CrossMark

The measurement of reflective function in adolescents with and without borderline traits

Carolyn Ha [a,b,*], Carla Sharp [a,b], Karin Ensink [c], Peter Fonagy [d], Paul Cirino [a]

[a] Department of Psychology, 126 Heyne Building, University of Houston, Houston, TX 77204, USA
[b] The Menninger Clinic, 12301 Main Street, Houston TX 77035, USA
[c] Ecole de psychologie, Universite Laval, Pavillon Félix-Antoine-Savard, 2325, rue des Bibliothèques, Québec, Québec G1V 0A6, Canada
[d] University College London, 1-19 Torrington Place, London WC1E 7HB, UK

ABSTRACT

Reflective function refers to the capacity to reflect on the mind of self and others in the context of the attachment relationship. Reflective function (and its conceptual neighbor, mentalizing) has been shown to be an important correlate of a variety of disorders, including borderline personality disorder (BPD). The current study examined the construct validity of the Reflective Function Questionnaire for Youths (RFQY) in an inpatient sample of adolescents. Adequate internal consistency was established for the RFQY. Significant positive associations with an interview-based measure of reflective function and an experimental-based assessment of mentalization were found for the RFQY. Strong negative relations with BPD features were found and adolescent patients who scored above clinical cut-off for BPD symptoms demonstrated significantly poorer reflective function compared to patients without the disorder. These findings provide preliminary support for the notion that reflective function can be validly and reliably assessed in adolescent populations.
© 2013 The Foundation for Professionals in Services for Adolescents. Published by Elsevier Ltd. All rights reserved.

Introduction

Borderline personality disorder (BPD) is characterized by deficits in multiple areas of functioning including cognitive, affective, and behavioral domains. BPD is a disorder of significant impact due to high prevalence in clinical settings (Chanen et al., 2004; Grant et al., 2008; Grilo et al., 1998; Swartz, Blazer, George, & Winfield, 1990; Widiger & Weissman, 1991), poor clinical and psychosocial functioning (Chanen, Jovev, & Jackson, 2007), poor achievement of academic and occupational milestones (Winograd, Cohen, & Chen, 2008), and increased rates of hospitalization (Guilé & Greenfield, 2004). Given the poor prognosis and negative impact of BPD, it is essential to investigate the developmental pathways that lead to the disorder so that early interventions may be developed to prevent the relatively stable trajectory usually associated with borderline pathology. One such pathway is proposed by Fonagy and colleagues (Fonagy, Target, Gergely, Allen, & Bateman, 2003; Sharp & Fonagy, 2008) whose model focuses on the development of social-cognitive processes (specifically mentalizing) that are at the core of interpersonal interactions in an attachment context (Bateman & Fonagy, 2004; Fonagy & Bateman, 2008; Fonagy, Gergely, Jurist, & Target, 2002; Fonagy & Luyten, 2009). In this model, BPD is viewed as an attachment-related disorder where

* Corresponding author. Tel.: +1 281 615 0200.
E-mail addresses: cha@uh.edu, carolynha22@gmail.com (C. Ha), csharp2@central.uh.edu (C. Sharp), Karin.Ensink@psy.ulaval.ca (K. Ensink), Peter.Fonagy2@annafreud.org (P. Fonagy), Paul.Cirino@times.uh.edu (P. Cirino).

insecure attachment potentiates the development of social-cognitive (mentalizing) dysfunction associated with the disorder (Fonagy & Luyten, 2009; Sharp & Fonagy, 2008).

Within the mentalization-based model of BPD, mentalizing is defined as an individual's ability to understand or reflect on the context of, or the causes of, self and others' thoughts and feelings (Fonagy, Steele, Moran, Steele, & Higgitt, 1991). In other words, it is the person's capacity to think about and reflect on his/her own experiences and formulate interpretations about their own and others' behavior. Mentalizing is a broad concept that subsumes different social-cognitive functions, including emotion recognition, theory of mind, mindreading as well as reflective function. Reflective function, the focus of the current paper, is defined as the capacity to reflect on the mind of self and others in the context of the attachment relationship (Fonagy, Steele, Moran, Steele, & Higgitt, 1991). While the concept of reflective function is traditionally more associated with psychodynamic literature, mentalizing has more often been used in mainstream developmental and neuroscience literature. However, the two terms are sometimes used interchangeably (Fonagy, Target, Steele, & Steele, 1998).

Several measures have been developed to assess reflective function in adults (Luyten, Fonagy, Lowyck, & Vermote, 2012). The Adult Reflective Function Scale (ARFS) was developed by Fonagy et al. (1998), and is coded from transcripts generated from the Adult Attachment Interview (AAI, Main & Goldwyn, 1985–1995). Adult BPD patients with a history of abuse were shown to differentiate from patients without BPD with a history of abuse (Fonagy et al., 1996, 1998) using the ARFS. In addition, a 46-item self-report measure, the Reflective Function Questionnaire (RFQ), was developed by Fonagy and Ghinai (unpublished manuscript) to assess mentalizing capacity in adults. Preliminary reports demonstrate promising psychometric properties with good reliability and validity for this measure (Perkins, 2010, 2011).

For assessment of reflective function in youths, the Child Reflective Function Scale (CRFS) was developed (Target, Oandasan, & Ensink, 2001) and modeled from the ARFS (Fonagy et al., 1998). Scores for the CRFS are rated using transcriptions from a semi-structured interview, the Child Attachment Interview (CAI; Shmueli-Goetz, Target, Fonagy, & Datta, 2008; Target, Fonagy, Shmueli-Goetz, Datta, & Schneider, 1998). The CAI is an interview-based assessment designed to assess a child's attachment styles with his/her primary caregivers. The interview was modeled from the AAI (George, Kaplan, & Main, 1985; Main, 1995) for use in youth populations. The CAI contains 15 open-ended questions that elicit detailed information on relationship episodes, tapping into a child's perspective of him/herself and of his/her primary attachment figures, and reactions in response to upsetting events involving separation and loss. Studies have found the CAI to be a reliable and valid measure for assessing attachment in youths (Shmueli-Goetz et al., 2008; Target, Fonagy, & Shmueli-Goetz, 2003). Good inter-rater reliability for the reflective function items has been reported, with intraclass coefficients (ICCs) ranging from .6 to 1.00 (Ensink, 2004).

Despite its good psychometric properties, the CRFS is time-consuming and expensive (requiring extensive coding resources). Therefore, an adolescent version of the 46-item adult RFQ has been adapted for use with adolescents (Reflective Function Questionnaire for Youths [RFQY]; Sharp et al., 2009). Like its adult counterpart, the RFQY asks adolescents to rate how much they agree or disagree with a statement of reflective function on a 6-point Likert scale. However, the psychometric properties of the RFQY are unknown. It was therefore the aim of the current study to evaluate the construct validity of the RFQY by examining its correlations with the CRFS as well as measures more traditionally considered to tap into mentalizing capacity in the more general sense.

While examining the correlations between the RFQY and CRFS is a good test of the RFQY's construct validity, we furthermore wanted to explore correlations between the RFQY and measures tapping into the broader concept of mentalizing, as reflective function is considered one way in which the broader construct of mentalizing is operationalized, as explained earlier. In choosing measures of mentalizing for evaluating the construct validity of the RFQY, we considered conceptualizations of mentalization that describes mentalizing as having three dimensions: (1) implicit or explicit, (2) in relation to self or other, and (3) in cognitive or affective aspects (Choi-Kain & Gunderson, 2008; Fonagy & Luyten, 2009). A well-used task purported to measure explicit mentalizing is the "Child Eyes Test" (CET; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) which assesses emotion-recognition based on the eye region of the face (Baron-Cohen, Wheelwright, Scahill, Lawson, & Spong, 2001). Because the CET assesses for mentalization through asking participants to identify emotions through the eye region of faces, it is viewed as an explicit-controlled measure of mentalizing (Sharp et al., 2013).

A recently developed experimental measure of mentalization is the Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006). The MASC is a real-time, video-based assessment of ToM which measures accurate mentalizing and dysfunctions in mentalizing including no mentalizing, hypermentalizing, and undermentalizing (Dziobek et al., 2006). This measure has demonstrated dysfunction in mentalizing in several adult patient populations, such as Autism Spectrum Disorder (Dziobek et al., 2006), bipolar disorder (Montag et al., 2009), Narcissistic Personality Disorder (Ritter et al., 2011) and BPD (Preißler, Dziobek, Ritter, Heekeren, & Roepke, 2010). More recently, the measure has also demonstrated impaired mentalizing in adolescent patients with BPD (e.g. Sharp et al., 2011, 2013). The MASC is a broad assessment of mentalization which taps into implicit mentalizing.

We also included a measure of empathy in order to evaluate the construct validity of the RFQY. Empathy can be defined as the capacity to experience and/or understand another individual's emotions (Decety & Jackson, 2004) and is therefore considered related to mentalizing although not completely overlapping with mentalizing (Choi-Kain & Gunderson, 2008).

In summary, reflective function has been identified as an important construct for the development of borderline traits. While an interview-based measure (CRFS) has been developed for assessing reflective function in adolescents, it is time-consuming and expensive. A questionnaire-based measure (RFQY) has been developed, but no psychometric data has thus far been reported. Against this background, the aim of the current study was to evaluate the construct validity of the RFQY in

an inpatient sample of adolescents. We first assessed reliability, expecting that the RFQY would display adequate properties in this regard. Second, we examined the construct validity of the RFQY by evaluating its correlations with an interview-based measure of reflective function (CRFS), and its convergent and discriminant validity with measures of explicit and implicit mentalizing, as well as empathy. Specifically, we expected significant but weak relationships between the RFQY and empathy as well as the Child Eyes Test (explicit mentalizing), and a strong relation with the MASC (implicit mentalizing) and CRFS, which, like the RFQY, was developed as a measure of reflective function.

Next, we investigated construct validity of the RFQY by examining its relation to BPD, given established relations between RF and BPD as discussed earlier (e.g. Sharp et al., 2013), first by examining the relation between the RFQY and a questionnaire-based measure of BPD symptoms dimensionally and then, through utilization of a cut-off score, categorically. We made use of both parent- and self-reported symptoms of borderline personality disorder symptomatology in this regard as parents and adolescents provide unique perspectives on adolescent BPD symptoms (Chang, Sharp, & Ha, 2011), and multi-method assessments provide a more comprehensive picture of the behavioral phenotype. For the dimensional BPD scores, we expected negative correlations with RFQY scores, and for the categorical BPD scores, we expected lower RFQY scores in the group scoring above cut-off on the borderline measure.

Method

Participants

This study included a sample of 12–17-year-olds admitted to the adolescent unit of a private psychiatric hospital between October 2008 and May 2011. Consent and assent for study participation were obtained from both parents and adolescents. Parents completed assessments regarding their child's symptoms within the first two days of admission, in a private room, through a computer web-based survey system developed by the hospital. Trained research coordinators were available to assist parents with questions.

After consent and assent, some exclusion and inclusion criteria were applied. Inclusions for study participation consisted of: (1) any adolescent patient between 12 and 17 years of age, and (2) sufficient fluency in English to complete all research. Exclusions for study participation comprised the following: (1) diagnosis of schizophrenia or any psychotic disorder, and/or (2) diagnosis of mental retardation. The dataset included a total of 276 consecutively admitted adolescents. Twenty-three patients and their families declined participation in the study, and two families revoked consent. Additionally, 15 adolescents were excluded for other reasons including active psychosis ($n = 7$), ineligibility for consent due to language barrier, or being wards of court ($n = 8$). Finally, an additional 36 adolescents were excluded because of missing CAI videos as a result of equipment errors, adolescent refusal, or incomplete assessments resulting from abrupt discharges. An additional 54 videos were not coded for reflective function, so were not used in the final analyses. After these exclusions, a total of 146 participants were used in subsequent analyses.

To see if exclusions introduced bias, analyses on demographics were conducted to see if adolescent patients who did not participate in the study or have complete data were significantly different from those who did have complete data and were included in the study. Adolescent patients were grouped into completers if they had a completed CRFS score and met study inclusion criteria, and non-completers if they did not have complete CRFS data, declined study participation, or met any exclusion criteria. Analyses of completers and non-completers revealed no significant differences in demographics including age ($t = -1.89$, $df = 274$, $p = .06$) and IQ ($t = -.314$, $df = 173$, $p = .75$).

Measures

Self-report measure of reflective function

The Reflective Function Questionnaire for Youths (RFQY; Sharp et al., 2009) was adapted from the adult version which was developed in the United Kingdom (Fonagy & Ghanai, unpublished manuscript), by rewording several items for a more appropriate developmental level and modified for use with populations in the United States. For example, "People's thoughts are a mystery to me" was modified to "People's thoughts are a secret to me" and "My intuition about a person is hardly ever wrong" was replaced with "My feelings about a person are hardly ever wrong". The RFQY is a 46-item self-report measure. Responses are scored on a 6-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree", with two subscales computed after eight items are reverse-scored.

Scale A consists of items where optimal reflective function was scored at the mid-point of the scale, and extreme responses indicated poor reflective function. In accordance with the scoring instructions from the adult version, items on Scale A were recoded so that higher scores reflected optimal reflective function. Thus, "Strongly Disagree" or "Strongly Agree" (the extremes of the scale) were assigned a value of (2), responses of "Disagree" or "Agree" were assigned a value of (4), and responses of "Disagree Somewhat" or "Agree Somewhat" (the mid-point of the scale) were assigned a value of (6). For example, one item for Scale A states: "I always know what I feel", and for this item, strongly agreeing or disagreeing would indicate low reflective function. The 23 items that formed Scale A were then averaged to compute the overall subscale score. Adolescents with optimal reflective function would receive a maximum averaged score of 6 on this scale.

Scale B consisted of items where a higher score indicated high reflective function, and was also formed of 23 items which were averaged to form an overall subscale score. An example item from Scale B stated: "In an argument, I keep the other

person's point of view in mind", with responses of "Strongly Agree" scored as high reflective function and responses of "Strongly Disagree" scored as poor reflective function. The eight reverse-scored items were on Scale B, including items 7, 13, 14, 15, 23, 26, 32, and 38. An example of a reverse-scored item on this scale was: "I find it difficult to see other people's points of view", with responses of "Strongly Agree" coded as poor reflective function. Adolescents with optimal reflective function would also receive a maximum averaged score of 6 for Scale B.

A total RFQY score was then derived by summing the scores for scales A and B, with higher scores indicating a high capacity for reflective function. The maximum optimal reflective function score for the total scale would be 12. Since items on Scales A and B were not designed to be substantially different in terms of the content of reflective function (i.e., self and other), a total score was used. Furthermore, analyses conducted with the adult RFQ have supported the use of a combined total score (Perkins, 2009), therefore only the total RFQY was used in the final analyses.

Interview-based measure of reflective function

The Child Reflective Function Scale (CRFS; Target et al., 2001) was used to evaluate construct validity of the RFQY. The CRFS is coded based on transcriptions from the Child Attachment Interview (CAI; Target et al., 1998). In the current study, trained clinical research staff and doctoral-level graduate students in clinical psychology conducted CAIs with patients. Adolescent reflective function was then coded by a team of trained coders who were directly trained on the coding system by the developer of the CRFS (the third author).

Reflective function ratings were coded on an 11-point dimensional scale, ranging from -1 to 9, and anchored at six points in terms of ability to reflect on self and others in mental state terms. In other words, a score between 5 and 6 reflects an overall average level of reflective function, with scores of 7 or higher indicating high reflective function, and scores of 4 or lower indicating low to impaired reflective function. The self-understanding scale was computed from four items on the CAI that elicit self-descriptions and reactions in upsetting situations. On the CAI, adolescents were asked to provide three words to describe themselves, and then prompted to provide examples. For instance, an adolescent may have described him/herself using the word "intelligent". An example of an average reflective function response would be "My teacher says I'm intelligent because I made an A on the math exam", which would be coded with a score of 6. An example of a high reflective function score (9) would be a response of "I feel intelligent when my big brother cannot complete a math problem and I help him figure it out. That makes me feel intelligent". An example of a response that would be coded with a (-1), would be one that clearly is attacking the interviewer, such as "I did not say that I'm intelligent. Why are you asking so many questions? Is this interview over yet?" The other-understanding scale consisted of the sum of nine items tapping into the child's relationships with his/her attachment figure and a description of the attachment figure's reactions when they are angry or when they argue. A global reflective function score was assigned to the interview as a whole. In this study, the global CRFS scale was used as criterion measure.

Other measures of mentalization

To further assess construct validity two measures of mentalizing were also included. The Movie for Assessment of Social Cognition (MASC; Dziobek et al., 2006) was used to assess implicit mentalizing in typical social situations involving peer and romantic relationships. The storyline of the movie involved four characters getting together, preparing dinner, and then playing a board game, with themes focused on peer and romantic relationships. In total, adolescents were presented with 46 video clips via slides on the computer, and were then asked to imagine what the characters thought or felt as soon as each scene ended (e.g., "What is Betty thinking?"). Answer choices were presented in a multiple-choice format with four response options. Each response was coded as hypermentalizing (e.g., "Angry, her friend forgot she doesn't like sardines"), undermentalizing (e.g., "Surprised, she didn't expect sardines"), no mentalizing (e.g., "Sardines are salty and slippery"), or accurate mentalizing (e.g., "Repelled, she doesn't like sardines"). A total mentalizing score was derived from summing the total correct responses. Additionally, three separate scales were computed for the extent to which incorrect mentalization occurred, including hypermentalizing, undermentalizing, and no mentalizing, by summing total responses for each subscale.

Adequate psychometric properties have been established for the MASC (Dziobek et al., 2006) and it has been shown to be sensitive in discriminating patients with BPD from individuals without the disorder (Preißler et al., 2010; Sharp et al., 2011, 2013). In the current study, total mentalizing and all subscales, including no mentalizing, undermentalizing, and hypermentalizing were used in the analyses.

The Child's Eyes Test (CET) was developed by Baron-Cohen et al. (2001), and was used to assess for explicit mentalizing in this study. Adolescents were presented with 28 pictures of the eye region of the face and instructed to examine each photo carefully to determine which word best fit what the person in the photo seemed to be thinking or feeling. For each image, four words were provided for the adolescents to choose from; reflecting what feelings the person in the photo may be experiencing (e.g., jealous, scared, relaxed, hate). A total score was derived from a sum of the correct items. Adequate psychometrics have been reported for this measure (Baron-Cohen et al., 2001). In the current study, the continuous total score was examined in all analyses.

Empathy

The Basic Empathy Scale (BES) is a self-report measure developed to assess the multidimensional aspects of empathy (Jolliffe & Farrington, 2006). Adolescents were asked to rate 40-items on a 5-point Likert scale, ranging from 1 = Strongly Disagree to 5 = Strongly Agree. Good convergent and divergent validity have been demonstrated for the BES (Jolliffe &

Farrington, 2006). A total score was computed for this study, after eight items were reverse-scored, with higher scores indicating higher levels of empathy. Internal reliability for this measure was good ($\alpha = .83$) for the current study.

Borderline personality disorder symptoms (dimensional and categorical)

The Borderline Personality Features Scale for Children (BPFSC; Crick, Murray-Close, & Woods, 2005) is a 24-item self-reported questionnaire measure that assesses borderline personality features in children aged 9 years and older, including adolescents. Responses are scored on a 5-point Likert scale, ranging from 1 (Not at all true) to 5 (Always true) with total scores indicating greater levels of borderline personality features. The measure was used in the present study both dimensionally and categorically. For the latter, adolescents were grouped into BPD versus non-BPD groups based on a cut-off score of 66, which was derived in previous work on the BPFSC (Chang et al., 2011). A parent-report version of the BPFS was adapted from the self-report version, with adequate parent–child concordance demonstrated (BPFSP; Sharp, Mosko, Chang, & Ha, 2010). A cut-off score of 72 was established for parent-reported borderline symptoms (Chang et al., 2011). In the present sample, internal consistency for both the self- and parent- reports were good with Cronbach's alpha of .89 for the BPFSC and .91 for the BPFSP.

Adolescent clinical characteristics

The Diagnostic Interview Schedule for Children – Computerized version (NIMH DISC-IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000) was used to provide a description of the clinical characteristics of this sample. The DISC-IV is a highly structured clinical interview which assesses for Axis I disorders in children and adolescents aged 9–17 years. It is a well-established measure of Axis I psychopathology in youth and has good reliability and validity (Shaffer et al., 2000). In this study, interviews were administered individually and in private by trained research staff and ranged in length of about 1.5–2 h.

To provide a description of clinical characteristics for this sample (Table 1), we used DISC-IV diagnoses for the past year. Diagnoses are assigned a code with no diagnosis scored as 0, intermediate diagnosis scored 1, or positive diagnosis scored 2. These were recoded so no or intermediate diagnoses were assigned a score of 0 and positive diagnoses were assigned a score of 1. Axis I diagnoses were then separated into four categories: "Any Mood Disorder", included patients who met a positive diagnosis for either Major Depressive Disorder, Hypomania, Mania, or Dysthymia in the past year. "Any Eating Disorder" included patients who met a positive diagnosis for either Anorexia or Bulimia. For the "Any Anxiety Disorder" category, if the patient met criteria for any of the anxiety disorders (Generalized Anxiety Disorder, Agoraphobia, Obsessive–Compulsive Disorder, Panic Disorder, Posttraumatic Stress Disorder, Social Phobia, or Specific Phobia), they were grouped in this category. The "Any Externalizing Disorder" group included those with a diagnosis of either Attention Deficit Hyperactivity Disorder, Conduct Disorder, or Oppositional Defiant Disorder.

Procedures

Ethics approval for the current study was obtained from local review boards. This sample was recruited from a private tertiary care inpatient psychiatric hospital specializing in the assessment and stabilization of adolescents who have failed to

Table 1
Descriptive data and clinical characteristics of the full sample ($N = 146$).

Study variable	N	Mean (SD)	Minimum	Maximum
Age	146	15.57 (1.39)	12	17
IQ	113	106.88 (13.84)	71	149
Admit GAF	145	39.22 (7.03)	20	55
CRFS Global RF	146	3.15 (1.18)	1	8
RFQY total (self-report)	146	8.73 (.92)	4.48	10.35
Empathy (BES)	146	74.40 (10.13)	44	99
CET	139	20.34 (2.21)	14	25
MASC total	146	32.15 (5.05)	10	41
MASC hypermz	146	7.86 (4.04)	1	26
MASC undermz	146	3.30 (2.12)	0	10
MASC no mz	146	1.68 (1.61)	0	8

	N			%
History of medical problems	70			48
History of psychiatric problems	143			98
Any anxiety disorder	75			51
Any mood disorder	66			45
Any eating disorder	9			6
Any externalizing disorder	66			45
BPFSC	89			61
BPFSP	86			62

*Note: Axis I psychopathology was determined using DISC-IV diagnoses, and BPD was determined by the self- (BPFSC) and parent-reported (BPFSP) Borderline Personality Features Scale. Abbreviations: GAF = Global Adaptive Functioning, CRFS Global RF = Child Reflective Function Scale Global Reflective Function, RFQY = Reflective Function Questionnaire for Youths, BES = Basic Empathy Scale, CET = Child Eyes Test, MASC = Movie for Assessment of Social Cognition; hypermz = hypermentalizing, undermz = undermentalizing, no mz = no mentalizing.

respond to previous treatments. All admissions received a comprehensive psychiatric evaluation at intake. The clinic accepts patients with a range of psychiatric disorders. Procedures and details of the research-based assessment protocol are provided in detail elsewhere (Sharp et al., 2009).

Results

Descriptive findings

Descriptive data for main study variables are reported in Table 1 along with clinical characteristics of the full sample. The mean age for this sample was 15.75 ($SD = 1.39$) and the average IQ was 106.88 ($SD = 13.84$). Fifty percent of the sample had a previous psychiatric hospitalization, with 25% having two or more previous psychiatric hospitalizations. The mean reflective function score for this sample as determined by the RFQY was 8.73 ($SD = .92$) with a minimum of 4.48 and a maximum of 10.35. Sample means and ranges are reported in Table 1 for all other measures.

Next, normality assumptions were examined for the dependent variable. The Kolmogorov–Smirnov test was calculated for the RFQY total score, indicating a non-normal distribution ($KS = .091$, $df = 146$, $p = .005$). The distribution of scores on the RFQY were negatively skewed (skewness $= -1.28$) and were highly peaked or leptokurtic (kurtosis $= 3.12$). Therefore, non-parametric tests were used for all analyses.

Internal reliability

To assess the RFQY's internal consistency, Cronbach's alpha was computed. The internal consistency was $\alpha = .71$ for the total RFQY score, which was in an acceptable range.

Construct validity

Before examining the other assessments of the construct validity of the RFQY, we conducted a Principal Component Analysis (PCA) to elucidate the structure of the RFQY. The results of the PCA did not cleave items along the lines of Scale A and Scale B, probably because Scale A and Scale B were not designed on substantive basis (see Measures section above). There were two subscales that accounted for the bulk of the variance in the items, but these appeared to be best characterized as self and other reflective function. They were found to be moderately correlated ($r = .39$), and therefore the total score was used for remaining analyses.

Construct validity was examined first through correlations between the RFQY total score and the interview-based measure of reflective function (the CRFS). Spearman correlations revealed significant positive relations between RFQY total and CRFS global reflective function ($r_s = .24$, $p = .004$). Correlations are shown in Table 2.

Construct validity was further investigated by examining associations between the RFQY total score and experimental measures of mentalization (the CET and MASC), and as well as a measure of empathy (the BES). As expected, significant positive relations were found for RFQY total with total scores on the MASC ($r_s = .28$, $p = .001$). A significant negative correlation was found between RFQY total and the hypermentalizing subscale on the MASC ($r_s = -.32$, $p < .001$). However, no significant relationships were found for RFQY total score with the CET measure ($r_s = .04$, $p = .59$), or with empathy ($r_s = .09$, $p = .28$), MASC no-mentalizing ($r_s = -.02$, $p = .80$) and undermentalizing ($r_s = .03$, $p = .71$). Results from Spearman correlations are presented in Table 2.

Next, construct validity was first examined through correlations between RFQY total score and a dimensional approach to BPD using the BPFSC and BPFSP total scores. Spearman's correlations revealed a significant inverse association between reflective function and borderline features as reported by adolescents, BPFSC ($r_s = -.48$, $p < .001$) and parents, BPFSP ($r_s = -.22$, $p = .01$). Next, a Mann–Whitney test was conducted to examine whether group differences existed between

Table 2
Spearman correlations for RFQY total with an interview-based measure of reflective function (CRFS), other mentalizing measures, and empathy.

	1	2	3	4	5	6	7	8
1. RFQY Total (self-report)	–							
2. CRFS Global RF	.24**	–						
3. Empathy (BES)	.09	.17*	–					
4. CET	.04	.02	−.09	–				
5. MASC total	.28**	.35**	.05	.08	–			
6. MASC hypermz	−.32**	−.27**	−.09	−.05	−.79**	–		
7. MASC undermz	.03	−.11	.03	.03	−.45**	−.02	–	
8. MASC no mz	−.02	−.17*	.01	.001	−.36**	.00	.10	–

$*p < .05$, $**p < .001$.
Abbreviations: RFQY = Reflective Function Questionnaire for Youths, CRFS Global RF = Child Reflective Function Scale Global Reflective Function, BES = Basic Empathy Scale, CET = Child Eyes Test, MASC = Movie for Assessment of Social Cognition; hypermz = hypermentalizing, undermz = undermentalizing, no mz = no mentalizing.

adolescent patients above and below cut-off on both the BPFSP and BPFSC (both parent- and self-reported symptoms) on reflective function using the RFQY. A significant difference was found for reflective function between patients above and below cut-off for self-reported symptoms of BPD ($U = 1431.50$, $p < .001$, $r = -.34$) using the BPFSC. Adolescents who scored above cut-off on the borderline features scale had significantly poorer reflective function ($Mdn = 8.59$) compared to adolescents who scored below the cut-off ($Mdn = 9.35$). Reflective function also differed significantly between patients above and below cut-off on parent-reported symptoms of BPD features ($U = 1783.50$, $p = .05$, $r = -.16$). Adolescents who scored above cut-off on the parent-reported borderline symptoms had significantly lower scores on reflective function ($Mdn = 8.74$) compared with patients who scored below the cut-off on parent-reported borderline symptoms ($Mdn = 9.13$).

Discussion

According to the mentalization-based model of BPD (Fonagy & Luyten, 2009), reflective function has been identified as a potential important etiological factor in the development of BPD. While adult measures of reflective function have been developed (Fonagy & Ghinai, unpublished manuscript) and scrutinized for validity (Perkins, 2010, 2011), and while an interview-based measure of reflective function has been validated for use in adolescents (Ensink, 2004), nothing is known about the psychometric properties of a questionnaire-based measure of reflective function for use in adolescents. Previously, we adapted the 46-item adult RFQ for use in adolescents (Sharp et al., 2009). The current study is the first to examine whether reflective function can be validly assessed through this questionnaire-based measure, which is less time- and labor-intensive than other measures of reflective functioning. Adequate internal reliability was found for all 46 items on the RFQY. This supports the use of the RFQY as a reliable measure to assess adolescent reflective function capacity. However, internal reliability was not completely optimal, and point to the need for further work on the internal construct validity of the RFQY.

Criterion validity was investigated by examining the relation between the RFQY total score and CRFS global score. Significant relations were found for RFQY total score with the CRFS global score. Effect sizes, were albeit moderate and we discuss this in more detail below when reviewing the limitations of the current study. Next, construct validity was examined between the RFQY total and two experimental measures of mentalizing and a measure of empathy. Our findings support the convergent validity of the RFQY, which related significantly to the MASC total score as predicted. In addition, we found a significant inverse relation between RFQY and the MASC hypermentalizing scale. In other words, a high score on hypermentalizing, or an over interpretation of mental states, was related to low reflective function capacity measured by the RFQY. No significant relations were found between the RFQY and the MASC no-mentalizing or undermentalizing subscales. The RFQY and the MASC hypermentalizing subscale therefore tap into similar aspects of mentalizing, in contrast to the MASC no-mentalizing and undermentalizing subscales, which relate to mental states in the experimental stimuli rarely being reported by test subjects. For example, in one scene in the MASC, Sandra offers Cliff a drink, but when she gets to the kitchen, she finds out that the dessert was ruined. The scene stops and adolescents are asked "What is Sandra feeling?" The correct response option was: "She is frustrated about the burnt cake", while a no-mentalizing response was: "She forgot to bring the coke", and an undermentalizing response was: "She is sure that they will have no dessert". In contrast, the hypermentalizing response was "She is afraid that the others will laugh at her".

Furthermore, no significant relation was found between the RFQY total and CET or with a measure of empathy. One explanation may be that the CET and RFQY tap into different aspects of mentalization. The CET taps into external and others' mental states by asking adolescents to rate eye regions of the face, as opposed to the RFQY, which assesses internal aspects of both self and other mental states. The RFQY also examines broader aspects of mentalization including more complex interpretative forms of mentalizing, which overlaps with mentalization components assessed for in tasks like the MASC, but not with the CET, which assesses more narrow aspects of mentalization related to emotion understanding or recognition. While empathy is a construct closely related to mentalizing, it can be distinguished from mentalizing, as our findings support no relations between the RFQY and empathy. Although empathy is not a form of mentalizing, it relates to mentalizing in that the ability to empathize relies on an individual's capacity to respond emotionally to another's mental state, which therefore involves mentalization (Sharp, 2006). Taken together, these findings support the discriminant validity of the RFQY with other measures of mentalizing (the CET) and with empathy (BES).

Construct validity was further examined by investigating the relation between the RFQY and BPD features. Using a dimensional approach to measuring borderline features, our results showed high levels of borderline features were associated with low reflective function. Additionally, when using a categorical approach to both self-reported and parent-reported BPD features, a significant group difference emerged for patients above cut-off on BPD symptoms compared to patients below cut-off on BPD symptoms, with the group above cut-off demonstrating poorer (lower) reflective function. This is consistent with adult research which has demonstrated poorer reflective function in individuals diagnosed with BPD (Fonagy et al., 1996, 1998; Perkins, 2009).

Taken together, our findings demonstrate that the RFQY is a promising self-report measure of mentalization in adolescent inpatients. The findings from this study must be interpreted with caution as there are several limitations. First, we did not conduct a thorough examination of the internal factor structure of the RFQY because our focus was on demonstrating convergence of the RFQY with other measures of mentalizing, as well as BPD. However, results of a PCA suggested much work is needed to further refine and develop the RFQY in terms of its internal factor structure. That the Cronbach's alpha was only in the acceptable range further points to necessary future work to refine the RFQY to be a more streamlined and effective measure of RF in adolescents. Ideally, this type of research should be conducted in a large unselected sample where

sophisticated factor analytic techniques can be employed. Second, our sample was composed primarily of predominantly Caucasian adolescents (91%) from well-educated and financially stable environments. These findings may not generalize to other adolescent populations including community and outpatient samples from diverse backgrounds.

Furthermore limiting generalizability is that this group of adolescent inpatients represents a severe end of the spectrum, as they have failed to respond to prior treatments. Another limitation to the findings is that the criterion validity was low although the relation between RFQY and the criterion measure (CRFS) was significant. This may be due to the typical relations between cross-method assessments utilizing a questionnaire and interview based measure which report correlations around .30 (Achenbach, McConaughy, & Howell, 1987), while questionnaire based measures correlate more highly with other questionnaire based measures at .70 (Achenbach et al., 1987). Finally, this study lacked the ability to demonstrate the clinical utility of the RFQY in identifying patients with poor reflective function due to the nature of the sample. Future studies should incorporate a community sample to establish a cut-off on the RFQY against CRFS for reflective function.

Despite these limitations, this study is the first to provide evidence in support of the reflective function construct in adolescents, as measured by the RFQY. The findings support the reliability and construct validity of the RFQY as a newly adapted measure of social cognition (mentalization) for adolescents, and shows promise as a useful tool for clinicians to assess mentalization in inpatient adolescents, especially when time and financial constraints limit the use of additional measures. Assessing reflective function is important – for many disorders, but in particular for BPD. Additionally, although it is currently unknown whether the RFQY is sufficiently sensitive to assess change in reflective function, it may be useful to track treatment outcome in treatment settings that use social-cognitive interventions. It is important for clinicians to have brief but adequate measures to effectively assess mentalizing in adolescent patients so that specific interventions may be developed to target problematic mentalizing in various treatment settings. The current findings also help to further reinforce the link between impaired mentalization and BPD in adolescents, which will aid in improving interventions for patients suffering from this challenging disorder.

Acknowledgments

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213–232. http://dx.doi.org/10.1037/0033-2909.101.2.213.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(2), 241–251.

Baron-Cohen, S., Wheelwright, S., Scahill, V., Lawson, J., & Spong, A. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger's syndrome. *Journal of Developmental and Learning Disorders, 5, 47–78.*

Bateman, A. W., & Fonagy, P. (2004). *Psychotherapy for borderline personality disorder: Mentalization-based treatment.* Oxford, UK: Oxford University Press.

Chanen, A. M., Jackson, H. J., McGorry, P. D., Allot, K. A., Clarkson, V., & Yuen, H. P. (2004). Two-year stability of personality disorder in older adolescent outpatients. *Journal of Personality Disorders, 18*, 526–541. http://dx.doi.org/10.1521/pedi.18.6.526.54798.

Chanen, A. M., Jovev, M., & Jackson, H. J. (2007). Adaptive functioning and psychiatric symptoms in adolescents with borderline personality disorder. *Journal of Clinical Psychiatry, 68*, 297–306. http://dx.doi.org/10.4088/JCP.v68n0217.

Chang, B., Sharp, C., & Ha, C. (2011). The criterion validity of the Borderline Personality Feature Scale for Children in an adolescent inpatient setting. *Journal of Personality Disorders, 25*, 492–503. http://dx.doi.org/10.1521/pedi.2011.25.4.492.

Choi-Kain, L. W., & Gunderson, J. G. (2008). Mentalization: ontogeny, assessment, and application in the treatment of borderline personality disorder. *American Journal of Psychiatry, 165*, 1127–1135. http://dx.doi.org/10.1176/appi.ajp.2008.07081360.

Crick, N. R., Murray-Close, D., & Woods, K. (2005). Borderline personality features in childhood: a short-term longitudinal study. *Development and Psychopathology, 17*, 1051–1070. http://dx.doi.org/10.1017/S0954579405050492.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*, 71–100. http://dx.doi.org/10.1177/1534582304267187.

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., et al. (2006). Introducing MASC: a Movie for the Assessment of Social Cognition. *Journal of Autism and Developmental Disorders, 36*, 623–636. http://dx.doi.org/10.1007/s10803-006-0107-0.

Ensink, K. (2004). *Assessing theory of mind, affective understanding and reflective functioning in primary school-aged children.* Unpublished PhD dissertation. UK: University of London.

Fonagy, P., & Bateman, A. (2008). The development of borderline personality disorder – a mentalizing model. *Journal of Personality Disorders, 22*, 4–21. http://dx.doi.org/10.1521/pedi.2008.22.1.4.

Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self.* New York, NY: Other Press.

Fonagy, P., & Ghinai, R. *A self-report measure of mentalizing: Development and preliminary test of the reliability and validity of the Reflective Function Questionnaire (RFQ).* Unpublished manuscript. UK: University College London.

Fonagy, P., Leigh, T., Steele, M., Steele, H., Kennedy, R., Mattoon, G., et al. (1996). The relation of attachment status, psychiatric classification, and response to psychotherapy. *Journal of Consulting and Clinical Psychology, 64*, 22–31. http://dx.doi.org/10.1037/0022-006X.64.1.22.

Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and Psychopathology, 21*, 1355–1381. http://dx.doi.org/10.1017/S0954579409990198.

Fonagy, P., Steele, H., Moran, G., Steele, M., & Higgitt, A. (1991). The capacity for understanding mental states: the reflective self in parent and child and its significance for security of attachment. *Infant Mental Health Journal, 13*, 200–217. http://dx.doi.org/10.1002/1097-0355(199123)12:3<201::AID-IMHJ2280120307>3.0.CO;2-7.

Fonagy, P., Target, M., Gergely, G., Allen, J. G., & Bateman, A. (2003). The developmental roots of borderline personality disorder in early attachment relationships: a theory and some evidence. *Psychoanalytic Inquiry, 23*, 412–459. http://dx.doi.org/10.1080/07351692309349042.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-functioning manual, version 5.0, for application to adult attachment interviews*. London, UK: University College London.

George, C., Kaplan, N., & Main, M. (1985). *Adult attachment interview*. Unpublished manuscript. Berkeley: University of California.

Grant, B. F., Chou, S. P., Goldstein, R. B., Huang, B., Stinson, F. S., Saha, T. D., et al. (2008). Prevalence, correlates, disability, and comorbidity of DSM-IV borderline personality disorder: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry, 69*, 533–545. http://dx.doi.org/10.4088/JCP.v69n0404.

Grilo, C. M., McGlashan, T. H., Quinlan, D. M., Walker, M. L., Greenfeld, D., & Edell, W. S. (1998). Frequency of personality disorders in two age cohorts of psychiatric inpatients. *American Journal of Psychiatry, 155*, 140–142.

Guilé, J. M., & Greenfield, B. (2004). Introduction. Personality disorders in childhood and adolescence. *Canadian Child and Adolescent Psychiatry Review, 13*, 51–52.

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence, 29*, 589–611. http://dx.doi.org/10.1016/j.adolescence.2005.08.010.

Luyten, P., Fonagy, P., Lowyck, B., & Vermote, R. (2012). Assessment of mentalization. In A. W. Bateman, & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (pp. 43–66). Washington, DC: American Psychiatric Press.

Main, M. (1995). Recent studies in attachment: overview, with selected implications for clinical work. In S. Goldberg, R. Muir, & J. Keer (Eds.), *Attachment theory: Social, developmental, and clinical perspectives*. Hillsdale, NJ: Analytic Press.

Main, M., & Goldwyn, R. (1985–1995). *The Adult Attachment Interview classification and scoring system*. Unpublished Manuscript. University of California at Berkeley.

Montag, C., Ehrlich, A., Neuhaus, K., Dziobek, I., Heekeren, H. R., Heinz, A., et al. (2009). Theory of mind impairments in euthymic bipolar patients. *Journal of Affective Disorders, 123*, 264–269. http://dx.doi.org/10.1016/j.jad.2009.08.017.

Perkins, A. (2009). *Feelings, faces and food: Mentalization in borderline personality disorder and eating disorders*. Dissertation. Guildford, UK: Department of Psychology, University of Surrey.

Perkins, A. (2010, March). *A new self-report measure of mentalization: The Reflective Function Questionnaire*. Presented at the UK Conference of the Society for Psychotherapy Research, Ravenscar, UK http://www.psychotherapyresearch.org/displaycommon.cfm?an=1&subarticlenbr=266.

Perkins, A. (2011, May). *Development and validation of a new self-report measure of mentalization: The 54-item Reflective Function Questionnaire*. Presented at the SPR/BACP Conference, Liverpool, UK.

Preißler, S., Dziobek, I., Ritter, K., Heekeren, H. R., & Roepke, S. (2010). Social cognition in borderline personality disorder: evidence for disturbed recognition of the emotions, thoughts, and intentions of others. *Frontiers in Behavioral Neuroscience, 4*, 1–8.

Ritter, K., Dziobek, I., Preißler, S., Rütter, A., Vater, A., Fydrich, T., et al. (2011). Lack of empathy in patients with narcissistic personality disorder. *Psychiatry Research, 187*, 241–247. http://dx.doi.org/10.1016/j.psychres.2010.09.013.

Shaffer, D., Fisher, P., Lucas, C., Dulcan, M., & Schwab-Stone, M. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 28–38.

Sharp, C. (2006). Mentalizing problems in childhood disorders. In J. G. Allen, & P. Fonagy (Eds.), *Handbook of mentalization-based treatments* (pp. 201–212). Chichester, UK: John Wiley & Sons.

Sharp, C., & Fonagy, P. (2008). Social cognition and attachment-related disorders. In C. Sharp, P. Fonagy, & I. M. Goodyer (Eds.), *Social cognition and developmental psychopathology* (pp. 269–302). Oxford, UK: Oxford University Press.

Sharp, C., Ha, C., Carbone, C., Kim, S., Perry, K., Williams, L., et al. (2013). Hypermentalizing in adolescent inpatients: treatment effects and association with borderline traits. *Journal of Personality Disorders, 27*(1), 3–18.

Sharp, C., Mosko, O., Chang, B., & Ha, C. (2010). The cross-informant concordance and construct validity of the Borderline Personality Features Scale for Children in a sample of male youth. *Clinical Child Psychology and Psychiatry, 16*, 1–15.

Sharp, C., Pane, H., Ha, C., Venta, A., Patel, B., Sturek, J., et al. (2011). Theory of mind and emotion regulation difficulties in adolescents with borderline traits. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*, 563–573. http://dx.doi.org/10.1016/j.jaac.2011.01.017.

Sharp, C., Williams, L., Ha, C., Baumgardner, J., Michonski, J., Seals, R., et al. (2009). The development of a mentalization-based outcomes and research protocol for an adolescent in-patient unit. *Bulletin of the Menninger Clinic, 73*, 311–338.

Shmueli-Goetz, Y., Target, M., Fonagy, P., & Datta, A. (2008). The Child Attachment Interview: a psychometric study of reliability and discriminant validity. *Developmental Psychology, 44*, 939–956. http://dx.doi.org/10.1037/0012-1649.44.4.939.

Swartz, M. S., Blazer, D., George, L., & Winfield, I. (1990). Estimating the prevalence of borderline personality disorder in the community. *Journal of Personality Disorders, 4*, 257–272. http://dx.doi.org/10.1521/pedi.1990.4.3.257.

Target, M., Fonagy, P., & Shmueli-Goetz, Y. (2003). Attachment representations in school-age children: the development of the child attachment interview (CAI). *Journal of Child Psychotherapy, 29*, 171–186. http://dx.doi.org/10.1080/0075417031000138433.

Target, M., Fonagy, P., Shmueli-Goetz, Y., Datta, A., & Schneider, T. (1998). *The Child Attachment Interview (CAI) protocol, Revised Edition VI*. Unpublished manuscript. UK: University College London.

Target, M., Oandasan, C., & Ensink, K. (2001). *Child Reflective Functioning Scale scoring manual: For application to the child attachment Interview*. Unpublished manuscript. UK: Anna Freud Centre/University College London.

Widiger, T. A., & Weissman, M. M. (1991). Epidemiology of borderline personality disorder. *Hospital and Community Psychiatry, 42*, 1015–1021.

Winograd, G., Cohen, P., & Chen, H. (2008). Adolescent borderline symptoms in the community: prognosis for functioning over 20 years. *Journal of Child Psychology and Psychiatry, 49*, 933–941. http://dx.doi.org/10.1111/j.1469-7610.2008.01930.x.