



## Probing commitment in individuals with borderline personality disorder

John Michael<sup>a,\*,1</sup>, Matthew Chennells<sup>b,1</sup>, Tobias Nolte<sup>c</sup>, Jinnie Ooi<sup>d</sup>, Julia Griem<sup>c</sup>, London Personality and Mood Disorder Research Network<sup>c</sup>, Wayne Christensen<sup>e</sup>, Janet Feigenbaum<sup>c</sup>, Brooks King-Casas<sup>f</sup>, Peter Fonagy<sup>c</sup>, P. Read Montague<sup>f</sup>

<sup>a</sup> Central European University, Hungary

<sup>b</sup> University of Warwick, UK

<sup>c</sup> University College London, UK

<sup>d</sup> University of East Anglia, UK

<sup>e</sup> University of Barcelona, Spain

<sup>f</sup> Virginia Tech Carilion, USA

### ARTICLE INFO

#### Keywords:

Borderline personality disorder  
Interpersonal problems  
Commitment  
Emotion regulation

### ABSTRACT

Interpersonal problems are a core symptom of borderline personality disorder (BPD). In particular, patients with BPD exhibit a heightened sensitivity to cues of acceptance or rejection in their relationships. The current study investigated the psychological processes underpinning this heightened responsiveness. In a between-subjects design, we implemented a reactivity induction designed to trigger either acceptance or rejection of a partner in two separate groups, and measured the effects which this manipulation had upon 49 patients with BPD, as well as 52 control participants. The experimental paradigm required participants to repeatedly choose whether to coordinate with their partner on a decision-making task. When both players coordinate on the same option, both are rewarded. The experiment probed participants' commitment to their partners: participants were sometimes presented with tempting opportunities to unilaterally defect from the coordination. The results show that participants in the BPD group were less committed than participants in the control group when exposed to the rejection manipulation.

### 1. Introduction

Borderline personality disorder (BPD) is a psychiatric diagnosis associated with significant psychosocial impairments, high rates of comorbidity with other psychiatric conditions<sup>1</sup>, high rates of suicide<sup>2</sup> and considerable economic costs due to intensive use of treatment and loss of productivity<sup>3</sup>. The disorder has biological and developmental roots and a lifetime prevalence of 2%–6% (Chanen and Kaess, 2012).

In addition to *affect dysregulation* and *behavioral dysregulation* (in particular impulsivity), *impairment in interpersonal functioning* has been highlighted as a core feature of psychopathology in BPD (Choi-Kain, Zanarini, Frankenburg, Fitzmaurice and Reich, 2010; Euler et al., 2019; Liebke et al., 2017; Skodol et al., 2005). Indeed, interpersonal problems have been identified as the most characteristic and discriminative feature of the disorder (Fossati et al., 1999; Gunderson, 2007; Johansen

et al., 2004). More specifically, interpersonal functioning among patients with BPD is marked by difficulties in being trusted (Hepp et al., 2018) and in trusting others (Botsford & Renneberg, B., 2020; Fertuck et al., 2019; King-Casas et al., 2008; Poggi et al., 2019), fears of abandonment (Clarkin et al., 1993; Kellett, S., Gausden and Gaskell, 2020), impairments in mentalizing (i.e., the ability to comprehend their own and others' behaviour in terms of emotions and other internal states (Fonagy et al., 2015; Luyten et al., 2020; Normann-Eide et al., 2019), and heightened responsiveness to cue of acceptance or rejection in relationships (American Psychiatric Association [APA], 2013; Cavicchioli and Maffei, 2020; Domes et al., 2009; Gratz et al., 2013; Leihener et al., 2003; Yeomans et al., 2015). Moreover, other characteristic symptoms of BPD, such as anger, affective instability, suicidal behavior, and impulsiveness, mainly manifest within interpersonal contexts (Sharp, 2016). It is therefore crucially important to develop a better

\* Corresponding author.

E-mail addresses: [Johnmichael.cogsci@gmail.com](mailto:Johnmichael.cogsci@gmail.com) (J. Michael), [M.Chennells@warwick.ac.uk](mailto:M.Chennells@warwick.ac.uk) (M. Chennells), [Tobias.NolteMD@annafreud.org](mailto:Tobias.NolteMD@annafreud.org) (T. Nolte), [Jinnie.Ooi@uea.ac.uk](mailto:Jinnie.Ooi@uea.ac.uk) (J. Ooi), [julia.griem.14@ucl.ac.uk](mailto:julia.griem.14@ucl.ac.uk) (J. Griem), [Tobias.NolteMD@annafreud.org](mailto:Tobias.NolteMD@annafreud.org) ( ), [wayne.christensen@gmail.com](mailto:wayne.christensen@gmail.com) (W. Christensen), [j.feigenbaum@ucl.ac.uk](mailto:j.feigenbaum@ucl.ac.uk) (J. Feigenbaum), [bkcasas@vtc.vt.edu](mailto:bkcasas@vtc.vt.edu) (B. King-Casas), [p.fonagy@ucl.ac.uk](mailto:p.fonagy@ucl.ac.uk) (P. Fonagy), [read@vtc.vt.edu](mailto:read@vtc.vt.edu) (P.R. Montague).

<sup>1</sup> Joint first authors.

<https://doi.org/10.1016/j.jpsychires.2021.02.062>

Received 2 December 2020; Received in revised form 24 February 2021; Accepted 26 February 2021

Available online 11 March 2021

0022-3956/© 2021 The Author(s).

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

understanding of the psychological processes underpinning the problems which patients with BPD experience in their interpersonal relations (Cristea et al., 2017).

To address this challenge, an emerging body of work has begun to explore the psychological processes underlying impaired interpersonal functioning in BPD. Specifically, researchers have increasingly adopted experimental approaches to investigating the behaviour and attitudes of patients with BPD within experimentally controlled social interactions designed to probe particular aspects of interpersonal functioning (Fineberg et al., 2018; Lis and Bohus, 2013 for a review). One approach, which is promising in view of the centrality of themes of acceptance and rejection in the interpersonal experience of patients with BPD, has been to probe the behaviour of BPD patients using the trust game (Berg et al., 1995). In a trust game, one participant (the ‘investor’) is given a small sum of money and can choose to transfer any portion of it to a second participant (the ‘trustee’). Then, the trustee receives triple the amount transferred, and can choose to return any portion of the money back to the investor. The trust game may be extended over multiple rounds, giving both players an incentive to transfer sufficient money to maintain trust and cooperation over time and thereby maximize their rewards (Montague et al., 2015). Using this paradigm, Unoka et al. (2009) found that patients with BPD (compared to controls) transferred less money when playing in the role of the investor and expected the trustees to return a smaller portion of the money. In a separate study, it was found that patients with BPD also showed difficulty in maintaining cooperation over multiple rounds of the game. King-Casas et al. (2008; see also Liebke et al., 2018; Fertuck et al., 2019; Abramov et al., 2020).

Taken together, these studies provide evidence that behavioral patterns characteristic of BPD can be elicited and measured using economic games such as the trust game. It is noteworthy however that patients with BPD do not *always* exhibit irregularities with respect to their willingness to trust others or in their approach to social interactions and relationships (Hepp et al., 2018). Rather, these symptoms are exhibited only when triggered, e.g. when social expectations are disappointed or when they feel rejected by another person (Levy, 2005; Doell et al., 2020). Moreover, it is also noteworthy that irregularities in the interpersonal functioning of patients with BPD are not always *negative* (e.g. demonstrating a lack of trust), but sometimes highly positive – e.g. when they perceive themselves to be accepted by another person (American Psychiatric Association [APA], 2013; Domes et al., 2009; Liebke et al., 2018; Liebke et al., 2018).

To gain a fuller understanding of the psychological processes underlying the problems in interpersonal functioning experienced by patients with BPD, it is therefore crucial to investigate patients’ differential responsiveness to cues of acceptance or rejection (Cavicchioli and Maffei, 2020; Gratz et al., 2013). To this end, we endeavored to elicit either acceptance or rejection within a controlled lab setting (in two separate groups), and to measure the effects which this reactivity induction had upon participants with BPD, compared to a control group.

In particular, we aimed to measure the effects of a reactivity induction upon participants’ degree of commitment to a partner. We focus on commitment as a prosocial attitude which serves to stabilize healthy relationships: in order to maintain healthy relationships, it is important to be committed in the sense of being willing to sacrifice time, effort, money and other resources (Michael et al., 2016; Ooi et al., 2018). This focus on commitment enables us to build upon the aforementioned research investigating trust in BPD. Trust and commitment are distinct constructs, but we propose that they are dynamically interrelated – i.e. that people tend to increase commitment towards partners who exhibit trustworthiness, and to withdraw commitment from partners who prove untrustworthy. And indeed, the heightened responsiveness to cues of acceptance and rejection which are characteristic of BPD may be driven in part by the dynamic interaction of commitment and trust: if a partner exhibits a level of commitment which is higher than expected, this may elicit trust, and consequently also a high level of commitment (i.e., as a response to perceived acceptance). On the other hand, if a partner

exhibits a level of commitment that is lower than expected, this may trigger mistrust, and consequently elicit a withdrawal of commitment (i.e., as a response to perceived rejection).

## 2. Method

### 2.1. Experimental approach

The experiment was designed to probe participants’ commitment to cooperative interactions with a partner *independently of trust*. To this end, we implemented a sequential joint decision-making task in which participants could choose whether or not to cooperate with a partner. We varied whether and to what degree the option not to cooperate constituted a temptation and measured the frequency with which participants chose to cooperate despite this temptation (cooperation rates). Crucially, the choices made by their partners could not affect them negatively, and participants were informed that their partners would receive no feedback about their choices. This ensured that participants’ willingness to cooperate could only be explained by commitment to their partner, not by trust or by any expectation of reciprocity.

Further, in a between-subjects design, once participants had made a choice as to their preferred interaction partner, we manipulated the induction process by which they were paired with a partner: in one condition, participants were paired with the partner they had chosen and told their choice was required – i.e. their partner had likewise chosen them as their preferred interaction partner (Acceptance Condition); in another condition, participants were paired with the partner they had chosen but, instead, told their choice was unrequired – i.e. their partner had *not* chosen them as their preferred interaction partner (Rejection Condition).

We predicted that patients with BPD in the Acceptance group would exhibit greater commitment than controls, leading to higher cooperation rates, whereas patients with BPD in the Rejection group would exhibit less commitment than controls, leading to lower cooperation rates. In other words, our main prediction was an interaction between Group (Patient versus Control) and reactivity induction (Acceptance versus Rejection), such that the Reactivity Induction would have a more pronounced *positive or negative* effect on patients compared to controls.

We also recorded reaction times (RTs) as well as movements of the computer mouse, enabling us to scrutinize participants’ decision-making processes at a finer-grain. We predicted that, when choosing to cooperate (i.e. to maintain commitment), patients would exhibit shorter RTs and more direct mouse trajectories than controls in the Acceptance group, indicating less inner conflict while making the decision. In contrast, we predicted that patients would exhibit longer RTs and less direct mouse trajectories than controls in the Rejection group, revealing greater inner conflict. We also predicted that the reactivity induction would have the largest impact for those participants with the most pronounced profiles of BPD traits.

The hypotheses, sample size, methods, exclusion criteria and planned analyses were pre-registered before data collection, and can be accessed at: <https://aspredicted.org/blind.php?x=bq8pn7>. All aspects of the study were carried out in accordance with the pre-registered protocol unless otherwise stated.

### 2.2. Participants

We recruited 101 participants, of which 49 were patients with a diagnosis of BPD (BPD Patient Group), and 52 were healthy control participants (Control Group). Of these, 52 participants were allocated to the Acceptance treatment condition and 49 to the Rejection condition. Additional descriptive statistics for the sample are provided in Table 2. Participants with BPD were recruited through seven different NHS Mental Health Trusts across London. They had a suspected or confirmed diagnosis of BPD according to the clinical assessments conducted in the relative trusts. Once interested individuals were consented and enrolled

**Table 2**  
Descriptive statistics of the study sample.

		BPD Patient Group	Control Group
Treatment condition	Acceptance	23	29
	Rejection	26	23
	Total	49	52
Gender	Females	38	37
	Males	8	15
	Prefer not to say	5	0
Age in years	Mean	29.57	30.4
	SD	8.59	11.55
	Range	18–53	18–59

in the study, trained and experienced researchers conducted the SCID-II interview (First et al., 1994) to confirm the extent of BPD symptomatology according to the DSM-IV-TR. Regular supervision and consensus meetings were held to ensure inter-rater reliability. Participants were excluded if they had serious comorbid mental disorders (bipolar disorder, schizophrenia), and if they had a recent psychotic episode. Healthy Control participants were recruited through public and online advertising. We pre-screened all potential HC participants prior to their participation in the study and excluded individuals who were currently diagnosed with or had a suspected mental illness. Next, all participants completed the SAPAS screening questionnaire, and if the score was 4+, then a SCID-II was conducted to check that no personality disorder was above threshold. The majority of the HC participants in this sample however still received the BPD and ASPD sections of the SCID regardless of their SAPAS score as this added another layer of control and was of potential interest for other parts of the study. All participants reported speaking and understanding English and participants provided their informed written consent prior to the testing. Ethics clearance for the experiment was obtained from Research Ethics Committee 3 for Wales, REC ref 12/WA/0283 and the research was conducted in accordance with the Declaration of Helsinki.

For each group (BPD/HC), we assigned numbers 1–50 to the two conditions (accept or reject) in an alternating manner. Whenever a new participant joined the study, they received the number that was next up in the sequence, and this determined their condition. This may have resulted in the conditions being unevenly distributed between the two groups – first, due to matching and/or missing data; and second, in some instances participants discontinued the study after their number had already been assigned to them but before they were able to complete the game.

2.3. Apparatus and stimuli

The program for the experiment was written in Javascript using the jsPsych toolbox (de Leeuw, J. R. (2015)). The two choice options were

presented as rectangular fields, separated and placed at opposite ends of the screen. The mouse start field was positioned at the bottom of the screen and at the midpoint between the two choice options. During trials, participants used the mouse to select by clicking on one of their choice options. The computer screen provided participants with real-time visual feedback on their inputs. The computer monitor used was consistent across participants and experiment sessions.

2.4. Design

We implemented a 2 between (Group: BPD patient vs. Control) x 2 between (Reactivity induction: Acceptance vs. Rejection) x 2 within (Partner block: Baseline vs. Treatment) design.

2.5. Procedure

The experiment consisted of two baseline and two treatment phases, with four phases in total (Fig. 1). Prior to beginning the two experiments, participants created their own avatars (see Supplementary Materials 2 for details) which ‘represented’ them in their future partner interactions; i.e. ostensible ‘partners’ saw these avatars when making choices of partners and, likewise, they saw their partners’ avatars. This enabled us to implement the Reactivity Induction (See below).

2.5.1. Baseline

Phases 1 and 2 involved a baseline phase with a new partner and the performance and completion of a baseline round of the joint Commitment Game (explained below and in Supplementary Materials 1). In the first phase, participants were told they had been randomly matched with a partner with whom they would subsequently perform a joint task for points. In reality, partners were virtual pre-programmed agents and all participants were ‘matched’ with the same partner profile. The purpose of the baseline phase was to establish participants’ cooperative behavior in the absence of a specific partner relationship.

2.5.2. Commitment Game

Participants performed 36 trials per block of the commitment game, a sequential joint decision-making task (see Fig. 2)

2.5.3. Reactivity induction treatment

Phases 3 and 4 involved an induction phase followed by the commitment game, but with a new partner. Crucially, while the commitment game was identical to that performed with the first partner, the induction phase with the second partner differed from the baseline with the first partner by involving the following procedure: participants were presented with five partner profiles and asked to select which of the possible partners they would prefer to perform a second round of the

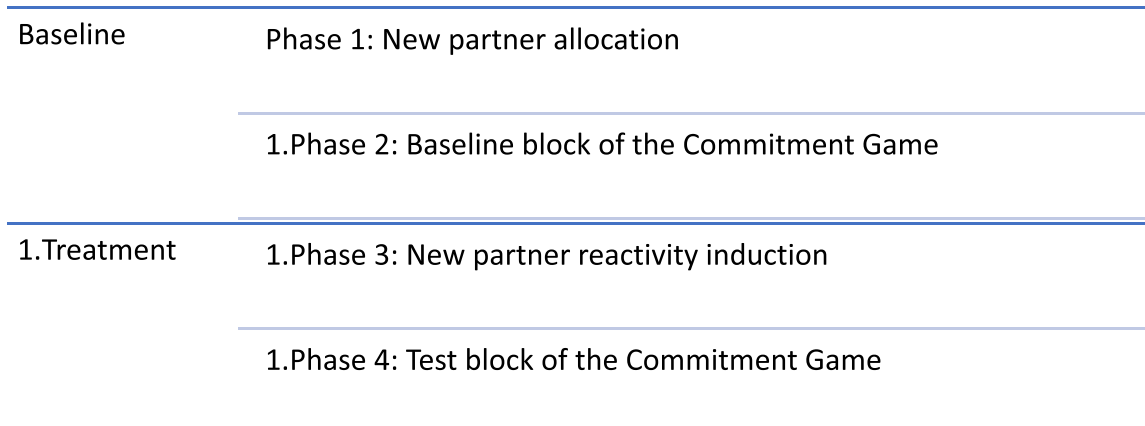
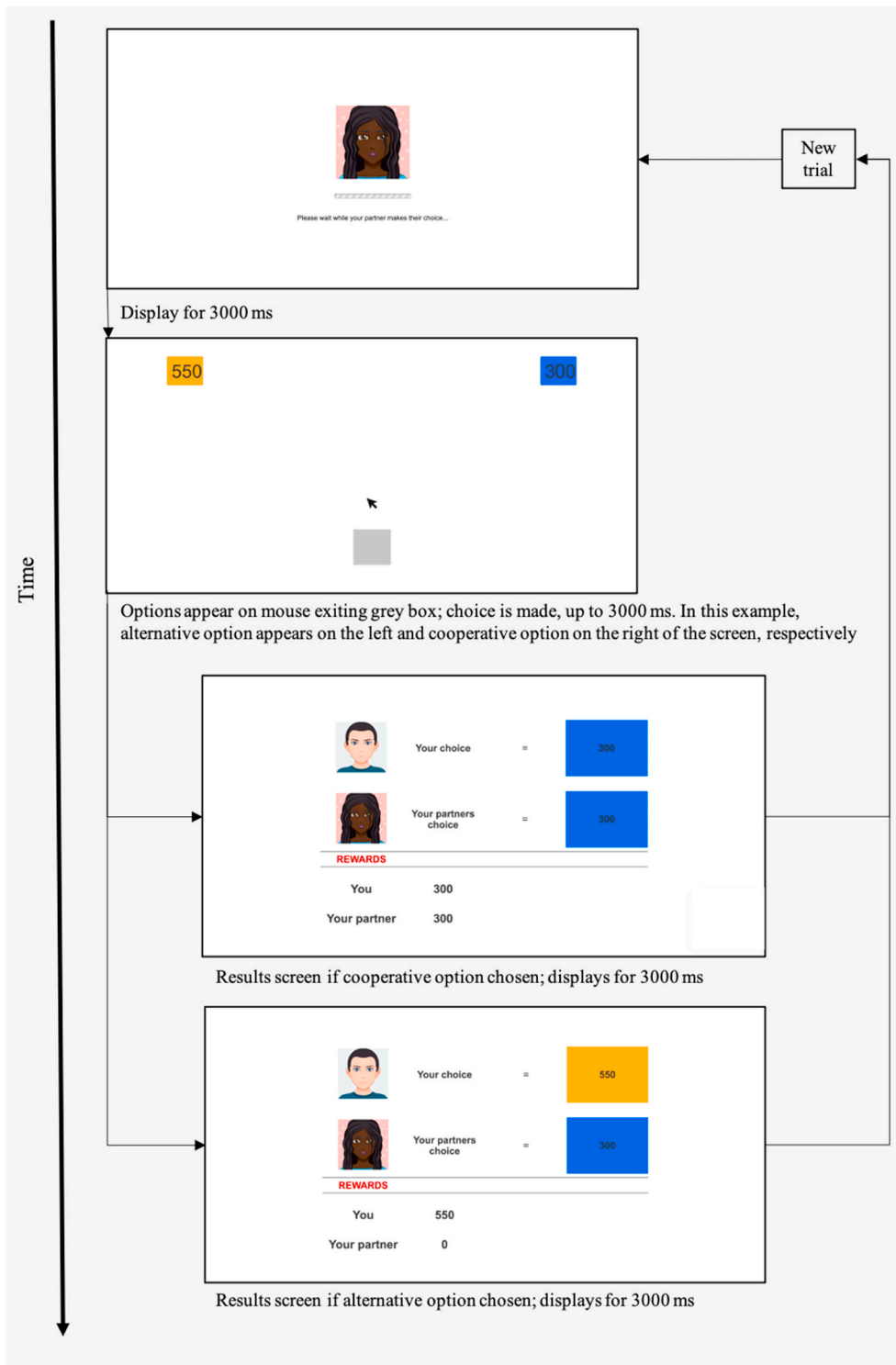


Fig. 1. The procedure.



**Fig. 2.** Trial Structure. At the beginning of each trial, an image of the partner’s avatar was displayed. Then, the partner chose one of two values. The participant did not see what the partner had chosen, but was then herself presented with two values to choose between. One of these, indicated in green or blue, was the same value that the partner had chosen (cooperative option); the other, indicated in orange, was an alternative value (alternative option). If the participant chose the alternative option, then the participant received the amount corresponding to the alternative option and their partner received no reward. Conversely, if the participant chose the cooperative option, then each received the corresponding amount. We varied whether- and to what extent the alternative option constituted a temptation, i.e. by varying its value relative to the cooperative option. In both baseline and treatment blocks, participants performed 8 practice trials followed by 36 trials. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

commitment game with. Participants were asked to choose up to three preferred partners before proceeding and were informed that possible partners were completing their own identical selection process.

Thereafter, we introduced a reactivity induction involving the following two conditions: for approximately half of the participants, the reactivity induction was designed to trigger an Acceptance of the partner (Acceptance Group), whereas for the other half it was designed to trigger a Rejection of the partner (Rejection Group) (for details on how this was presented to participants, see [Supplementary Materials 1](#)). In the Acceptance Group, participants were randomly matched with one of

their preferred partners and told that their matched partner had also chosen them as their preferred interaction partner. In the Rejection Group, participants were similarly matched with one of their preferred partners but told, instead, that their partner had *not* chosen them as their preferred interaction partner.

After completing the reactivity induction phase, participants proceeded to play a second and final round of the commitment game (36 trials) with the new partner. Once completed, they were thanked for their participation and shown their payout, automatically calculated based on their behavior during the two commitment games.

2.5.4. Analyses

Our predictions focused on behavioral differences in commitment between groups and treatment conditions. We operationalized commitment using several measures and used these as dependent variables in our analysis.

For our primary analysis, our measure of focus was participant cooperation rate. We test our predictions using an ANOVA (Type 3 tests, LRT-method) on a mixed-effects logistic regression model on participants' trial-by-trial decision whether to cooperate (DV = 1) or not (DV = 0). We ran this using the *afex* package in R (Singman et al., 2020) and used effects coding (Singmann and Kellen, 2019) rather than treatment contrasts to ensure correct measurement of main effects. To explore the results of this model further, we also conducted post-hoc pairwise comparisons of cooperation rates between Rejection and Acceptance induction conditions, for both Control and Patient groups. For these tests, cooperation rate is calculated for each individual as the change in the proportion of all trials in which a participant chose the cooperative option between baseline and treatment blocks.

For our secondary analysis, we used mousetracking data to collect several measures related to participants' movement of the computer mouse, which was used to begin trials and make choices (For details and results, see Supplementary Materials 2).

In all regression models we include three independent variables of interest. First, a between-subject indicator of Treatment Group equal to 1 if the subject was in the BPD Patient Group and 0 if in the Control Group. Second, a between-subject indicator for Induction Reactivity Condition, equal to 1 and 0 if the subject was in the Rejection and Acceptance treatment reactivity condition, respectively. Third, a within-subject indicator equal to 1 if the subject's trial in question was in the Treatment Phase and 0 if in the Baseline Phase. In addition, we control for but do not interpret the following covariates in our model: a random effect to allow the intercept to vary by subject; and the trial-by-trial Temptation Level, a re-scaled variable, allowed to randomly vary by

subject, indicating the attractiveness of the subject's alternative option.

We made two deviations from our pre-registered analysis plan. First, we excluded participants' age and gender as covariates, as the inclusion of these variables fails to allow any of our regression models to converge. We note that we made no predictions regarding the effects of these variables on cooperation rates. Further, to support our decision to exclude these variables, we conducted a simple regression of age and gender on cooperation choices and found no effects of either. Secondly, we decided not to include an analysis of earnings (points gained) by condition. This was because we determined that this additional analysis would not add any information over and above the analysis of cooperation rates, as earnings are closely determined by whether a participant cooperated or not.

3. Results

Fig. 3 shows participant cooperation rates from the Commitment Game. In our regression model on cooperation rate (that is, participants' trial-by-trial decision whether to cooperate or not), the interaction term *Induction Condition x Treatment Phase* shows a significant overall treatment effect ( $\chi^2 = 14.69, p < .001$ ), indicating that, overall, the Rejection condition is associated with lower cooperation rates than the Acceptance condition. Further, the results on the interaction term *Group x Induction Condition x Treatment Phase* corroborate our prediction that this effect on cooperation rates between Acceptance and Rejection conditions was significantly greater ( $\chi^2 = 5.52, p = .019$ ) for the BPD Patient group than the Control group.

Though our analysis focuses on differences between treatment and baseline phases, for robustness we checked for any differences in baseline cooperation rates that may indicate important, unseen factors driving our findings. We therefore limited the same regression model to include only baseline trials, but found no effect of either Group ( $\chi^2 = 0.40, p = .527$ ) or Induction condition ( $\chi^2 = 0.18, p = .670$ ) on

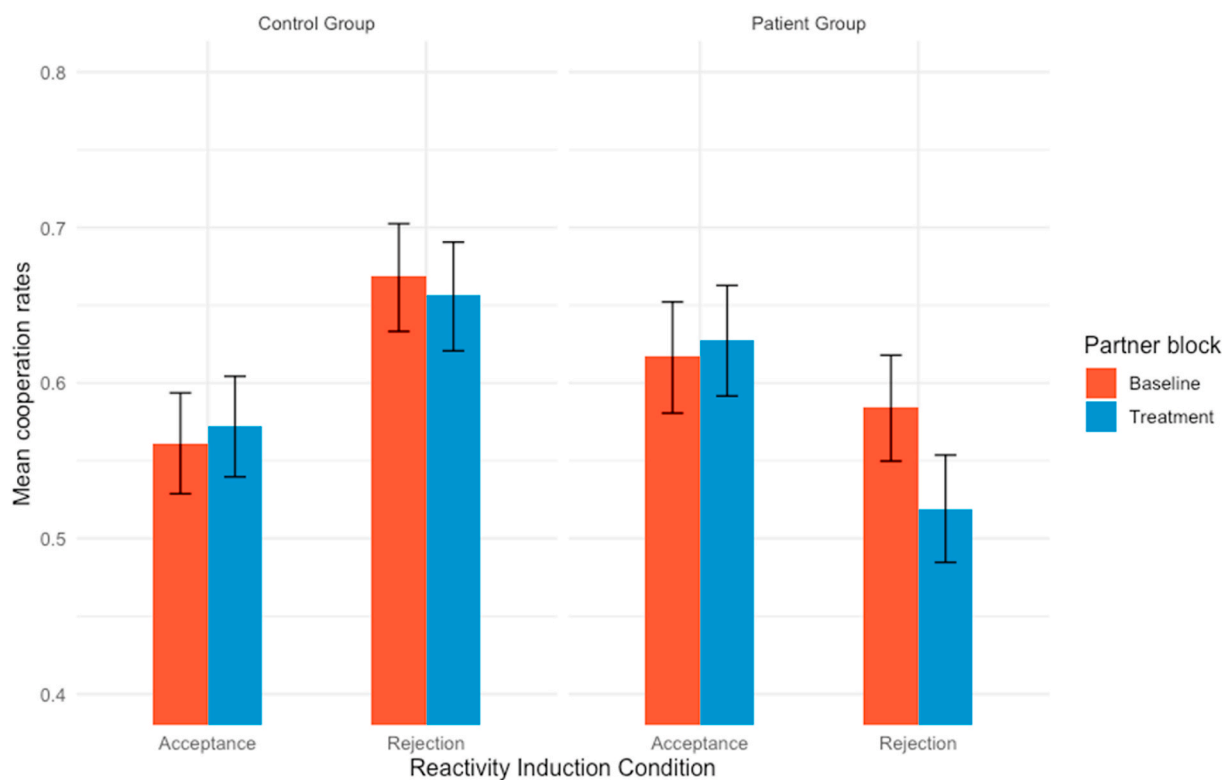


Fig. 3. Cooperation Rates by Condition. Comparing mean cooperation rates in baseline and induction treatment phases of the Commitment Game. The baseline phase precedes the treatment phase and involves partners automatically allocated to participants with no choosing process. Error bars represent 95% confidence intervals.

cooperation rates, nor an effect of their interaction ( $\chi^2 = 2.04, p = .154$ ).

For our post-hoc analysis, we first tested the data for normality and homogeneity of variance by conducting Shapiro-Wilk tests on participants' change in mean cooperation rate for each of our four groups, revealing evidence of only limited deviations in normality: Control group in Rejection condition,  $p = 0.214$ ; Control group in Acceptance condition,  $p = 0.135$ ; Patient group in Acceptance condition,  $p = 0.077$ ; and Patient group in Rejection condition,  $p < 0.000$ . Given these findings, and given our unequal sample sizes and no assumption of equal variances between groups, we therefore conducted post-hoc pairwise comparisons using two-tailed Welch's Two Sample t-tests with Bonferroni correction testing the equality of means between induction conditions for each group. Effect sizes are given by calculating Cohens  $d$  (Cohen, 1988). Participants in the BPD Patient group cooperated significantly less in the Rejection condition ( $Mean = -0.065$ ) than patients in the Acceptance condition ( $Mean = 0.011$ ),  $t(45.56) = 2.41, p = .039, d = 0.68$ . Participants in the Control group, however, showed no significant difference in cooperation between the Rejection condition ( $Mean = -0.012$ ) than the Acceptance condition ( $Mean = 0.011$ ),  $t(46.40) = 0.743, p = .923, d = 0.21$ .

#### 4. Discussion

As predicted, the reactivity induction had a greater impact on patients with BPD than on controls. Patients in the Rejection condition (but not controls in the Rejection condition) exhibited significantly lower cooperation rates than in the baseline condition. We did not however observe a significant increase in cooperation rates in the Acceptance condition for patients or controls.

Analyses of the RTs and of the motion data were consistent with this picture, although these results were not statistically significant: there were longer RTs and greater AUC for patients when defecting than when cooperating in the Acceptance condition, and greater conflict when cooperating than when defecting in the Rejection condition. The high variability in this data raises the possibility that this measure was not sufficiently sensitive to detect any potential differences between conditions or groups.

Taken together, the findings suggest a specific deficit in BPD patients' ability to resolve complex social scenarios in which the degree of commitment to a new partner must be calibrated – in particular when a partner is perceived to have exhibited a lower degree of commitment. This would be consistent with the hypothesis that BPD patients generally expect less commitment from partners (i.e. they have lower priors on partner commitment), and are therefore quick to infer from scant evidence that their partners are in fact disloyal to them, and consequently to devalue the corresponding relationships. Such a tendency may be compounded by a deficit in the ability to mentalize, and accordingly to consider generous, benign explanations of their partners' behaviour (Fonagy et al., 2015; Luyten et al., 2020; Normann-Eide et al., 2019). These findings corroborate and build upon earlier findings suggesting a link between BPD traits and disruptions of the sense of commitment (Ooi et al., 2018).

The current research also extends previous research in several other important ways. First, the innovative reactivity induction which we implemented here makes it possible to elicit, in a controlled settings, the irregularities in interpersonal functioning which are characteristic of BPD – but which are not always exhibited, and which can therefore (in the absence of such a reactivity induction) elude experimental investigation (Lazarus et al., 2014). Relatedly, our reactivity induction makes it possible to investigate positive as well as negative irregularities in the interpersonal functioning of patients with BPD. This is important insofar as these responses may be triggered by heightened responsiveness to cues of acceptance and rejection (American Psychiatric Association [APA], 2013; Cavicchioli and Maffei, 2020; Domes et al., 2009; Gratz et al., 2013).

The experimental design developed here is also innovative insofar as

it makes it possible to isolate participants' degree of commitment to a partner, and test for this independently of trust. We believe that commitment is an important construct which is distinct from, but dynamically related to, trust: people tend to increase commitment towards partners who exhibit trustworthiness, and to withdraw commitment from partners who prove untrustworthy. And indeed, the extreme responses to acceptance and rejection which are characteristic of BPD may be driven in part by the dynamic interaction of commitment and trust: if a partner exhibits a level of commitment which is higher than expected, this may elicit trust, and consequently also a high level of commitment (a response to perceived acceptance). On the other hand, if a partner exhibits a level of commitment that is lower than expected, this may trigger mistrust, and consequently elicit a withdrawal of commitment (a response to perceived rejection). Both processes require the realistic appraisal of the other and of one's relationship with them. Future research may benefit from linking these with individual differences in mentalizing capacities.

It is important to acknowledge that the sample size may not have been large enough to capture the effects of our manipulations, in particular given that we had two between-groups factors. It is therefore important for future research to incorporate larger sample sizes, or if this is not feasible, to attempt to implement within-group or mixed designs.

The current research opens up a new perspective from which to consider the interpersonal problems faced by patients with BPD. In particular, if a heightened sensitivity to cues of rejection leads patients to withdraw commitment from interpersonal relationships, they may be unwilling to invest time, effort or other resources necessary to repair their relationships. This, in turn, may cause others to turn away from them. The pattern of results mirrors the observations of King-Casas et al. (2008) where the greater reluctance to engage in post-rupture repair distinguished patients with BPD from controls. This pattern of fragility in the sense of commitment is likely to serve patients with BPD poorly in the everyday ups and downs that characterize relationships. For example, a mild rupture initiated by a clinician (e.g. forgetting a personal detail, a slight delay in expressing empathy about an adverse experience) may trigger a dramatic withdrawal of commitment to the relationship with the therapist. Of course, if the therapist is naive to this vulnerability, they may react to such a dramatic response as if it carried a serious indication of hostility rather than of a specific social-cognitive deficit, deepening the rift in the therapeutic relationship. We hope that future research may further illuminate the many ways in which a disruption in the sense of commitment may contribute to the interpersonal difficulties of patients with BPD, both in their personal lives and in therapeutic relationships.

#### Author contributions

The study design was developed by J. Michael, M. Chennells, T. Nolte, J. Ooi, W. Christensen, W., P. Fonagy, and the London Personality and Mood Disorder Research Network. The data collection was carried out by J. Griems. The analysis was carried out by M. Chennells, under the supervisions of J. Michael, T. Nolte, T. B. King-Casas, P.R. Montague, and P. Fonagy. J. Michael drafted the manuscript. All authors contributed revisions and approved the final version of the manuscript.

#### Declaration of competing interest

The authors declare no conflicts of interests.

#### Acknowledgment

This research was supported by a Starting Grant (679092. Sense of Commitment) awarded by the European Research Council to John Michael.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jpsychires.2021.02.062>.

## References

- Abramov, G., Miellel, S., Kautz, J., Grenyer, B.F.S., Deane, F.P., 2020. The paradoxical decline and growth of trust as a function of borderline personality disorder trait count: using discontinuous growth modelling to examine trust dynamics in response to violation and repair. *PLoS One* 15 (7), e0236170. <https://doi.org/10.1371/journal.pone.0236170>.
- American Psychiatric Association, 2013. *Diagnostic and Statistical Manual of Mental Disorders*, fifth ed. American Psychiatric Publishing, Arlington, VA.
- Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, reciprocity, and social history. *Game Econ. Behav.* 10 (1), 122–142.
- Botsford, J., Renneberg, B., 2020. Autobiographical memories of interpersonal trust in borderline personality disorder. *Borderline Personal Disord. Emot. Dysregul* 7, 14. <https://doi.org/10.1186/s40479-020-00130-w>.
- Cavicchioli, M., Maffei, C., 2020. Rejection sensitivity in borderline personality disorder and the cognitive–affective personality system: a meta-analytic review. *Personality Disord.: Theor. Res. Treat.* 11 (1), 1.
- Chanen, A.M., Kaess, M., 2012. Developmental pathways to borderline personality disorder. *Curr. Psychiatr. Rep.* 14 (1), 45–53.
- Clarkin, J.F., Hull, J.W., Hurt, S.W., 1993. Factor structure of borderline personality disorder criteria. *J. Pers. Disord.* 7 (2), 137–143.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*, second ed. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Cristea, I.A., Gentili, C., Cotet, C.D., Palomba, D., Barbui, C., Cuijpers, P., 2017. Efficacy of psychotherapies for borderline personality disorder: a systematic review and meta-analysis. *JAMA Psychiatr.* 74, 319–328. <https://doi.org/10.1001/jamapsychiatry.2016.4287>.
- de Leeuw, J.R., 2015. jsPsych: a JavaScript library for creating behavioral experiments in a web browser. *Behav. Res. Methods* 47 (1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>.
- Doell, K.C., Olić, E., Courtet, P., Corradi-Dell'Acqua, C., Perroud, N., Schwartz, S., 2020. Atypical processing of social anticipation and feedback in borderline personality disorder. *Neuroimage: Clinical* 25, 102126.
- Domes, G., Schulze, L., Herpertz, S.C., 2009. Emotion recognition in borderline personality disorder—a review of the literature. *J. Pers. Disord.* 23, 6–19. <https://doi.org/10.1521/pedi.2009.23.1.6>.
- Euler, S., Nolte, T., Constantinou, M., Griem, J., Montague, P.R., Fonagy, P., 2019. & Personality and Mood Disorders Research Network.
- Fertuck, E.A., Grinband, J., Mann, J.J., Hirsch, J., Ochsner, K., Pilkonis, P., et al., 2019. Trustworthiness appraisal deficits in borderline personality disorder are associated with prefrontal cortex, not amygdala, impairment. *Neuroimage Clin.* 21, 101616. <https://doi.org/10.1016/j.nicl.2018.101616>.
- Fineberg, S.K., Leavitt, J., Stahl, D.S., Kronemer, S., Landry, C.D., Alexander-Bloch, A., Corlett, P.R., 2018. Differential valuation and learning from social and nonsocial cues in borderline personality disorder. *Biol. Psychiatr.* 84 (11), 838–845. <https://doi.org/10.1016/j.biopsych.2018.05.020>.
- Fonagy, P., Luyten, P., Bateman, A.W., 2015. Translation: mentalizing as treatment target in borderline personality disorder. *Personality Disord.: Theor. Res. Treat.* 6, 380–392.
- Fossati, A., Maffei, C., Bagnato, M., Donati, D., Namia, C., Novella, L., 1999. Latent structure analysis of *DSM-IV* borderline personality disorder criteria. *Compr. Psychiatr.* 40, 72–79.
- Gratz, K.L., Dixon-Gordon, K.L., Breetz, A., Tull, M., 2013. A laboratory-based examination of responses to social rejection in borderline personality disorder: the mediating role of emotion dysregulation. *J. Pers. Disord.* 27 (2), 157–171. Chicago.
- Gunderson, J.G., 2007. Disturbed relationships as a phenotype for borderline personality disorder. *Am. J. Psychiatr.* 164, 1637–1640. <https://doi.org/10.1176/appi.ajp.2007.07071125>.
- Hepp, J., Storkel, L.M., Kieslich, P.J., Schmahl, C., Niedtfeld, I., 2018. Negative evaluation of individuals with borderline personality disorder at zero acquaintance. *Behav. Res. Ther.* 111, 84–91. <https://doi.org/10.1016/j.brat.2018.09.009>.
- Johansen, M., Karterud, S., Pedersen, G., Gude, T., Falkum, E., 2004. An investigation of the prototype validity of the borderline *DSM-IV* construct. *Acta Psychiatr. Scand.* 109, 289–298.
- 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.
- Kellett, S., Gausden, J., Gaskell, C., 2020. The effectiveness of cognitive analytic therapy for borderline personality disorder: utilizing a withdrawal experimental design to improve sensitivity to abandonment. *Psychol. Psychother.* <https://doi.org/10.1111/papt.12278>.
- Lazarus, S.A., Cheavens, J.S., Festa, F., Rosenthal, M.Z., 2014. Interpersonal functioning in borderline personality disorder: a systematic review of behavioral and laboratory-based assessments. *Clin. Psychol. Rev.* 34 (3), 193–205.
- Leihener, F., Wagner, A., Haaf, B., Schmidt, C., Lieb, K., Stieglitz, R., Bohus, M., 2003. Subtype differentiation of patients with borderline personality disorder using a circumplex model of interpersonal behavior. *J. Nerv. Ment. Dis.* 191 (4), 248–254. <https://doi.org/10.1097/01.NMD.0000061150.38924.2A>.
- Levy, K.N., 2005. The implications of attachment theory and research for understanding borderline personality disorder. *Dev. Psychopathol.* 17 (4), 959–986.
- Liebke, L., Koppe, G., Bungert, M., Thome, J., Hauschild, S., Defiebre, N., Lis, S., 2018. Difficulties with being socially accepted: an experimental study in borderline personality disorder. *J. Abnorm. Psychol.* 127 (7), 670–682. <https://doi.org/10.1037/abn0000373>.
- Liebke, L., Bungert, M., Thome, J., Hauschild, S., Gescher, D.M., Schmahl, C., et al., 2017. Loneliness, social networks, and social functioning in borderline personality disorder. *Personality Disord.: Theor. Res. Treat.* 8, 349–356.
- Lis, S., Bohus, M., 2013. Social interaction in borderline personality disorder. *Curr. Psychiatr. Rep.* 15 (2), 338–345. <https://doi.org/10.1007/s11920-012-0338-z>.
- Luyten, P., Campbell, C., Allison, E., Fonagy, P., 2020. The mentalizing approach to psychopathology: state of the art and future directions. *Annu. Rev. Clin. Psychol.* <https://doi.org/10.1146/annurev-clinpsy-071919-015355>.
- Michael, J., Sebanz, N., Knoblich, G., 2016. The sense of commitment: a minimal approach. *Front. Psychol.* 6, 1968.
- Montague, P.R., Lohrenz, T., Dayan, P., 2015. The three R's of trust. *Curr. Opin. Behav. Sci.* (3), 102–106.
- Normann-Eide, E., Antonsen, B.T., Kvarstein, E.H., Pedersen, G., Vaskinn, A., Wilberg, T., 2019. Are impairments in theory of mind specific to borderline personality disorder? *J. Pers. Disord.* 1–15. <https://doi.org/10.1521/pedi.2019.33.417>.
- Ooi, J., Francová, A., Székely, M., Michael, J., 2018. The sense of commitment in individuals with borderline personality traits in a non-clinical population. *Front. Psychiatr.* 9, 519.
- Poggi, A., Richetin, J., Preti, E., 2019. Trust and rejection sensitivity in personality disorders. *Curr. Psychiatr. Rep.* 21 (8), 69.
- Sharp, C., 2016. Current trends in BPD research as indicative of a broader sea-change in psychiatric nosology. *Personality Disord.: Theor. Res. Treat.* 7, 334–343.
- Singmann, H., Kellen, D., 2019. An introduction to mixed models for experimental psychology. In: Spieler, D.H., Schumacher, E. (Eds.), *New Methods in Cognitive Psychology*. Psychology Press, pp. 4–31.
- Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M.S., 2020. *Afex: Analysis Of Factorial Experiments*. R Package Version 0.27-2. <https://CRAN.R-project.org/package=afex>.
- Skodol, A.E., Oldham, J.M., Bender, D.S., Dyck, I.R., Stout, R.L., Morey, L.C., et al., 2005. Dimensional representations of *DSM-IV* personality disorders: relationships to functional impairment. *Am. J. Psychiatr.* 162, 1919–1925.
- Unoka, Z., Seres, I., Áspán, N., Bódi, N., Kéri, S., 2009. Trust game reveals restricted interpersonal transactions in patients with borderline personality disorder. *J. Pers. Disord.* 23 (4), 399–409.
- Yeomans, F.E., Clarkin, J.F., Kernberg, O.F., 2015. *Transference-focused Psychotherapy for Borderline Personality Disorder: A Clinical Guide*.