011; Boldero et al., 2009; Kim L. Gratz, Dixon-Gordon, Breetz, & Tull, 2013; Miano et al., 2013; Staebler et al., 2011). Based on these previous studies and on clinical knowledge, it would make sense to expect that negative interpersonal scenarios (i.e., a movie clip depicting a verbal fight in Study 1 and imagining a rejecting partner in Study 2) would trigger fears of rejection in individuals with high BP features and in turn elicit negative evaluation biases. However, the overall pattern of results did not confirm this view. One possible explanation for these counter-intuitive findings is that individuals with high BP features have more negative evaluations at baseline. In other words, they may generally have negative feelings and thoughts about other people even in the absence of any apparent negative situation; in this context, it may be more difficult to negatively modulate evaluations that are already at a more negative level relative to control participants. This interpretation is in line with evidence showing that individuals with BPD exhibit heightened baseline intensity of physiological response and increased intensity of self-reported negative emotions at baseline, but that they show no sign of heightened reactivity compared to controls (Kuo & Linehan, 2009). Moreover, cognitive distortions among

individuals with high BP features may lead them to see signs of distrust in behaviors that would otherwise be interpreted positively or neutrally (for a study on cognitive distortions associated with BP features, see for example Geiger, Peters, & Baer, 2015). This possibility converges with evidence showing that individuals with BPD assigned more negative traits than controls to an individual presented in a neutral context (Barnow et al., 2009) and that BP features are associated with rating neutral faces as less trustworthy (Miano et al., 2013).

Implications for splitting. We found little evidence for splitting. In fact, although there was some evidence for instability in the first study, the result did not replicate across analyses and there was no evidence for polarity in the second study. This stands in contrast with clinical observation, which emphasizes the key role of instability and polarity in interpersonal and identity disturbances of individuals with BPD and is reflected in cognitive theories (A. T. Beck & Freeman, 1990; Aaron T. Beck et al., 2004) and psychodynamic theories (Kernberg, 1986) as well as in the DSM-5 (Criterion 2: “close relationships often viewed in extremes of idealization and devaluation”; Criterion 3: “unstable self-image”; American Psychiatric Association, 2013). In addition, the fact that we did not obtain clearer evidence for splitting contrasts with previous empirical studies. Indeed, past research has obtained empirical support for the view that BPD involves splitted evaluations of self and others, as evidenced by polarized relationship experiences (Coifman et al., 2012), less integration in self evaluations, less consistency in evaluations of self and others (Beeney, Hallquist, Ellison, & Levy, 2016), compartmentalized self-concept structure (Vater et al., 2015) and greater diffusion of positive self evaluations and greater interconnection of negative self evaluations (Evans et al., 2015). The lack of evidence for splitting also diverges with studies that measured real-life variations in affect in individuals with BPD and observed affective instability (Ebner-Priemer et al., 2015; Henry et al., 2001; Koenigsberg et al., 2002; Myin-Germeys et al., 2009; Santangelo et al., 2012; Trull et al., 2008); in fact, sudden, large decreases from a very positive mood are consistent with the notion of splitting and with criteria 2 and 3 of DSM-5. Given evidence for affective instability, one might expect to find a similar pattern of instability in the way individuals with BPD evaluate self and others. In fact, Kernberg (1986) argues that affect and evaluations of self and others are intertwined when splitting occurs such that extreme switches in affective valence (e.g., all-negative affects) is accompanied by corresponding extreme switches in evaluations of self and other (e.g., all-negative evaluations of self and other). Yet,

despite evidence for such switches in affective valence, there is no direct empirical evidence supporting the notion that BP features also involves switches between all-negative and all-positive evaluations of self and others. Although Coifman et al. (2012) measured relationship evaluations at different times, their statistical procedure did not aim to capture instability across time. Instead, it focused on the polarity of the positive and negative valence in BPD. In addition, Arntz and al. tested the splitting hypothesis in several laboratory studies and failed to find supporting evidence (Arntz & Haaf, 2012; Arntz & Veen, 2001; Sieswerda et al., 2005; Sieswerda et al., 2013; Veen & Arntz, 2000). However, given that this study mostly relied on movie clips in the laboratory (as the first study of this thesis), it is possible that the lack of support for splitting is method-dependent. Nevertheless, despite a lack of direct support for instability in evaluations of others or for splitting in some studies, other studies did find support for splitting, as evidenced by greater polarization and compartmentalization in evaluations of self and others in individuals with BP features or BPD (Beeney et al., 2016; Coifman et al., 2012; Evans et al., 2015; Vater et al., 2015). In sum, with respect to splitting, previous research suggests that BPD indeed involves extreme evaluations and difficulty integrating positive and negative evaluations of self and others, but there is no support for the view that it involves switches from negative to positive and from positive to negative evaluations of self and others, as noted in DSM-5. In order to examine such switches, research designs must include different measurement times. For instance, one study found evidence for greater instability in evaluations of self and others in BPD across a 3-hour period (Beeney et al., 2016), but examined instability of evaluations in general without focusing specifically on switches between positive and negative evaluations, as described in the splitting hypothesis (Kernberg, 1986) and in Criterion 2 of BPD in DSM-5 (American Psychiatric Association, 2013, pp. 766). Coifman et al. (2012)'s statistical procedure specifically examined the relationship between the two valenced dimensions, that is, they estimated the degree of association or polarization between negative and positive affect as well as between negative and positive relational experiences. They found that individuals with BPD had significantly greater polarity in reports of affective and relational experience across a 21-day diary period. However, notwithstanding the contribution of this method to examine the polarization of positive and negative experiences, Coifman et al. (2012)'s statistical procedure was not designed to examine switches from negative to positive experiences at different measurement

times. Finally, Arntz & al.'s procedure using movie clips did not calculate switches between different measurement times (Sieswerda et al., 2013). Hence, no study to date has specifically found evidence that BPD involves switches from negative to positive evaluations of self and others. For this purpose, our Study 1 included two measurement times (i.e. one measurement following the positive movie clip and one measurement following the negative movie clip), but again we did not find evidence for instability. Considering the amount of theorizing on instability in BPD, we argue that more research should be conducted to specifically examine whether BPD involves instability in evaluations of self and others and to distinguish between and simultaneously test the notions of negativity biases and splitting. A promising avenue for future exploration would be to examine whether instability is best characterized by evaluative switches from neutral to more negative evaluations of self and others in BPD (which would converge with the negativity bias hypothesis) or whether such switches indeed involve highly positive evaluations (idealization) that alternate with highly negative evaluations (devaluation), consistent with the splitting hypothesis. To test the splitting hypothesis, two complementary approaches can be used in future studies: (1) laboratory studies with multiple measurement times (e.g., ABA design) and (2) field-based approaches that assess within- and between-days variations in evaluations, affects and interpersonal disturbances as they occur in real-life interactions of individuals with BPD or BP features. Moreover, given that psychodynamic theories insist on the role of unconscious representations (Kernberg, 1986), such research should not be limited to self-report measures and should also include implicit measures.

Implications of the different patterns of negativity biases. It is important to also note that the obtained patterns of negativity biases differed between Studies 1-2 as well as between the conditions of Study 1. In fact, although both studies found evidence for negativity biases, Study 1 found *context-specific* negativity biases in individuals with high BP features, whereas Study 2 found *generalized* negativity biases in individuals with high BP features. Moreover, Study 1 found different patterns of negativity biases between the two conditions of the study. Because Studies 1-2 examined different interpersonal contexts, namely impression formation (Study 1) versus partner evaluation (Study 2), such inconsistencies could be attributable to genuine differences in the way BP features influence evaluations of others depending on the context, that is, impression formation versus partner evaluations. Alternatively, these

inconsistencies could reflect methodological differences between the two studies rather than genuine differences between people with low vs. high BP features. Notably, the use of different implicit measures in Studies 1-2 may have potentially contributed to inconsistencies in the findings. Specifically, although the name-letter task and the evaluative priming task are considered implicit measures because the impact of evaluations on participants' responses is deemed unintentional, resource-independent, unconscious, or uncontrollable, they still differ on the procedural level. In fact, the evaluative priming task captures the extent that the presentation of a given prime stimulus facilitates quick responses to positive/negative target words. In contrast, the name-letter task does not rely on time pressure or response facilitation; instead, it involves rating how aesthetically pleasing each letter of the alphabet is, based on the principle that individuals prefer partner's and self's initials. Based on these procedural differences, whether the same processes influence both tasks is an empirical question that needs to be examined as such. The picture becomes even more complex when one considers that negativity biases emerged in different conditions for spontaneous and deliberate evaluations in Study 1. Specifically, for spontaneous evaluations the negativity bias emerged only after watching the negative clip, whereas for deliberate evaluations the negativity bias emerged only after watching the positive clip first. As we previously described, this pattern of results could suggest that negative evaluation biases in individuals with BP features might emerge in different situations depending on which kind of evaluations is at play (spontaneous or deliberate). However, the pattern of negativity obtained in Study 2 does not confirm this claim. In fact, in Study 2, the effects of BP features on partner evaluations were unqualified by the conditions in the study and this was true for the two kinds of partner evaluations (i.e., spontaneous and deliberate). Moreover, given that BPD has been shown to involve greater reactivity of negative affect to negative perceptions (Berenson et al., 2011; Sieswerda et al., 2013), it would have made sense to find a pattern where the negativity bias would be strongest after the negative clip. However, past research that used movie clips disconfirms this claim, showing that negative evaluations of caring characters in an intimate context particularly characterized BPD (Sieswerda et al., 2013). An alternative interpretation of the different patterns of negativity biases is that low reliability estimates of implicit measures might have hindered the replicability of the results between and within our studies. In fact, using mathematical simulations, Lebel and Paunonen (2011) showed that the use of implicit

measures with low reliability as dependent variables in experimental studies is associated with substantially lower levels of replicability. In fact, reliability estimates are typically low in the evaluative priming task (Herring et al., 2013) and in the name-letter task (LeBel & Gawronski, 2009), which were used as dependent measures in Study 1 and Study 2, respectively. “Lower levels of reliability are associated with decreasing probabilities of detecting a statistically significant effect, given one exists in the population” (Sutcliffe, 1958, cited in Lebel and Paunonen, 2011, p. 573). In short, the association between low reliability and low replicability could explain why the results that we obtained with the implicit measures emerged only in some conditions of our studies and not in other conditions. Finally, given that the main conclusion pertaining to spontaneous evaluations relies on a marginal three-way interaction in Study 1 and that the expected main effects of experimental conditions were not observed in Study 2, the results obtained in the present thesis may not accurately reflect how BP features influence evaluations of others. More research is needed to examine replicability issues and to determine whether, why and when individuals with BP features show negativity biases and/or splitting. Toward this goal, complementary designs and measures should be combined. In fact, a limit of the BPD literature is the lack of integration of implicit measures with ecological designs that involve personally relevant emotional material. Implicit measures can and have been used in real-life settings to examine variations in spontaneous self evaluations, but such studies remain scarce (e.g., DeHart & Pelham, 2007). Past BPD studies conducted in the laboratory have used implicit measures more extensively, but they relied for a large part on stimuli with little personal significance or ecological validity (e.g., words). The present thesis addressed these shortcomings by including an implicit measure as well as vivid emotional stimuli that echo real-life interactions, that is, a character presented in realistic interpersonal movie clips (Study 1) and partners (Study 2). Our studies yielded preliminary evidence that BP features may involve negativity biases at the spontaneous and deliberate levels of evaluations. Although these results remain provisional, they nonetheless suggest that future studies may gain deeper insights into the mechanisms underlying interpersonal difficulties in individuals with BP features by examining the interplay between spontaneous and deliberate evaluations. We hope that future research will help clarify the existence and nature of negativity biases and splitting in individuals with BP features and BPD and examine their

implications and antecedents, with a specific focus on spontaneous and deliberate evaluations and complementary research designs.

Références citées dans l'introduction et dans la conclusion

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Abdul-Hamid, S., Denman, C., & Dudas, R. B. (2014). Self-relevant disgust and self-harm urges in patients with borderline personality disorder and depression: a pilot study with a newly designed psychological challenge. *Plos One*, *9*(6). doi: 10.1371/journal.pone.0099696
- Ackerman, S. J., Clemence, A. J., Weatherill, R., & Hilsenroth, M. J. (1999). Use of the TAT in the assessment of DSM-IV cluster B personality disorders. *Journal of Personality Assessment*, *73*(3), 422-448. doi: 10.1207/s15327752jpa7303_9
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Amir, N., McNally, R. J., Riemann, B. C., Burns, J., Lorenz, M., & Mullen, J. T. (1996). Suppression of the emotional Stroop effect by increased anxiety in patients with social phobia. *Behaviour Research and Therapy*, *34*(11-12), 945-948. doi: 10.1016/s0005-7967(96)00054-x
- Arntz, A., Appels, C., & Sieswerda, S. (2000). Hypervigilance in borderline disorder: A test with the emotional Stroop paradigm. *Journal of Personality Disorders*, *14*(4), 366-373.
- Arntz, A., Dreessen, L., Schouten, E., & Weertman, A. (2004). Beliefs in personality disorders: a test with the personality disorder belief questionnaire. *Behaviour Research and Therapy*, *42*(10), 1215-1225. doi: 10.1016/j.brat.2003.08.004
- Arntz, A., & Haaf, J. T. (2012). Social cognition in borderline personality disorder: evidence for dichotomous thinking but no evidence for less complex attributions. *Behaviour research and therapy*, *50*(11), 707-718. doi: 10.1016/j.brat.2012.07.002
- Arntz, A., & Veen, G. (2001). Evaluations of others by borderline patients. *Journal of Nervous and Mental Disease*, *189*(8), 513-521. doi: 10.1097/00005053-200108000-00004
- Arntz, A., Weertman, A., & Salet, S. (2011). Interpretation bias in Cluster-C and borderline personality disorders. *Behaviour Research and Therapy*, *49*(8), 472-481. doi: 10.1016/j.brat.2011.05.002
- Asmundson, G. J. G., & Stein, M. B. (1994). Selective processing of social threat in patients with generalized social phobia - evaluation using a dot-probe paradigm. *Journal of Anxiety Disorders*, *8*(2), 107-117. doi: 10.1016/0887-6185(94)90009-4
- Ayduk, O., Zayas, V., Downey, G., Cole, A. B., Shoda, Y., & Mischel, W. (2008). Rejection sensitivity and executive control: joint predictors of borderline personality features. *Journal of Research in Personality*, *42*(1), 151-168. doi: 10.1016/j.jrp.2007.04.002
- Baccus, J. R., Baldwin, M. W., & Packer, D. J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science*, *15*, 498-502.
- Bagge, C., Nickell, A., Stepp, S., Durrett, C., Jackson, K., & Trull, T. J. (2004). Borderline personality disorder features predict negative outcomes 2 years later. *Journal of Abnormal Psychology*, *113*(2), 279-288. doi: 10.1037/0021-843x.113.2.279
- Bargh, J. A. (1994). The four horsemen of automaticity: awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1-40). Hillsdale, NJ: Erlbaum.

- Barnow, S., Stopsack, M., Grabe, H. J., Meinke, C., Spitzer, C., Kronmueller, K., & Sieswerda, S. (2009). Interpersonal evaluation bias in borderline personality disorder. *Behaviour Research and Therapy*, *47*(5), 359-365. doi: 10.1016/j.brat.2009.02.003
- Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, *130*(4), 553-573. doi: 10.1037/0033-2909.130.4.553
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252-1265. doi: 10.1037/0022-3514.74.5.1252
- Beck, A. T., & Freeman, A. M. (1990). *Cognitive therapy of personality disorders*. New York, NY: Guilford Press.
- Beck, Aaron T., Freeman, Arthur, & Davis, Denise D. (2004). *Cognitive therapy of personality disorders (2nd ed.)*. New York, NY: Guilford Press; US.
- Beeney, Joseph E., Hallquist, Michael N., Ellison, William D., & Levy, Kenneth N. (2016). Self-other disturbance in borderline personality disorder: Neural, self-report, and performance-based evidence. *Personality disorders*, *7*(1), 28-39. doi: 10.1037/per0000127
- Berenson, K. R., Downey, G., Rafaeli, E., Coifman, K. G., & Paquin, N. L. (2011). The rejection-rage contingency in borderline personality disorder. *Journal of Abnormal Psychology*, *120*(3), 681-690. doi: 10.1037/a0023335
- Bhatia, V., Davila, J., Eubanks-Carter, C., & Burckell, L. A. (2013). Appraisals of Daily Romantic Relationship Experiences in Individuals With Borderline Personality Disorder Features. *Journal of Family Psychology*, *27*(3), 518-524. doi: 10.1037/a0032870
- Blackburn, R., Donnelly, J. P., Logan, C., & Renwick, S. J. D. (2004). Convergent and discriminative validity of interview and questionnaire measures of personality disorder in mentally disordered offenders: A multitrait-multimethod analysis using confirmatory factor analysis. *Journal of Personality Disorders*, *18*(2), 129-150. doi: 10.1521/pedi.18.2.129.32779
- Boldero, J. M., Hulbert, C. A., Bloom, L., Cooper, J., Gilbert, F., Mooney, J. L., & Salinger, J. (2009). Rejection sensitivity and negative self-beliefs as mediators of associations between the number of borderline personality disorder features and self-reported adult attachment. *Personality and Mental Health*, *3*, 248-262.
- Bornstein, R. F. (2002). A process dissociation approach to objective-projective test score interrelationships. *Journal of Personality Assessment*, *78*(1), 47-68. doi: 10.1207/s15327752jpa7801_04
- Bouchard, S., Godbout, N., & Sabourin, S. (2009). Sexual attitudes and activities in women with borderline personality disorder involved in romantic relationships. *Journal of Sex & Marital Therapy*, *35*(2), 106-121. doi: 10.1080/00926230802712301
- Bouchard, S., Sabourin, S., Lussier, Y., & Villeneuve, E. (2009). Relationship quality and stability in couples when one partner suffers from borderline personality disorder. *Journal of Marital and Family Therapy*, *35*(4), 446-455. doi: 10.1111/j.1752-0606.2009.00151.x
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49-59. doi: 10.1016/0005-7916(94)90063-9

- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6, 3-5.
- Cavallo, J. V., Holmes, J. G., Fitzsimons, G. M., Murray, S. L., & Wood, J. V. (2012). Managing motivational conflict: how self-esteem and executive resources influence self-regulatory responses to risk. *Journal of Personality and Social Psychology*, 103(3), 430-451. doi: 10.1037/a0028821
- Clarkin, J. F., & Posner, M. (2005). Defining the mechanisms of borderline personality disorder. *Psychopathology*, 38(2), 56-63. doi: 10.1159/000084812
- Coifman, K. G., Berenson, K. R., Rafaeli, E., & Downey, G. (2012). From negative to positive and back again: polarized affective and relational experience in borderline personality disorder. *Journal of Abnormal Psychology*, 121(3), 668-679. doi: 10.1037/a0028502
- Crowell, S. E., Beauchaine, T. P., & Linehan, M. M. (2009). A biosocial developmental model of borderline personality: Elaborating and extending Linehan's theory. *Psychological Bulletin*, 135(3), 495-510. doi: 10.1037/a0015616
- Daley, S. E., Burge, D., & Hammen, C. (2000). Borderline personality disorder symptoms as predictors of 4-year romantic relationship dysfunction in young women: addressing issues of specificity. *Journal of Abnormal Psychology*, 109(3), 451-460. doi: 10.1037//0021-843x.109.3.451
- Dammann, G., Huegli, C., Selinger, J., Gremaud-Heitz, D., Sollberger, D., Wiesbeck, G. A., . . . Walter, M. (2011). The self-image in borderline personality disorder: an in-depth qualitative research study. *Journal of Personality Disorders*, 25(4), 517-527.
- Dannowski, U., Kersting, A., Lalee-Mentzel, J., Donges, U. S., Arolt, V., & Suslow, T. (2006). Subliminal affective priming in clinical depression and comorbid anxiety: a longitudinal investigation. *Psychiatry Research*, 143(1), 63-75. doi: 10.1016/j.psychres.2005.08.022
- De Houwer, J. (in press). Why a propositional single-process model of associative learning deserves to be defended. In J. W. Sherman, B. Gawronski & Y. Trope (Eds.), *Dual processes in social psychology*. NY: Guilford.
- DeHart, Tracy, & Pelham, Brett W. (2007). Fluctuations in state implicit self-esteem in response to daily negative events. *Journal of Experimental Social Psychology*, 43(1), 157-165. doi: 10.1016/j.jesp.2006.01.002
- Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, 48(6), 1267-1278. doi: 10.1016/j.jesp.2012.06.003
- Domes, G., Schulze, L., & Herpertz, S. C. (2009). Emotion recognition in borderline personality disorder-a review of the literature. *Journal of Personality Disorders*, 23(1), 6-19.
- Domes, G., Winter, B., Schnell, K., Vohs, K., Fast, K., & Herpertz, S. C. (2006). The influence of emotions on inhibitory functioning in borderline personality disorder. *Psychological Medicine*, 36(8), 1163-1172. doi: 10.1017/s0033291706007756
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82(1), 62-68. doi: 10.1037//0022-3514.82.1.62

- Dozois, D. J. A., & Dobson, K. S. (2001). Information processing and cognitive organization in unipolar depression: Specificity and comorbidity issues. *Journal of Abnormal Psychology, 110*(2), 236-246. doi: 10.1037//0021-843x.110.2.236
- Dreessen, L., & Arntz, A. (1995a). *Personality Disorder Belief Questionnaire (PDBQ) (short version)*. The Netherlands: Maastricht University.
- Dreessen, L., & Arntz, A. (1995b). *The personality disorder beliefs questionnaire*. Maastricht.
- Ebner-Priemer, U. W., Houben, M., Santangelo, P., Kleindienst, N., Tuerlinckx, F., Oravecz, Z., . . . Kuppens, P. (2015). Unraveling affective dysregulation in borderline personality disorder: a theoretical model and empirical evidence. *Journal of Abnormal Psychology, 124*(1), 186-198. doi: 10.1037/abn0000021
- Edens, J. F., Marcus, D. K., & Ruiz, M. A. (2008). Taxometric analyses of borderline personality features in a large-scale male and female offender sample. *Journal of Abnormal Psychology, 117*(3), 705-711. doi: 10.1037/0021-843x.117.3.70.5
- Evans, D., Dalgleish, T., Dudas, R. B., Denman, C., Howard, M., & Dunn, B. D. (2015). Examining the shared and unique features of self-concept content and structure in borderline personality disorder and depression. *Cognitive Therapy and Research, 39*(5), 613-626. doi: 10.1007/s10608-015-9695-3
- Fales, Christina L., Barch, Deanna M., Rundle, Melissa M., Mintun, Mark A., Snyder, Abraham Z., Cohen, Jonathan D., . . . Sheline, Yvette I. (2008). Altered emotional interference processing in affective and cognitive-control brain circuitry in major depression. *Biological Psychiatry, 63*(4), 377-384. doi: 10.1016/j.biopsych.2007.06.012
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial-attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology, 69*(6), 1013-1027. doi: 10.1037//0022-3514.69.6.1013
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: their meaning and use. *Annual Review of Psychology, 54*, 297-327. doi: 10.1146/annurev.psych.54.101601.145225
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*(2), 229-238. doi: 10.1037//0022-3514.50.2.229
- Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. *Dual-Process Theories in Social Psychology, 97*-116.
- Gardner, K., & Qualter, P. (2009). Reliability and validity of three screening measures of borderline personality disorder in a nonclinical population. *Personality and Individual Differences, 46*(5-6), 636-641.
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692-731. doi: <http://dx.doi.org/10.1037/0033-2909.132.5.692>
- Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: theory, evidence, and open questions. *Advances in Experimental Social Psychology, 44*, 59-127. doi: 10.1016/b978-0-12-385522-0.00002-0
- Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in*

- social and personality psychology* (2nd ed., pp. 283-310). New York, NY: Cambridge University Press.
- Gawronski, B., Deutsch, R., & Banse, R. (2011). Response interference tasks as indirect measures of automatic associations. In K. C. Klauer, A. Voss & C. Stahl (Eds.), *Cognitive methods in social psychology* (pp. 78-123). New York, NY: Guilford Press.
- Gawronski, B., Deutsch, R., Mbirikou, S., Selbt, B., & Strack, F. (2008). When "Just Say No" is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology, 44*(2), 370-377. doi: 10.1016/j.jesp.2006.12.004
- Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: associations influence the construal of individuating information. *European Journal of Social Psychology, 33*(5), 573-589. doi: 10.1002/ejsp.166
- Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology, 44*(5), 1355-1361. doi: 10.1016/j.jesp.2008.04.005
- Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of implicit social cognition: measurement, theory, and applications*. New York, NY: Guilford Press.
- Geiger, P. J., Peters, J. R., & Baer, R. A. (2015). Cognitive distortions under cognitive load mediate the relationship between severity of borderline personality features and goal-interfering emotion dysregulation. *Journal of Experimental Psychopathology, 6*(4), 389-397. doi: 10.5127/jep.040413
- Graham, S. M., & Clark, M. S. (2006). Self-esteem and organization of valenced information about others: The "Jekyll and Hyde"-ing of relationship partners. *Journal of Personality and Social Psychology, 90*(4), 652-665. doi: 10.1037/0022-3514.90.4.652
- Gratz, K. L., Tull, M. T., & Gunderson, J. G. (2008). Preliminary data on the relationship between anxiety sensitivity and borderline personality disorder: The role of experiential avoidance. *Journal of Psychiatric Research, 42*(7), 550-559. doi: 10.1016/j.jpsychires.2007.05.011
- Gratz, Kim L., Dixon-Gordon, K. L., Breetz, A., & Tull, M. T. (2013). A laboratory-based examination of responses to social rejection in borderline personality disorder: The mediating role of emotion dysregulation. *Journal of Personality Disorders, 27*(2), 157-171.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480. doi: 10.1037/0022-3514.74.6.1464
- Greenwald, Anthony G., Poehlman, T. Andrew, Uhlmann, Eric Luis, & Banaji, Mahzarin R. (2009). Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity. *Journal of Personality and Social Psychology, 97*(1), 17-41. doi: 10.1037/a0015575
- Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology, 90*(1), 1-20. doi: 10.1037/0022-3514.90.1.1
- Gunderson, J. G., Bender, D., Sanislow, C., Yen, S., Rettew, J. B., Dolan-Sewell, R., . . . Skodol, A. E. (2003). Plausibility and possible determinants of sudden "remissions" in borderline patients. *Psychiatry-Interpersonal and Biological Processes, 66*(2), 111-119. doi: 10.1521/psyc.66.2.111.20614

- Gunderson, J. G., Daversa, M. T., Grilo, C. M., McGlashan, T. H., Zanarini, M. C., Shea, M. T., . . . Stout, R. L. (2006). Predictors of 2-year outcome for patients with borderline personality disorder. *American Journal of Psychiatry*, *163*(5), 822-826. doi: 10.1176/appi.ajp.163.5.822
- Hagenhoff, M., Franzen, N., Gerstner, L., Koppe, G., Sarnmer, G., Netter, P., . . . Lis, S. (2013). Reduced sensitivity to emotional facial expressions in borderline personality disorder: effects of emotional valence and intensity. *Journal of Personality Disorders*, *27*(1), 19-35.
- Haslam, N. (2003). Categorical versus dimensional models of mental disorder: the taxometric evidence. *Australian and New Zealand Journal of Psychiatry*, *37*(6), 696-704. doi: 10.1080/j.1440-1614.2003.01258.x
- Hawes, D. J., Helyer, R., Herlianto, E. C., & Willing, J. (2013). Borderline personality features and implicit shame-prone self-concept in middle childhood and early adolescence. *Journal of Clinical Child and Adolescent Psychology*, *42*(3), 302-308. doi: 10.1080/15374416.2012.723264
- Hedrick, A. N., & Berlin, H. A. (2012). Implicit self-esteem in borderline personality and depersonalization disorder. *Frontiers in Psychology*, *3*. doi: 10.3389/fpsyg.2012.00091
- Henry, C., Mitropoulou, V., New, A. S., Koenigsberg, H. W., Silverman, J., & Siever, L. J. (2001). Affective instability and impulsivity in borderline personality and bipolar II disorders: similarities and differences. *Journal of Psychiatric Research*, *35*(6), 307-312. doi: 10.1016/s0022-3956(01)00038-3
- Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., . . . Crites, S. L. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin*, *139*(5), 1062-1089. doi: 10.1037/a0031309
- Hill, J., Pilkonis, P., Morse, J., Feske, U., Reynolds, S., Hope, H., . . . Broyden, N. (2008). Social domain dysfunction and disorganization in borderline personality disorder. *Psychological Medicine*, *38*(1), 135-146. doi: 10.1017/s0033291707001626
- Hill, J., Stepp, S. D., Wan, M. W., Hope, H., Morse, J. Q., Steele, M., . . . Pilkonis, P. A. (2011). Attachment, borderline personality, and romantic relationship dysfunction. *Journal of Personality Disorders*, *25*(6), 789-805.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*(10), 1369-1385. doi: 10.1177/0146167205275613
- Hurlbert, D. F., Apt, C., & White, L. C. (1992). An empirical-examination into the sexuality of women with borderline personality-disorder. *Journal of Sex & Marital Therapy*, *18*(3), 231-242.
- Hyer, S. E. (1994). *Personality Diagnostic Questionnaire-4*. New York: New York State Psychiatric Institute.
- Hyer, S. E., Skodol, A. E., Kellman, H. D., Oldham, J. M., & Rosnick, L. (1990). Validity of the personality diagnostic questionnaire revised - comparison with 2 structured interviews. *American Journal of Psychiatry*, *147*(8), 1043-1048.
- Jackson, K. M., & Trull, T. J. (2001). The factor structure of the personality assessment inventory-borderline features (PAI-BOR) scale in a nonclinical sample. *Journal of Personality Disorders*, *15*(6), 536-545. doi: 10.1521/pedi.15.6.536.19187

- Johnson, S. M. (2003). The revolution in couple therapy: A practitioner-scientist perspective. *Journal of Marital and Family Therapy*, 29(3), 365-384.
- Kernberg, O. F. (1986). *Severe personality disorders: psychotherapeutic strategies*. New Haven, CT: Yale University Press.
- Kernberg, O. F., Yeomans, F. E., Clarkin, J. F., & Levy, K. N. (2008). Transference focused psychotherapy: Overview and update. *International Journal of Psychoanalysis*, 89(3), 601-620. doi: 10.1111/j.1745-8315.2008.00046.x
- King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., & Montague, P. R. (2008). The rupture and repair of cooperation in borderline personality disorder. *Science*, 321(5890), 806-810. doi: 10.1126/science.1156902
- Klein, S. R., Renshaw, K. D., & Curby, T. W. (2016). Emotion regulation and perceptions of hostile and constructive criticism in romantic relationships. *Behavior Therapy*, 47(2), 143-154.
- Koenigsberg, H. W., Harvey, P. D., Mitropoulou, V., Schmeidler, J., New, A. S., Goodman, M., . . . Siever, L. J. (2002). Characterizing affective instability in borderline personality disorder. *American Journal of Psychiatry*, 159(5), 784-788. doi: 10.1176/appi.ajp.159.5.784
- Kopala-Sibley, D. C., Zuroff, D. C., Russell, J. J., Moskowitz, D. S., & Paris, J. (2012). Understanding heterogeneity in borderline personality disorder: differences in affective reactivity explained by the traits of dependency and self-criticism. *Journal of Abnormal Psychology*, 121(3), 680-691. doi: 10.1037/a0028513
- Krause, S., Back, M. D., Egloff, B., & Schmukle, S. C. (2014). Implicit Interpersonal Attraction in Small Groups: Automatically Activated Evaluations Predict Actual Behavior Toward Social Partners. *Social Psychological and Personality Science*, 5(6), 671-679. doi: 10.1177/1948550613517723
- Krause-Utz, A., Oei, N. Y. L., Niedtfeld, I., Bohus, M., Spinhoven, P., Schmahl, C., & Elzinga, B. M. (2012). Influence of emotional distraction on working memory performance in borderline personality disorder. *Psychological Medicine*, 42(10), 2181-2192. doi: 10.1017/s0033291712000153
- Kuhlken, K., Robertson, C., Benson, J., & Nelson-Gray, R. (2014). The interaction of borderline personality disorder symptoms and relationship satisfaction in predicting affect. *Personality Disorders-Theory Research and Treatment*, 5(1), 20-25. doi: 10.1037/per0000013
- Kuo, J. R., & Linehan, M. A. (2009). Disentangling emotion processes in borderline personality disorder: Physiological and self-reported assessment of biological vulnerability, baseline intensity, and reactivity to emotionally evocative stimuli. *Journal of Abnormal Psychology*, 118(3), 531-544. doi: 10.1037/a0016392
- LeBel, E. P., & Campbell, L. (2009). Implicit partner affect, relationship satisfaction, and the prediction of romantic breakup. *Journal of Experimental Social Psychology*, 45(6), 1291-1294. doi: 10.1016/j.jesp.2009.07.003
- LeBel, E. P., & Campbell, L. (2013). The interactive role of implicit and explicit partner evaluations on ongoing affective and behavioral romantic realities. *Social Psychological and Personality Science*, 4(2), 167-174. doi: 10.1177/1948550612448196
- LeBel, E. P., Campbell, L., & Loving, T. J. (in press). Benefits of open and high-powered research outweigh costs. *Journal of Personality and Social Psychology*.

- LeBel, E. P., & Gawronski, B. (2009). How to find what's in a name: Scrutinizing the optimality of five scoring algorithms for the name-letter task. *European Journal of Personality, 23*, 85–106.
- LeBel, E. P., & Paunonen, S. V. (2011). Sexy but often unreliable: The impact of unreliability on the replicability of experimental findings with implicit measures. *Personality and Social Psychology Bulletin, 37*(4), 570-583. doi: 10.1177/0146167211400619
- Lee, S., Rogge, R. D., & Reis, H. T. (2010). Assessing the seeds of relationship decay: using implicit evaluations to detect the early stages of disillusionment. *Psychological Science, 21*(6), 857-864. doi: 10.1177/0956797610371342
- Leichsenring, F., Leibing, E., Kruse, J., New, A. S., & Leweke, F. (2011). Borderline personality disorder. *Lancet, 377*(9759), 74-84. doi: 10.1016/s0140-6736(10)61422-5
- Lenzenweger, M. F., Clarkin, J. F., Fertuck, E. A., & Kernberg, O. F. (2004). Executive neurocognitive functioning and neurobehavioral systems indicators in borderline personality disorder: A preliminary study. *Journal of Personality Disorders, 18*(5), 421-438. doi: 10.1521/pedi.18.5.421.51323
- Lieb, K., Zanarini, M. C., Schmahl, C., Linehan, M. M., & Bohus, M. (2004). Borderline personality disorder. *Lancet, 364*(9432), 453-461. doi: 10.1016/s0140-6736(04)16770-6
- Linehan, M. M. (1987). Dialectical behavior-therapy for borderline personality-disorder - Theory and method. *Bulletin of the Menninger Clinic, 51*(3), 261-276.
- Lobbestael, J., & Arntz, A. (2010). Emotional, cognitive and physiological correlates of abuse-related stress in borderline and antisocial personality disorder. *Behaviour Research and Therapy, 48*(2), 116-124. doi: 10.1016/j.brat.2009.09.015
- Lynch, T. R., Chapman, A. L., Rosenthal, M. Z., Kuo, J. R., & Linehan, M. M. (2006). Mechanisms of change in dialectical behavior therapy: Theoretical and empirical observations. *Journal of Clinical Psychology, 62*(4), 459-480. doi: 10.1002/jclp.20243
- Matzke, B., Herpertz, S. C., Berger, C., Fleischer, M., & Domes, G. (2014). Facial reactions during emotion recognition in borderline personality disorder: a facial electromyography study. *Psychopathology, 47*(2), 101-110. doi: 10.1159/000351122
- McMain, S. F., Links, P. S., Gnam, W. H., Guimond, T., Cardish, R. J., Korman, L., & Streiner, D. L. (2009). A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *American Journal of Psychiatry, 166*(12), 1365-1374. doi: 10.1176/appi.ajp.2009.09010039
- Miano, A., Fertuck, E. A., Arntz, A., & Stanley, B. (2013). Rejection sensitivity is a mediator between borderline personality disorder features and facial trust appraisal. *Journal of Personality Disorders, 27*(4), 442-456.
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297-326. doi: 10.1037/0033-2909.132.2.297
- Morey, L. C. . (1991). *The personality assessment inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Murray, S. L., Bellavia, G. M., Rose, P., & Griffin, D. W. (2003). Once hurt, twice hurtful: How perceived regard regulates daily marital interactions. *Journal of Personality and Social Psychology, 84*(1), 126–147.
- Murray, S. L., Derrick, J. L., Leder, S., & Holmes, J. G. (2008). Balancing connectedness and self-protection goals in close relationships: A levels-of-processing perspective on risk

- regulation. *Journal of Personality and Social Psychology*, *94*(3), 429-459. doi: 10.1037/0022-3514.94.3.429
- Murray, S. L., Gomillion, S., Holmes, J. G., & Harris, B. (2015). Inhibiting Self-Protection in Romantic Relationships: Automatic Partner Attitudes as a Resource for Low Self-Esteem People. *Social Psychological and Personality Science*, *6*(2), 173-182. doi: 10.1177/1948550614549386
- Murray, S. L., Gomillion, S., Holmes, J. G., Harris, B., & Lamarche, V. (2013). The Dynamics of Relationship Promotion: Controlling the Automatic Inclination to Trust. *Journal of Personality and Social Psychology*, *104*(2), 305-334. doi: 10.1037/a0030513
- Murray, S. L., Holmes, J. G., Griffin, D. W., & Derrick, J. L. (2015). The Equilibrium Model of Relationship Maintenance. *Journal of Personality and Social Psychology*, *108*(1), 93-113. doi: 10.1037/pspi0000004
- Murray, S. L., Holmes, J. G., MacDonald, G., & Ellsworth, P. C. (1998). Through the looking glass darkly? When self-doubts turn into relationship insecurities. *Journal of Personality and Social Psychology*, *75*(6), 1459-1480. doi: 10.1037//0022-3514.75.6.1459
- Murray, S. L., Holmes, J. G., & Pinkus, R. T. (2010). A smart unconscious? Procedural origins of automatic partner attitudes in marriage. *Journal of Experimental Social Psychology*, *46*(4), 650-656. doi: 10.1016/j.jesp.2010.03.003
- Murray, S. L., Lupien, S. P., & Seery, M. D. (2012). Resilience in the face of romantic rejection: The automatic impulse to trust. *Journal of Experimental Social Psychology*, *48*(4), 845-854. doi: 10.1016/j.jesp.2012.02.016
- Murray, S. L., Pinkus, R. T., Holmes, J. G., Harris, B., Gomillion, S., Aloni, M., . . . Leder, S. (2011). Signaling When (and When Not) to Be Cautious and Self-Protective: Impulsive and Reflective Trust in Close Relationships. *Journal of Personality and Social Psychology*, *101*(3), 485-502. doi: 10.1037/a0023233
- Murray, S. L., Rose, P., Bellavia, G. M., Holmes, J. G., & Kusche, A. G. (2002). When rejection stings: How self-esteem constrains relationship-enhancement processes. *Journal of Personality and Social Psychology*, *83*(3), 556-573. doi: 10.1037//0022-3514.83.3.556
- Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: opening the black box of daily life. *Psychological Medicine*, *39*(9), 1533-1547. doi: 10.1017/s0033291708004947
- Napolitano, L. A., & McKay, D. (2007). Dichotomous thinking in borderline personality disorder. *Cognitive Therapy and Research*, *31*(6), 717-726. doi: 10.1007/s10608-007-9123-4
- Nater, U. M., Bohus, M., Abbruzzese, E., Ditzen, B., Gaab, J., Kleindienst, N., . . . Ehlert, U. (2010). Increased psychological and attenuated cortisol and alpha-amylase responses to acute psychosocial stress in female patients with borderline personality disorder. *Psychoneuroendocrinology*, *35*(10), 1565-1572. doi: 10.1016/j.psyneuen.2010.06.002
- Nosek, B. A. (2007). Implicit - explicit relations. *Current Directions in Psychological Science*, *16*(2), 65-69. doi: 10.1111/j.1467-8721.2007.00477.x
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-go Association Task. *Social Cognition*, *19*(6), 625-666. doi: 10.1521/soco.19.6.625.20886

- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, *31*(2), 166-180. doi: 10.1177/0146167204271418
- Nuttin, M. J. Jr. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology*, *64*, 723-739.
- Olson, M. A., & Fazio, R. H. (2008). Implicit and explicit measures of attitudes: The perspective of the MODE model. In R. E. Petty, R. H. Fazio & P. Brinol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 19-63). Mahwah, NJ: Erlbaum.
- Oshio, Atsushi. (2012). An all-or-nothing thinking turns into darkness: Relations between dichotomous thinking and personality disorders. *Japanese Psychological Research*, *54*(4), 424-429. doi: 10.1111/j.1468-5884.2012.00515.x
- Payne, B. K., Burkley, M., & Stokes, M. B. . (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology*, *94*(1), 16-31.
- Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*, *44*(1), 29-45.
- Portella, M. J., Soler, J., Tiana, T., Barrachina, J., Pascual, J. C., Tejero, A., . . . Perez, V. (2011). Slow processing in borderline personality disorder: the emotional Stroop paradigm. *Actas Espanolas De Psiquiatria*, *39*(6), 356-362.
- Renneberg, B., Herm, K., Hahn, A., Staebler, K., Lammers, C.-H., & Roepke, S. (2012). Perception of social participation in borderline personality disorder. *Clinical Psychology & Psychotherapy*, *19*(6), 473-480. doi: 10.1002/cpp.772
- Ritter, K., Vater, A., Rusch, N., Schroder-Abe, M., Schutz, A., Fydrich, T., . . . Roepke, S. (2014). Shame in patients with narcissistic personality disorder. *Psychiatry Research*, *215*(2), 429-437. doi: 10.1016/j.psychres.2013.11.019
- Ross, J. M., & Babcock, J. C. (2009). Proactive and reactive violence among intimate partner violent men diagnosed with antisocial and borderline personality disorder. *Journal of Family Violence*, *24*(8), 607-617. doi: 10.1007/s10896-009-9259-y
- Rothschild, L., Cleland, C., Haslam, N., & Zimmerman, M. (2003). A taxometric study of borderline personality disorder. *Journal of Abnormal Psychology*, *112*(4), 657-666. doi: 10.1037/0021-843x.112.4.657
- Rullkotter, N., Markowitsch, H. J., & Driessen, M. (2011). Interference Control in Borderline Personality Disorder: The individual emotional Stroop Test in fMRI - Neuropsychology of BPD: Experimental Findings and clinical Significance. *Psychotherapie Psychosomatik Medizinische Psychologie*, *61*(2), 111-111.
- Rusch, N., Corrigan, P. W., Bohus, M., Kuehler, T., Jacob, G. A., & Lieb, K. (2007). The impact of posttraumatic stress disorder on dysfunctional implicit and explicit emotions among women with borderline personality disorder. *Journal of Nervous and Mental Disease*, *195*(6), 537-539. doi: 10.1097/NMD.0b013e318064e7fc
- Rusch, N., Lieb, K., Gottler, I., Hermann, C., Schramm, E., Richter, H., . . . Bohus, M. (2007). Shame and implicit self-concept in women with borderline personality disorder. *American Journal of Psychiatry*, *164*(3), 500-508. doi: 10.1176/appi.ajp.164.3.500
- Rusch, N., Schulz, D., Valerius, G., Steil, R., Bohus, M., & Schmahl, C. (2011). Disgust and implicit self-concept in women with borderline personality disorder and posttraumatic

- stress disorder. *European Archives of Psychiatry and Clinical Neuroscience*, 261(5), 369-376. doi: 10.1007/s00406-010-0174-2
- Sadikaj, G., Moskowitz, D. S., Russell, J. J., Zuroff, D. C., & Paris, J. (2013). Quarrelsome behavior in borderline personality disorder: influence of behavioral and affective reactivity to perceptions of others. *Journal of Abnormal Psychology*, 122(1), 195-207. doi: 10.1037/a0030871
- Sadikaj, G., Moskowitz, D. S., & Zuroff, D. C. (2015). Intrapersonal variability in negative affect as a moderator of accuracy and bias in interpersonal perception. *Journal of Personality Disorders*, 29(4), 468-485.
- Sadikaj, G., Russell, J. J., Moskowitz, D. S., & Paris, J. (2010). Affect dysregulation in individuals with borderline personality disorder: persistence and interpersonal triggers. *Journal of Personality Assessment*, 92(6), 490-500. doi: 10.1080/00223891.2010.513287
- Salekin, R. T., Trobst, K. K., & Krioukova, M. (2001). Construct validity of psychopathy in a community sample: A nomological net approach. *Journal of Personality Disorders*, 15(5), 425-441. doi: 10.1521/pedi.15.5.425.19196
- Santangelo, P., Bohus, M., & Ebner-Priemer, U. W. (2012). Ecological momentary assessment in borderline personality disorder: A review of recent findings and methodological challenges. *Journal of Personality Disorders*, 28(4), 555-576.
- Scott, L. N., Levy, K. N., & Granger, D. A. (2013). Biobehavioral reactivity to social evaluative stress in women with borderline personality disorder. *Personality Disorders-Theory Research and Treatment*, 4(2), 91-100. doi: 10.1037/a0030117
- Sheffield, M. W., Barlow, S. H., Lambert, M. J., Hoyal, H., Thompson, K. C., & Garbutt, M. (1999). A qualitative examination of borderline personality disordered (BPD) patients: Interpersonal dynamics and underlying paradoxes. *Journal of Personality Disorders*, 13(3), 287-296.
- Sieswerda, S., Arntz, A., & Kindt, M. (2007). Successful psychotherapy reduces hypervigilance in borderline personality disorder. *Behavioural and Cognitive Psychotherapy*, 35(4), 387-402. doi: 10.1017/s1352465807003694
- Sieswerda, S., Arntz, A., Mertens, I., & Vertommen, S. (2007). Hypervigilance in patients with borderline personality disorder: Specificity, automaticity, and predictors. *Behaviour Research and Therapy*, 45(5), 1011-1024. doi: 10.1016/j.brat.2006.07.012
- Sieswerda, S., Arntz, A., & Wolfis, M. (2005). Evaluations of emotional noninterpersonal situations by patients with borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 36(3), 209-225. doi: 10.1016/j.jbtep.2005.05.004
- Sieswerda, S., Barnow, S., Verheul, R., & Arntz, A. (2013). Neither dichotomous nor split, but schema-related negative interpersonal evaluations characterize borderline patients. *Journal of Personality Disorders*, 27(1), 36-52.
- Silbersweig, David, Clarkin, John F., Goldstein, Martin, Kernberg, Otto F., Tuescher, Oliver, Levy, Kenneth N., . . . Stern, Emily. (2007). Failure of frontolimbic inhibitory function in the context of negative emotion in borderline personality disorder. *American Journal of Psychiatry*, 164(12), 1832-1841. doi: 10.1176/appi.ajp.2007.06010126
- Simpson, J. A., Ickes, W., & Grich, J. (1999). When accuracy hurts: Reactions of anxious-ambivalent dating partners to a relationship-threatening situation. *Journal of Personality and Social Psychology*, 76(5), 754-769.

- Skowron, E. A., & Friedlander, M. L. (1998). The differentiation of self inventory: Development and initial validation. *Journal of Counseling Psychology, 45*(3), 235-246. doi: 10.1037/0022-0167.45.3.235
- Soloff, Paul H., Lis, Judith A., Kelly, Thomas, Cornelius, Jack, & et al. (1994). Self-mutilation and suicidal behavior in borderline personality disorder. *Journal of Personality Disorders, 8*(4), 257-267.
- Staebler, K., Helbing, E., Rosenbach, C., & Renneberg, B. (2011). Rejection sensitivity and borderline personality disorder. *Clinical Psychology & Psychotherapy, 18*(4), 275-283. doi: 10.1002/cpp.705
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*(3), 220-247. doi: 10.1207/s15327957pspr0803_1
- Stuart, J., Westen, D., Lohr, N. E., Benjamin, J., & et al. (1990). Object relations in borderlines, depressives, and normals: An examination of human responses on the Rorschach. *Journal of Personality Assessment, 55*(1-2), 296-318.
- Swartz, Marvin, Blazer, Dan, George, Linda, & Winfield, Idee. (1990). Estimating the prevalence of borderline personality disorder in the community. *Journal of Personality Disorders, 4*(3), 257-272.
- Teachman, B. A., Joormann, J., Steinman, S. A., & Gotlib, I. H. (2012). Automaticity in anxiety disorders and major depressive disorder. *Clinical Psychology Review, 32*(6), 575-603. doi: 10.1016/j.cpr.2012.06.004
- Tragesser, S. L., & Benfield, J. (2012). Borderline personality disorder features and mate retention tactics. *Journal of Personality Disorders, 26*(3), 334-344.
- Tragesser, S. L., Lippman, L. G., Trull, T. J., & Barrett, K. C. (2008). Borderline personality disorder features and cognitive, emotional, and predicted behavioral reactions to teasing. *Journal of Research in Personality, 42*(6), 1512-1523. doi: 10.1016/j.jrp.2008.07.003
- Trull, T. J. (1995). Borderline personality disorder features in non-clinical young-adults: I. Identification and validation. *Psychological Assessment, 7*(1), 33-41. doi: 10.1037/1040-3590.7.1.33
- Trull, T. J., Solhan, M. B., Tragesser, S. L., Jahng, S., Wood, P. K., Piasecki, T. M., & Watson, D. (2008). Affective instability: Measuring a core feature of borderline personality disorder with ecological momentary assessment. *Journal of Abnormal Psychology, 117*(3), 647-661. doi: 10.1037/a0012532
- Trull, T. J., Useda, J. D., Conforti, K., & Doan, B. T. (1997). Borderline personality disorder features in nonclinical young adults: 2. Two-year outcome. *Journal of Abnormal Psychology, 106*(2), 307-314. doi: 10.1037//0021-843x.106.2.307
- Valentiner, D. P., Hiraoka, R., & Skowronski, J. J. (2014). Borderline personality disorder features, self-verification, and committed relationships. *Journal of Social and Clinical Psychology, 33*(5), 463-480.
- Vater, A., Ritter, K., Schroder-Abe, M., Schutz, A., Lammers, C. H., Bosson, J. K., & Roepke, S. (2013). When grandiosity and vulnerability collide: Implicit and explicit self-esteem in patients with narcissistic personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 44*(1), 37-47. doi: 10.1016/j.jbtep.2012.07.001
- Vater, A., Schroeder-Abe, M., Weissgerber, S., Roepke, S., & Schuetz, A. (2015). Self-concept structure and borderline personality disorder: evidence for negative

- compartmentalization. *Journal of Behavior Therapy and Experimental Psychiatry*, 46, 50-58. doi: 10.1016/j.jbtep.2014.08.003
- Veen, G., & Arntz, A. (2000). Multidimensional dichotomous thinking characterizes borderline personality disorder. *Cognitive Therapy and Research*, 24(1), 23-45. doi: 10.1023/a:1005498824175
- von Ceumern-Lindenstjerna, I. A., Brunner, R., Parzer, P., Mundt, C., Fiedler, P., & Resch, F. (2010). Attentional bias in later stages of emotional information processing in female adolescents with borderline personality disorder. *Psychopathology*, 43(1), 25-32. doi: 10.1159/000255960
- von Ceumern-Lindenstjerna, I.-A., Brunner, R., Parzer, P., Mundt, C., Fiedler, P., & Resch, F. (2010). Attentional bias in later stages of emotional information processing in female adolescents with borderline personality disorder. *Psychopathology*, 43(1), 25-32. doi: 10.1159/000255960
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070. doi: 10.1037/0022-3514.54.6.1063
- Wentura, D., & Degner, J. (2010). A practical guide to sequential priming and related tasks In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications*. New York, NY: Guilford Press.
- Westen, D. (1991). Social cognition and object relations. *Psychological Bulletin*, 109(3), 429-455. doi: 10.1037/0033-2909.109.3.429
- Westen, D., Lohr, N., Silk, K. R., Gold, L., & Kerber, K. (1990). Object relations and social cognition in borderlines, major depressives, and normals: A Thematic Apperception Test analysis. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 2(4), 355-364.
- Whisman, M. A., & Schonbrun, Y. C. (2009). Social consequences of borderline personality disorder symptoms in a population-based survey: marital distress, marital violence, and marital disruption. *Journal of Personality Disorders*, 23(4), 410-415.
- Wingenfeld, K., Mensebach, C., Rullkoetter, N., Schlosser, N., Schaffrath, C., Woermann, F. G., . . . Beblo, T. (2009). Attentional bias to personally relevant words in borderline personality disorder is strongly related to comorbid posttraumatic stress disorder. *Journal of Personality Disorders*, 23(2), 141-155.
- Wingenfeld, K., Rullkoetter, N., Mensebach, C., Beblo, T., Mertens, M., Kreisel, S., . . . Woermann, F. G. (2009). Neural correlates of the individual emotional Stroop in borderline personality disorder. *Psychoneuroendocrinology*, 34(4), 571-586. doi: 10.1016/j.psyneuen.2008.10.024
- Winter, D., Herbert, C., Koplín, K., Schmahl, C., Bohus, M., & Lis, S. (2015). Negative evaluation bias for positive self-referential information in borderline personality disorder. *Plos One*, 10(1). doi: 10.1371/journal.pone.0117083
- Winter, D., Krause-Utz, A., Lis, S., Chiu, C. D., Lanius, R. A., Schrinier, F., . . . Schmahl, C. (2015). Dissociation in borderline personality disorder: Disturbed cognitive and emotional inhibition and its neural correlates. *Psychiatry Research-Neuroimaging*, 233(3), 339-351. doi: 10.1016/j.pychresns.2015.05.018
- Witthoft, M., Borgmann, E., White, A., & Dyer, A. (2015). Body-related attentional biases in patients with posttraumatic stress disorder resulting from childhood sexual abuse with

- and without co-occurring borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 46, 72-77. doi: 10.1016/j.jbtep.2014.07.010
- Zanarini, M. C., Frankenburg, F. R., Dubo, E. D., Sickel, A. E., Trikha, A., Levin, A., & Reynolds, V. (1998). Axis I comorbidity of borderline personality disorder. *American Journal of Psychiatry*, 155(12), 1733-1739.
- Zanarini, M. C., Frankenburg, F. R., Khera, G. S., & Bleichmar, J. (2001). Treatment histories of borderline inpatients. *Comprehensive Psychiatry*, 42(2), 144-150. doi: 10.1053/comp.2001.19749
- Zayas, V., & Shoda, Y. (2005). Do automatic reactions elicited by thoughts of romantic partner, mother, and self relate to adult romantic attachment? *Personality and Social Psychology Bulletin*, 31(8), 1011-1025. doi: 10.1177/0146167204274100
- Zeigler-Hill, V., & Abraham, J. (2006). Borderline personality features: Instability of self-esteem and affect. *Journal of Social and Clinical Psychology*, 25(6), 668-687. doi: 10.1521/jscp.2006.25.6.668

Annexe 1 : Consent form for Study 1

FORMULAIRE DE CONSENTEMENT

Titre de la recherche : Instabilité des perceptions spontanées en relation avec les traits du trouble de la personnalité limite.

Chercheur : Jean Gagnon, professeur agrégé, Département de psychologie, Université de Montréal.

Étudiant co-chercheur : Félix Gauthier Mongeon, étudiant au doctorat, Département de psychologie, Université de Montréal.

A) RENSEIGNEMENTS AUX PARTICIPANTS

1. Objectifs de la recherche

Ce projet vise à mieux comprendre comment les femmes adultes organisent leurs perceptions d'autres individus en interaction lors de conflits relationnels. Le recrutement est effectué parmi les étudiantes au baccalauréat en psychologie.

2. Participation à la recherche

Votre participation à ce projet consiste en deux volets. Le premier volet est effectué en classe et consiste à remplir un bref questionnaire sur la personnalité (9 questions, 2 minutes). En laissant vos coordonnées, vous pourriez être recontacté dans les prochaines semaines et il sera à votre discrétion de déterminer si vous êtes toujours intéressé à participer au second volet. Le second volet se déroule au département de psychologie. Votre participation implique de visualiser trois extraits de film de 5 minutes chacun mettant en scène des interactions conjugales; vous répondrez également à quatre brefs questionnaires et effectuerez une tâche informatisée sur vos perceptions des personnages de l'extrait, sur vos émotions pendant l'extrait et sur votre personnalité. Le tout devrait durer environ 40 minutes.

3. Confidentialité

Les renseignements personnels que vous nous donnerez demeureront confidentiels. Chaque participant à la recherche se verra attribuer un numéro et seul le chercheur et son équipe auront la liste des participants et des numéros correspondants. De plus, les données seront conservées dans un lieu sûr. Aucune information permettant de vous identifier d'une façon ou d'une autre ne sera publiée. Toute information personnelle sera détruite sept ans après la fin du projet. Seules les données ne permettant pas de vous identifier seront conservées après cette période.

4. Avantages et inconvénients

En participant à cette recherche, vous pourrez contribuer à une meilleure compréhension de la façon dont les individus adultes organisent leurs perceptions d'autrui en relation. Il n'y a pas de risque particulier à participer à ce projet. Il est possible cependant que les extraits de film évoquent chez vous des émotions légèrement inconfortables.

5. Droit de retrait

Votre participation à ce projet est entièrement volontaire et vous pouvez à tout moment vous retirer de la recherche sur simple avis verbal et sans devoir justifier votre décision. Si vous choisissez de ne pas participer, vous n'avez qu'à remettre le questionnaire non rempli. À votre demande, tous les renseignements qui vous concernent pourront aussi être détruits. Cependant, après le déclenchement du processus de publication (où seules pourront être diffusées des informations ne permettant pas de vous identifier), il sera impossible de détruire les analyses et les résultats portant sur vos données. Les résultats pourraient être utilisés pour la thèse doctorale du chercheur et pour des fins de publication et communication scientifiques.

B) CONSENTEMENT

J'ai pris connaissance des informations ci-dessus et je n'ai pas d'autres questions concernant ce projet ainsi que ma participation. Je consens librement à prendre part à cette recherche et je sais que je peux me retirer en tout temps sans avoir à justifier ma décision:

Oui Non

Je consens à être recontacté par courriel pour le deuxième volet de l'étude:

Oui Non

Le cas échéant, veuillez s'il vous plaît indiquer votre adresse courriel.

Adresse courriel: _____

Signature : _____ Date : _____

Nom : _____ Prénom : _____

Je déclare avoir expliqué le but, la nature, les avantages et les inconvénients de l'étude et avoir répondu au meilleur de ma connaissance aux questions posées.

Signature du chercheur : _____ Date : _____
(ou de son représentant)

Nom : _____ Prénom : _____

Annexe 2 : Ethics approval for Study 1



N° de certificat
CERAS-2014-15-175-D

Comité d'éthique de la recherche en arts et en sciences

CERTIFICAT D'APPROBATION ÉTHIQUE

Le Comité d'éthique de la recherche en arts et en sciences (CÉRAS), selon les procédures en vigueur, en vertu des documents qui lui ont été fournis, a examiné le projet de recherche suivant et conclu qu'il respecte les règles d'éthique énoncées dans la Politique sur la recherche avec des êtres humains de l'Université de Montréal.

Projet	
Titre du projet	Instabilité des perceptions spontanées en relation avec les traits du trouble de la personnalité limite
Étudiant requérant	Félix Gauthier-Mongeon Étudiant au doctorat, FAS-Département de psychologie
Sous la direction de	Jean Gagnon, Professeur agrégé, FAS-Département de psychologie
Autres membres de l'équipe:	Bertram Gawronski (University of Austin)
Financement	
Organisme	Non financé
Programme	
Titre de l'octroi si différent	
Numéro d'octroi	
Chercheur principal	
No de compte	

MODALITÉS D'APPLICATION

Tout changement anticipé au protocole de recherche doit être communiqué au CÉRAS qui en évaluera l'impact au chapitre de l'éthique.

Toute interruption prématurée du projet ou tout incident grave doit être immédiatement signalé au CÉRAS.

Selon les règles universitaires en vigueur, un suivi annuel est minimalement exigé pour maintenir la validité de la présente approbation éthique, et ce, jusqu'à la fin du projet. Le questionnaire de suivi est disponible sur la page web du CÉRAS.

Annexe 3 : Material for Study 1

Reconnaissance des visages :



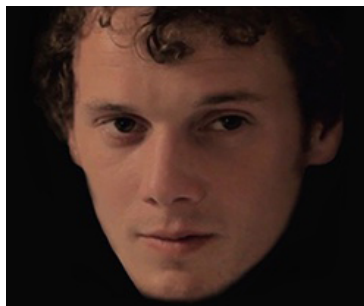
a. À quel point pensez-vous reconnaître la personne ci-haut?

Je ne le reconnais pas du tout	Un petit peu	Beaucoup	Je suis certaine de le reconnaître
1	2	3	4



b. À quel point pensez-vous reconnaître la personne ci-haut?

Je ne le reconnais pas du tout	Un petit peu	Beaucoup	Je suis certaine de le reconnaître
1	2	3	4



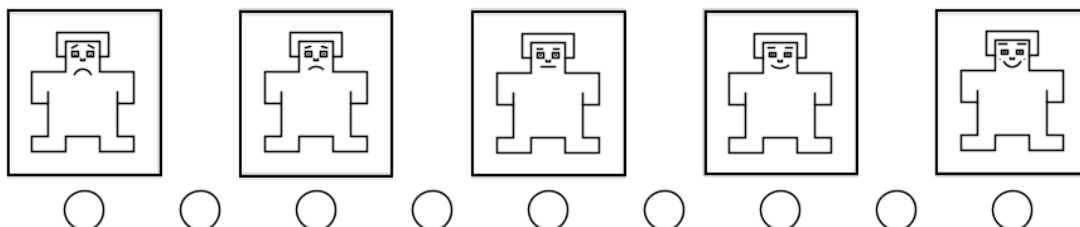
c. À quel point pensez-vous reconnaître la personne ci-haut?

Je ne le reconnais pas du tout	Un petit peu	Beaucoup	Je suis certaine de le reconnaître
1	2	3	4

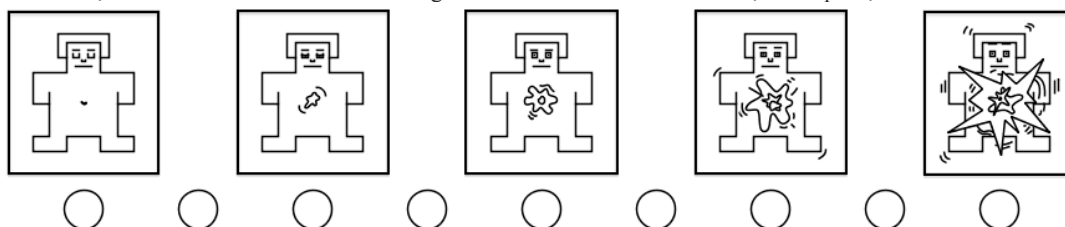
Manikins – émotions :

T1 :

La première série de Manikins représente une émotion sur un continuum de positif (Manikin qui sourit) à négatif (Manikin qui fait la moue). Veuillez cocher le cercle à l'extrême droite si vous vous sentez heureux, joyeux, satisfait ou content ou, au contraire, à l'extrême gauche si vous vous sentez malheureux, triste, mélancolique ou déprimé. Si vous ne ressentez rien en particulier, vous pouvez répondre de façon neutre en cochant le cercle situé au milieu.

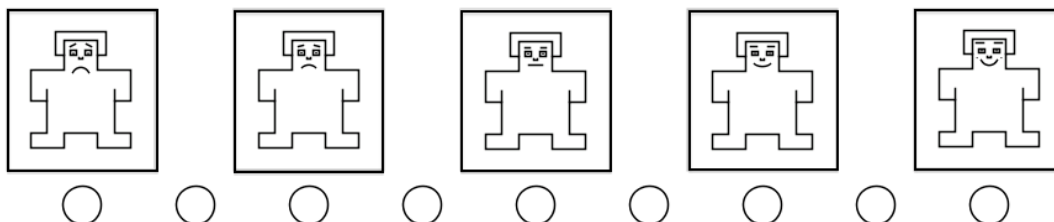


La deuxième série représente une émotion de très calme (Manikin avec les yeux fermés et qui dort) à très intense (Manikin qui semble sur le point d'exploser). Veuillez cocher le cercle à l'extrême droite si vous vous sentez excité, très éveillé ou très intense ou, au contraire le cercle à l'extrême gauche si vous vous sentez calme, très reposé, inactif ou bien même endormi.

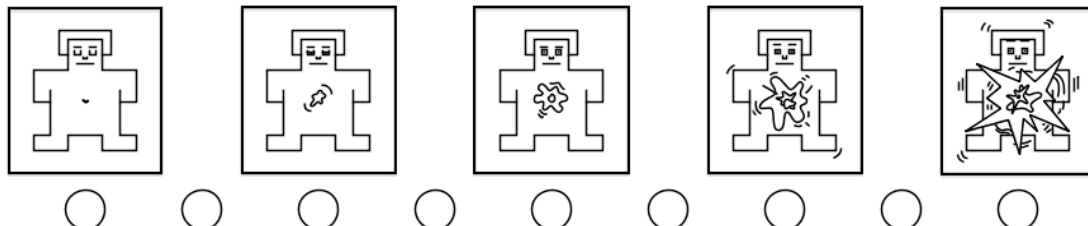


T2 :

La première série de Manikins représente une émotion sur un continuum de positif (Manikin qui sourit) à négatif (Manikin qui fait la moue). Veuillez cocher le cercle à l'extrême droite si vous vous sentez heureux, joyeux, satisfait ou content ou, au contraire, à l'extrême gauche si vous vous sentez malheureux, triste, mélancolique ou déprimé. Si vous ne ressentez rien en particulier, vous pouvez répondre de façon neutre en cochant le cercle situé au milieu.



La deuxième série représente une émotion de très calme (Manikin avec les yeux fermés et qui dort) à très intense (Manikin qui semble sur le point d'exploser). Veuillez cocher le cercle à l'extrême droite si vous vous sentez excité, très éveillé ou très intense ou, au contraire le cercle à l'extrême gauche si vous vous sentez calme, très reposé, inactif ou bien même endormi.



Questionnaire PANAS – émotions :

Ce questionnaire contient des adjectifs qui décrivent des sentiments et des émotions. Lisez chacun de ces adjectifs. Pour chacun de ces adjectifs, vous devez indiquer à quel point il décrit comment vous vous sentez présentement. N'oubliez pas, il n'y a pas de bonnes ou de mauvaises réponses. Nous voulons savoir comment vous vous sentez présentement.

1	2	3	4	5
Jamais	Un peu	Modérément	Souvent	Toujours

_____ 1. Intéressé	_____ 11. Irritable
_____ 2. Perturbé	_____ 12. Vigilant
_____ 3. Excité	_____ 13. Honteux
_____ 4. Bouleversé	_____ 14. Inspiré
_____ 5. Fort	_____ 15. Nerveux
_____ 6. Coupable	_____ 16. Déterminé
_____ 7. Effrayé	_____ 17. Attentif
_____ 8. Hostile	_____ 18. Agité
_____ 9. Enthousiaste	_____ 19. Actif
_____ 10. Fier	_____ 20. Apeuré

Familiarité de l'extrait :

Consigne: À quel point pensez-vous reconnaître l'extrait de film?

	Pas du tout	Un petit peu	Beaucoup	Tout à fait
a) Je reconnais l'extrait	1	2	3	4

Questionnaire sur la personnalité – PAI-BOR :

Ce questionnaire contient un certain nombre d'énoncés. Pour chacun, décidez à quel degré l'énoncé décrit votre fonctionnement général. Indiquez votre réponse en cochant la case appropriée. Il est important de répondre à tous les énoncés.

Passez à la page suivante lorsque vous avez terminé.

1. Mon humeur peut changer très subitement.	2. Mon attitude à propos de moi-même change beaucoup.
0. Faux, pas du tout vrai	0. Faux, pas du tout vrai
1. Un peu vrai	1. Un peu vrai
2. Surtout vrai	2. Surtout vrai
3. Très vrai	3. Très vrai

3. Mes relations ont été orageuses.	4. Mes humeurs sont assez intenses.
0. Faux, pas du tout vrai	0. Faux, pas du tout vrai
1. Un peu vrai	1. Un peu vrai
2. Surtout vrai	2. Surtout vrai
3. Très vrai	3. Très vrai

5. Parfois je me sens terriblement vide en dedans.	6. Je tiens à ce que certaines personnes sachent à quel point elles m'ont blessé(e).
0. Faux, pas du tout vrai	0. Faux, pas du tout vrai
1. Un peu vrai	1. Un peu vrai
2. Surtout vrai	2. Surtout vrai
3. Très vrai	3. Très vrai

7. Mon humeur est très stable.	8. Ça m'inquiète beaucoup de penser que les autres peuvent m'abandonner.
0. Faux, pas du tout vrai	0. Faux, pas du tout vrai
1. Un peu vrai	1. Un peu vrai
2. Surtout vrai	2. Surtout vrai
3. Très vrai	3. Très vrai

9. Une fois devenus proches de moi, les gens m'ont laissé(e) tomber.	10. J'ai peu de contrôle sur ma colère
0. Faux, pas du tout vrai	0. Faux, pas du tout vrai
1. Un peu vrai	1. Un peu vrai
2. Surtout vrai	2. Surtout vrai

3. Très vrai

3. Très vrai

11. Je me demande souvent ce que je devrais faire de ma vie. 12. Je me sens rarement très seule.

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

13. Je fais parfois les choses de façon si impulsive que je m'attire des ennuis. 14. J'ai toujours été une personne très heureuse.

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

15. Je supporte difficilement la séparation.

16. J'ai fait quelques graves erreurs dans le choix de mes amis.

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

17. Typiquement, quand je suis très bouleversé(e), je fais quelque chose pour me blesser.

18. J'ai eu des moments où j'étais tellement furieux(se) que je ne pouvais pas en faire assez pour exprimer toute ma colère.

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

19. Je ne m'ennuie pas très facilement.

20. Une fois que quelqu'un est devenu mon ami, nous restons amis.

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

21. Je suis trop impulsif(ve) à mon goût.

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

23. Je suis un(e) casse-cou.

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

22. Je dépense l'argent trop facilement.

- 0. Faux, pas du tout vrai
- 1. Un peu vrai
- 2. Surtout vrai
- 3. Très vrai

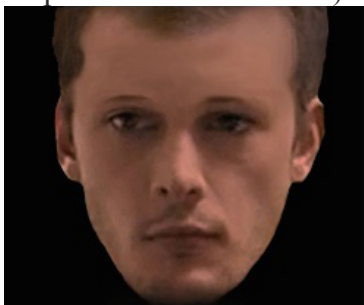
24. Je fais attention à la façon dont je dépense mon argent.

- 0. Faux, pas du tout vrai
 - 1. Un peu vrai
 - 2. Surtout vrai
 - 3. Très vrai
-

Évaluations du personnage :

Consigne:

En vous basant sur votre première impression, à quel point pensez-vous que cet individu est acceptant? (cette consigne est présentée pour chacun des 20 items en contrebalançant l'ordre des visages. La mesure implicite emploie les mêmes stimuli).



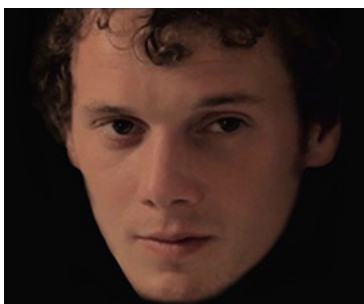
a. À quel point pensez-vous reconnaître la personne ci-haut?

Pas du tout acceptant						Tout à fait acceptant
1	2	3	4	5	6	7



b. À quel point pensez-vous reconnaître la personne ci-haut?

Je ne le reconnais pas du tout	Un petit peu	Beaucoup	Je suis certaine de le reconnaître
1	2	3	4



c. À quel point pensez-vous reconnaître la personne ci-haut?

Je ne le reconnais pas du tout	Un petit peu	Beaucoup	Je suis certaine de le reconnaître
1	2	3	4

List of positive and negative words used in the explicit and implicit measures in Study 1

Positive words	Negative words
Acceptant	Distant
Agréable	Indifférent
Doux	Blâmant
Affectueux	Froid
Attentionné	Colérique
Sensible	Malhonnête
Gentil	Rejetant
Aimable	Agressif
Ouvert	Déplaisant
Chaleureux	Insensible

These are the original words (in French) that used in the task

Annexe 4 : Consent form for Study 2

LETTER OF INFORMATION

Title of Project: Personality and Romantic Relationships

We are inviting you to participate in a research study on the role of personality characteristics in romantic relationships, conducted by Dr. Bertram Gawronski, Ph.D., Professor and Canada Research Chair, Félix Gauthier Mongeon, and Sarah Stanton of the Department of Psychology at The University of Western Ontario.

The experiment involves completing an online survey comprising a few questions about your personality and your relationship as well as a simple evaluation task. This study will take approximately 10-15 minutes to complete.

Please note that you have to be in a romantic relationship for at least 6 months to be eligible for this study. There are no known risks involved with participating in this study. You will receive a total 1\$ in return for your participation. You may refuse to participate, refuse to answer any questions, or withdraw from the study at any time.

You will receive written information about the purpose of the study at the end of the session. If you have any questions about the research or want to obtain information about results, feel free to contact Dr. Bertram Gawronski.

All data collected will be used for research purposes only and stored on a secure online server viewed only by authorized researchers. All possible measures will be taken to protect your confidentiality and you will not be identified individually with any responses you provide during your participation.

If you have any questions about the conduct of this study or your rights as a research participant you may contact the Director of the Office of Research Ethics, The University of Western Ontario.

You may print a copy of this letter.

Please indicate your consent by clicking the NEXT button at the bottom of the screen.

Annexe 5 : Ethics approval for Study 2

Department of Psychology The University of Western
Ontario
Room 7418
Social Sciences
Centre, London,
ON, Canada
N6A 5C1

Use of Human Subjects - Ethics Approval Notice

Principal Investigator	Bertram Gawronski/Felix Gauthier Mongeon/Sarah Stanton
Protocol	Personality and romantic relationships

This is to notify you that The University of Western Ontario Department of Psychology Research Ethics Board (PREB) has granted expedited ethics approval to the above named research study on the date noted above.

The PREB is a sub-REB of The University of Western Ontario's Research Ethics Board for Non-Medical Research Involving Human Subjects (NMREB) which is organized and operates according to the Tri-Council Policy Statement and the applicable laws and regulations of Ontario. (See Office of Research Ethics web site: <http://www.uwo.ca/research/ethics/>)

This approval shall remain valid until end date noted above assuming timely and acceptable responses to the University's periodic requests for surveillance and monitoring information.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the PREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of research assistant, telephone number etc). Subjects must receive a copy of the information/consent documentation.

Investigators must promptly also report to the PREB:

- changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
- all adverse and unexpected experiences or events that are both serious and unexpected;
- new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to the PREB for approval.

Members of the PREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the PREB.

Clive Seligman Ph.D.

Chair, Psychology Expedited Research Ethics Board (PREB)

The other members of the 2012-2013 PRES are: Mike Atkinson (Introductory Psychology Coordinator), Rick Goffin, Riley Hinson, Albert Katz (Department Chair), Steve Lupker, and Adam Piraino (Graduate Student Representative)

CC: UWO Office of Research Ethics

This is an official document. Please retain the original in your files

SECTION 1: PROJECT REGISTRATION

1.1.1 Project Title: Personality and Romantic Relationships					
1.2a Anticipated Project dates Academic Year 2012-2013, 2013-2014	<table border="1"> <tr> <td>Start Date 2013</td> <td>June 13,</td> </tr> <tr> <td>Completion Date 2013</td> <td>Dec. 31,</td> </tr> </table>	Start Date 2013	June 13,	Completion Date 2013	Dec. 31,
Start Date 2013	June 13,				
Completion Date 2013	Dec. 31,				

1.4a	Signature of Local Principal Investigator or Sponsor of Student/Post-Doc/Visiting Scholar attestin that:
	<p>a) <i>allco-investigators have reviewed the protocol contents and are in agreement with the protocol as submitted;</i></p> <p>b) all investigators have read the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans (1998) and the UWO Guidelines on Non-Medical Research Involving Human Subjects and agree to abide by the guidelines therein;</p> <p>c) the investigator(s) will adhere to the Protocol and Consent Form as approved by the REB; and</p> <p>d) the Principal Investigator will notify the REB of any changes or adverse events/experiences in a timely manner;</p> <p>e) the study, if funded by an external sponsor, will not start until the contract agreement has been approved by the appropriate university or research institute official.</p> <p style="text-align: center;">D the research is no or minimal risk</p> <p style="text-align: center;">_____</p>

1.4b	Signature(s) of Student(s), Post-Doc, or Scholar attest that the:
	<p>a) have read the Tri-Council Policy Statement and the UWO Guidelines on Non-Medical Research Involving Human Subjects and agree to abide by the guidelines therein;</p> <p>b) will adhere to the Protocol and Consent Form as approved by the REB; and</p> <p>c) will report to their supervisor and the REB of any changes or adverse events/experiences in a timely manner;</p>

SECTION 2: PROJECT INFORMATION

2.1 **Objectives and Hypotheses:** Provide a clear statement of the purpose and objectives of the project (please maximize)

In this project, we will investigate the relationship between borderline personality features and emotional responses to positive and negative relationship events.

Borderline personality disorder is marked by pervasive emotional dysregulation and relationship conflicts (Gratz, Dixon-Gordon, Breetz, & Tull, 2013; Sadikaj, Moskowitz, Russell, Zuroff, & Paris, 2013). The discovery of ubiquitous subclinical features of borderline personality disorder brought greater attention toward the role of borderline characteristics in the normal population (Jackson & Trull, 2001; Trull, 1995; Trull, Useda, Conforti, & Doan, 1997). Meanwhile, advances in social psychology have provided fruitful insights in the structure of social cognition: Abounding empirical evidence supports a distinction between two levels of social-cognitive outputs, that is, spontaneous and deliberate responses (Gawronski & Bodenhausen, 2006, 2011; Strack & Deutsch, 2004). Yet, no study has examined how spontaneous and deliberate responses to relationship events might be related with borderline features. Studying such responses are particularly relevant given the detrimental impact of dysregulated emotional responses in close relationships in borderline personality disorder (Gratz et al., 2013; Sadikaj et al., 2013).

Hence, the present study aims at examining the relationship between borderline features in normal adult participants and spontaneous/deliberate emotional reactions to positive and negative relationship events. The focus is on romantic relationships because those are particularly central in adulthood (Hazan & Shaver, 1987). The now well-established presence of subclinical borderline features in the normal population justifies using a pool of normal participants (Jackson & Trull, 2001; Trull, 1995; Trull et al., 1997).

We expect that borderline features will concurrently be related with more intense spontaneous and deliberate emotional responses to the partner for those participants who imagine a positive or a negative event occurring within their relationship, while controlling for depression and anxiety symptoms. We also examine whether borderline features will show different relations with spontaneous and deliberate measures.

This research will involve first collecting information on participants' borderline features. Then, participants will be assigned to one of three conditions. One third of the participants will be asked to imagine a hypothetical situation where their partner rejects them (negative condition); another third will be asked to imagine a hypothetical situation where their partner seeks connection (positive condition); the last third will complete the dependent measures without being asked to imagine a hypothetical situation (control condition).

All participants will then complete a questionnaire and letter evaluation task that respectively tap deliberate and spontaneous responses to the partner. Finally, participants will complete a questionnaire on depression and anxiety. We expect a specific pattern of relationship between borderline features and emotional responses to the partner rather than a relationship that is transversal to other well-known psychopathological features (in this case, depression and anxiety).

2.2	Research Participants:
-----	-------------------------------

If you are requesting to use the department's subject pool, please indicate the number of credits per participant _n/a_ and the total number of credits for the study _n/a_.

Briefly describe the sample, number of participants, and any exclusionary criteria, e.g., exclude non-English speaking participants.

300 English speaking American adults; exclusion criteria: participants must be in a romantic relationship for at least 6 months

Describe the **method of recruiting participants**, and any **compensation** offered.

300 participants will be recruited by means of the Mturk participation website. Participants will receive 1\$ as compensation for their participation in a 10-15 minutes session, which is comparable to compensation for other projects on Mturk.

Include one copy of the **sign-up poster** or advertisements if used.

See Enclosures

Include one copy of the **letter of information** and one copy of the **informed consent** sheet. See

Enclosures

Include one copy of your **debriefing sheet**.

See Enclosures

If the research will not be conducted in the Social Science Centre, please indicate the location:

Participants will complete the survey online:

https://uwopsych.qualtrics.com/SE/?SID=SV_OHuymlnXqcpXPxj

2.3	<p>Methodology - Describe the study design and procedure, that is, what participants will be asked to do at each stage of the research, e.g. manipulations. (2 page maximum)</p> <p>Include one copy of all measures, e.g., questionnaires, scales. Indicate if sensitive questions are being asked, e.g., sexual behavior, religious beliefs, suicide ideation, and the like.</p> <p>Does your research involve deception? If so, please describe the deception and the reasons for it, and indicate how the participants will be debriefed.</p> <p>Describe any risks and/or discomforts to the participants and how you would deal with them.</p> <p>Describe the procedures to be used to ensure confidentiality of participants and for preserving the confidentiality of data during the research, storage, disposition and in the release of the findings.</p>
-----	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Participants will be invited to complete an online survey by means of the Mturk participation website. The survey should take 10-15 minutes to complete and will be completed through qualtrics.

In this experiment, participants will be informed that they will be asked to answer a few questions about their personality and their relationship (see Materials).

After completing the borderline features' questionnaire, one third of the participants will be asked to imagine a hypothetical situation where their partner rejects them (negative condition); another third will be asked to imagine a hypothetical situation where their partner seeks connection (positive condition); the last third will complete the dependent measures without being asked to imagine a hypothetical situation (control condition).

The study does not involve deception.

Research participants will complete the study at their leisure and will be advised to do so in a private setting. Participants will be identified by code number, not by name. Records will be completely anonymous. All data will be recorded by on a safe secure server to which only the Principal Investigators will have access. If the results of the project are published, research subjects will not be personally identified. Data generated will not be used for purposes other than this research project. No nominal information will be asked about any other person. Data will be stored in electronic format, only for as long as the retention period mandated by the American Psychological Association.

All data will be collected electronically via an online survey. There will be no written records or videotapes. The data from this study will be stored on a secure server and when

downloaded will be kept on a computer in a locked room supervised by the Principal Investigator. In accordance with standards set by the American Psychological Association, all data will be retained for a period of 5 years after publication of the results.

Annexe 6 : Material for Study 2

Demographic Instructions and Materials

Before we begin, please answer the following demographic questions?

How old are you? _____ years

What is your gender? male / female

What is your relationship status? (married, single, dating, etc.) [ask Sarah]

What is sexual orientation? straight, gay/lesbian, bisexual, prefer not to answer

How long have you been in your current relationship? _____ years _____ months

What is your race/ethnicity? _____

What is your highest level of education? _____

Description of the study

The present study examines the role of personality characteristics in romantic relationships. In the following tasks, we will ask you to answer a few questions about your personality and your relationship. The study will take approximately 10-15 minutes.

Personality Questionnaire

In the first part, we are interested in aspects of your personality.

The purpose of the following questionnaire is for you to describe the kind of person you are. When answering the questions, think about how you have tended to feel, think, and act over the past several years.

Please answer either **True** or **False** to each item.

Where:

T (True) means that the statement is generally true for you.

F(False) means that the statement is generally false for you.

Even if you are not entirely sure about the answer, indicate “T” or “F” for every question.

For example, for the question:

xx. I tend to be stubborn.

T F

If, in fact you have been stubborn over the past several years, you would answer True by circling T.

If, this were not true at all for you, you would answer False by circling F.

There are no correct answers.

You may take as much time as you wish.

- | | | | |
|---|---------------------------------------------------------------------------------------------|---|---|
| 1 | I'll go to extremes to prevent those who I love from ever leaving me. | T | F |
| 2 | I either love someone or hate them, with nothing in between. | T | F |
| 3 | I often wonder who I really am. | T | F |
| 4 | I have tried to hurt or kill myself. | T | F |
| 5 | I am a very moody person. | T | F |
| 6 | I feel that my life is dull and meaningless. | T | F |
| 7 | I have difficulty controlling my anger, or temper. | T | F |
| 8 | When stressed, things happen. Like I get paranoid or just "black out." | T | F |
| 9 | I have done things on impulse (such as those below) that could have gotten me into trouble. | T | F |

Check all that apply to you:

- | | |
|-----------------------------------------|--------------------------|
| a. Spending more money than I have | <input type="checkbox"/> |
| b. Having sex with people I hardly know | <input type="checkbox"/> |
| c. Drinking too much | <input type="checkbox"/> |
| d. Taking drugs | <input type="checkbox"/> |
| e. Eating binges | <input type="checkbox"/> |
| f. Reckless driving | <input type="checkbox"/> |

Relationship scenarios

In the following part, we are interested in individual aspects of your romantic relationship.

For this purpose, we will show you a brief description of a hypothetical scenario that might happen in a romantic relationship. Your task is to vividly imagine this scenario as if it was real.

1) *Rejection condition*: You have breakfast with your partner and he/she tells you that he/she wants to be alone this evening and also the entire weekend.

Please imagine this scenario vividly as if it were a real event. Please describe what exactly your partner says when he/she tells you that he/she wants to be alone this evening and the entire weekend. In addition, please describe his/her voice tone, gesture, and facial expression.

Please describe what you are feeling when your partner tells you that he/she wants to be alone this evening and the entire weekend.

2) *Closeness condition*: You have breakfast with your partner and he/she tells you that he/she wants to spend this evening with you and also the entire weekend.

Please imagine this scenario vividly as if it were a real event. Please describe what exactly your partner says when he/she tells you that he/she wants to be with you this evening and the entire weekend. In addition, please describe his/her voice tone, gesture, and facial expression.

Please describe what you are feeling when your partner tells you that he/she wants to be with you this evening and the entire weekend.

3) *Control condition*: No manipulation.

How does your partner make you feel?

The following questionnaire includes several adjectives that may describe how people feel toward their romantic partners. Please indicate how your partner makes you feel on each adjectives on a scale from 1 to 9. Please focus on your current feelings toward your partner. If you choose the score 1, it means that the adjective does not at all reflect how your partner makes you feel. Conversely, if you choose the score 9, it means that the adjective very much reflects how your partner makes you feel. Please choose a number between 1 and 9 for each adjective. Again, remember to focus on your current feelings toward your partner.

My partner makes me feel...

	not at all				very much				
Distressed	1	2	3	4	5	6	7	8	9
Nervous	1	2	3	4	5	6	7	8	9
Upset	1	2	3	4	5	6	7	8	9
Relaxed	1	2	3	4	5	6	7	8	9
Content	1	2	3	4	5	6	7	8	9
Irritable	1	2	3	4	5	6	7	8	9
Afraid	1	2	3	4	5	6	7	8	9
Happy	1	2	3	4	5	6	7	8	9
Angry	1	2	3	4	5	6	7	8	9
Delighted	1	2	3	4	5	6	7	8	9
Agreeable	1	2	3	4	5	6	7	8	9

Gloomy	1	2	3	4	5	6	7	8	9
Calm	1	2	3	4	5	6	7	8	9
Excited	1	2	3	4	5	6	7	8	9

Spontaneous Emotional Responses Task Instructions and Materials (LeBel & Campbell, 2009)

In the next questions, your task will be to indicate **how much *you* like** each letter of the alphabet using the scales below and selecting the appropriate number. Please complete your judgments **quickly and try to focus on your gut feelings** toward each letter.

Scale:

not at all	1	2	3	4	5	6	7	ery much
------------	---	---	---	---	---	---	---	----------

Question before each letter:

How much do you like this letter?

G Z I B Q
P S V C J

E O N L R

D Y M K W

T X U A H

F

Please indicate your first and last name initials, without periods. For example, if your name is John Smith, you would enter "JS"

Please indicate your intimate partner's first and last name initials, without periods (or best friend if you do not have a partner). For example, if your partner's name is Olivia Brown, you would enter "OB"

Mood questionnaire Instructions and Materials (Biel/and, Dahl. Haug, & Neckelmann, 2002; Zigmond & Snaith, 1983)

In the final part, we are interested in your mood during the last year.

In the following questionnaire, please describe how you felt on *average* during the last year. Please read each sentence *very* carefully. Chose the answer that best describes how you have been feeling during the last year. You do not have to think too much. In this questionnaire, spontaneous answers are more important.

(Scale from 0 to 3)

- A I feel tense or 'wound up'
- D I still enjoy the things I used to enjoy
- A I get a sort of frightened feeling as if something awful is about to happen
- D I can laugh and see the funny side of things
- A Worrying thoughts go through my mind
- D I feel cheerful
- A I can sit at ease and feel relaxed
- D I feel as if I am slowed down
- A I get a sort of frightened feeling like 'butterflies' in the stomach
- D I have lost interest in my appearance
- A I feel restless as I have to be on the move
- D I look forward with enjoyment to things
- A I get sudden feelings of panic
- D I can enjoy a good book or radio or TV program