Neuroscience and Biobehavioral Reviews 71 (2016) 849–864

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Rationality and self-interest as economic-exchange strategy in borderline personality disorder: Game theory, social preferences, and interpersonal behavior

Haang Jeung [a,*], Christiane Schwieren [b], Sabine C. Herpertz [a]

[a] Department of General Psychiatry, Heidelberg University Hospital, Vossstr. 4, 69115 Heidelberg, Germany
[b] Alfred-Weber-Institute of Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany

ARTICLE INFO

Article history:
Received 15 March 2016
Received in revised form
12 September 2016
Accepted 31 October 2016
Available online 5 November 2016

Keywords:
Borderline Personality Disorder (BPD)
Game theory
Homo economicus
Social interaction

ABSTRACT

Borderline Personality Disorder (BPD) is characterized by severe and persistent impairments in interpersonal functioning. Given the complexity of social interactions, studying the interactive behavior of BPD patients is challenging. One way to implement both tight experimental control and realistic, externally valid settings is to use game-theoretical experiments. This review discusses findings from economic exchange studies in BPD against the background of game-theoretical literature. BPD patients do not seem to derive utility from mutual cooperation with others and appear not to "forgive" a partner's unfairness. By pursuing a strategy of negative reciprocity, BPD patients seem to act mostly "rationally" and in their own self-interest. Their "grim trigger strategy" resembles the theoretical ideal of the rational and self-interested agent *homo economicus*. Finally, we summarize how research findings from economics and clinical psychiatry may be mutually enriching and propose new research ideas in this fascinating field.

Contents

1. Introduction

Borderline Personality Disorder (BPD) is a severe mental disorder that occurs in about 1% of the population (Paris, 2010). Individuals with BPD experience negative affect and poor self-concept, act impulsively, and engage in unstable relationships (Lieb et al., 2004). Moreover, according to the largest epidemiological study of mental disorders in U.S. adults, about 85% of individuals with BPD are likely to have co-occurring lifetime mental disorders such as anxiety disorders, major depression, substance abuse, eating disorders, and other personality disorders (Tomko et al., 2014). Despite its high symptom burden and co-morbidity rates, the 10-year course of BPD shows a high cumulative remission rate of 91% and a low cumulative relapse rate of 11% (Gunderson et al., 2011). However, compared to other clinical samples, severe and persistent impairments remain in social functioning, e.g. with only a third

* Corresponding author.
E-mail address: haang.jeung@med.uni-heidelberg.de (H. Jeung).

http://dx.doi.org/10.1016/j.neubiorev.2016.10.030
0149-7634/

of BPD patients working full-time. Finally, in a 27-year follow-up cohort, 10% of individuals with BPD were found to have committed suicide (Paris and Zweig-Frank, 2001).

1.1. Interpersonal behavior in Borderline Personality Disorder (BPD)

Difficulties in social interactions are of major significance in BPD (Gunderson, 2007). In the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, APA, 2013), two of the nine diagnostic criteria address maladaptive interpersonal behavior in BPD, namely "frantic efforts to avoid real or imagined abandonment" and "pervasive pattern of unstable and intense interpersonal relationships". In the new model of DSM-5, Section III, these two criteria have been included in BPD patients' incapacity for "intimacy" as part of interpersonal dysfunction. Additionally, the pathological personality trait "antagonism" has been proposed as a stable and consistent criterion for BPD, characterized by "hostility" that partly manifests as "anger or irritability in response to minor slights and insults".

In contrast to healthy subjects, individuals with BPD are trapped in a vicious circle of negative social interactions sustained by interpersonal hypersensitivity, affect dysregulation, and quarrelsome behavior (Sadikaj et al., 2013). The interpersonal difficulties of BPD patients have been related to impairments in sensing and responding to social signals. For instance, subjects with BPD tend to attribute negative emotions to neutral facial expressions (Daros et al., 2012; Domes et al., 2009). Moreover, in these individuals, seemingly minor daily events may trigger feelings of rejection, loneliness, and failure followed by frequent, intense, and persistent aversive tension (Stiglmayr et al., 2005).

When focusing on experimental studies that explicitly address interactive behavior and its modulating factors, the literature suggests an enhanced perception of and emotional distress in response to social exclusion in BPD (Lazarus et al., 2014; Lis and Bohus, 2013). In the following text, we would like to recapitulate the few experimental studies in which participants interacted with fictitious or human confederates. Among the existing paradigms, the Cyberball paradigm has been employed the most frequently to study the perception of social exclusion in BPD. Cyberball (Williams and Jarvis, 2006) is a virtual ball-tossing game in which two to three unknown confederates include or exclude the participant from the game. It is usually presented as an internet web page which depicts three to four animated ball-tossers standing in a circle, one of which represents the participant. When receiving the ball, the participant is asked to click on one of the others in order to throw the ball to him/her. In the inclusion condition, the participant receives as many ball passes as each of the other fictitious participants. In the exclusion condition, the participant receives either a smaller proportion or no participation at all.

BPD patients showed a bias towards the perception of exclusion independently of their factual participation (De Panfilis et al., 2015; Domsalla et al., 2013; Renneberg et al., 2012; Staebler et al., 2011). When excluded, BPD patients experienced a higher intensity of negative emotions than healthy controls (De Panfilis et al., 2015; Renneberg et al., 2012; Staebler et al., 2011). Hence, the paradigm has been used for mood induction, e.g. before reward-based decision making (Lawrence et al., 2010). However, it has not been explored whether the perception of social exclusion and related emotional distress have direct consequences on the interaction behavior of BPD patients.

When comparing BPD patients to healthy controls, one should take into account a different baseline in hostility and negative affect towards interaction partners, which might result in more reactive aggressive behavior. To evaluate aggressive behavior of BPD patients in response to another person's prior unfair behavior, the Point Subtraction Aggression Paradigm (PSAP) (Cherek et al., 1997) has been employed. The PSAP is a computer game in which participants can earn points by pressing a button 100 times. They can also subtract points from their interaction partner by pressing another button 10 times, but if they do this, these points will not be added to the their own earnings. The participants are told that their interaction partner might also take points from them. This narration blames the fictitious partner for the participants' losing points but prevents the participants from aggressive responding in order to earn money instead of earning money by pressing the other button. In two studies using the PSAP, BPD patients subtracted more points from the fictitious interaction partner than healthy controls (Dougherty et al., 1999; McCloskey et al., 2009). Moreover, point-subtracting responses significantly correlated with self-reported hostility (Dougherty et al., 1999) and self-reported trait aggression, especially physical aggression (McCloskey et al., 2009).

Even in the condition in which the participants did not lose any points due to their fictitious partners, female and male patients with BPD and intermittent explosive disorder (BPD-IED) subtracted more points from their opponents than did healthy controls (New et al., 2009). In general, BPD-IED patients pressed the subtraction button more often than healthy controls, and both groups subtracted more points from their opponents in response to pretended unfair behavior. However, in contrast to the authors' expectation, there was no group-by-condition interaction, and there were no correlations between point subtracting and clinical measures of anger or aggression in either the BPD-IED group or the HC group.

Besides the experimental settings with fictitious co-players, two studies employed real-life social interactions with human confederates to evaluate social feedback processing in BPD. In a modified "analogue" version of Cyberball, participants were asked to play cards with two attendant interaction partners (Ruocco et al., 2010). In contrast to the results from the virtual Cyberball paradigm, subjects with BPD did not differ from healthy controls in terms of the perception of inclusion and exclusion. However, the neural processing of social exclusion seemed to differ between BPD patients and healthy controls, with left medial prefrontal hyperactivation suggesting potential dysfunction of frontolimbic circuitry, as measured by functional near-infrared spectroscopy during the game. In a study in which participants played a well-known board game ("Monopoly"), BPD patients interacted with four healthy participants and were subsequently asked to rate themselves and one other participant on 80 character traits (Korn et al., 2016). The rating was conducted before and after receiving desirable and undesirable feedback from their interaction partners. Before the feedback, BPD patients rated themselves and others less favorably than did healthy controls. While healthy controls showed a positivity bias for self- and other-relevant feedback, BPD patients demonstrated a negativity bias for self-relevant feedback but not for other-relevant feedback. Especially after receiving negative feedback, BPD patients rated themselves more negatively than before the feedback. However, while both studies demonstrated alterations in either neural or behavioral feedback processing, the actual interactive behavior of BPD patients and their confederates was not described.

Given the complexity of social interaction between humans, studying the interactive behavior of BPD patients, with tight experimental control on the one hand and realistic, externally valid settings on the other hand, is challenging. One way of overcoming this challenge is to use well-validated game-theoretical experiments.

1.2. Game-theoretical background of social exchange

Starting from basic computations of social exchange (Axelrod and Hamilton, 1981), game theory has fostered our understanding of the evolutionary and individual origins of cooperation and