Reliability and validity of the Mentalization Questionnaire (MZQ) in forensic care

Pleuni Paridaens

Department of Development Psychology, Tilburg University

Author note

Pleuni Paridaens, Master Psychology and Mental Health- Forensic Psychology, Tilburg University

Under the supervision of Erik Bulten and Jelle J. Sijtsema.

Abstract

This study further validates the Mentalization Questionnaire (MZQ), an easy, fast, and cost-efficient self-report questionnaire originally developed by Hausberg et al. (2012). Participants were retrieved from the general population (N = 173, 42,8% men). An exploratory factor analyses, Pearson correlations, and four multiple regression analyses were conducted for testing the validity and reliability of the MZQ. Exploratory factor analyses showed a one-factor solution. Pearson correlations showed relations with mindfulness and alexithymia. Contrary to expectations, this study showed a relation with the cognitive side of empathy, however not with the affective side of empathy. Multiple regression analyses showed that mentalization was a unique predictor of antisocial behavior in addition to measures of empathy, mindfulness, and alexithymia. It can be concluded that the MZQ is an appropriate and distinct measure of at least the cognitive component of mentalization.

Introduction

Mentalization is defined as the ability with which humans can read the mental states of other humans in order to predict the behavior of others (Blakemore & Robins, 2012; Frith, 2006). These mental states include thinking about thoughts, emotions, desires, feelings, beliefs, wishes, and needs in oneself and others (Hausberg et al., 2012). The brain is able to represent these mental states of the self and the other and the relationship between these mental states, making it therefore possible for humans to communicate (Frith, 2006).

Mentalizing is a construct that is used within several therapies, such as cognitive behavioral therapy (Björgvinsson & Hart, 2006), psychodynamic therapy (Bateman & Tyrer, 2004), transference-focused therapy (Kernberg et al., 2008) and psychoanalytic therapy (Taubner, Kessler, Buchheim, Kächele & Staun, 2011). Treatments that focus on learning to mentalize are called mentalization-based treatments (MBT; Bateman & Fonagy, 1999) and were initially developed for patients with borderline personality disorder (Bateman & Fonagy, 1999). Several studies showed its effectiveness on patients with borderline personality disorder (e.g., Bateman & Fonagy, 2008; Bateman & Fonagy, 2010; Choi-Kain & Gunderson, 2008; Fonagy, Luyten & Strathearn, 2011; Rossouw & Fonagy, 2012). Bateman and Fonagy (2008) stated that mentalization-based treatment has the best treatment outcome (e.g., lower suicidality rates, less need of service use, less use of medication, higher global functioning and higher vocational status) for patients with borderline personality disorder compared to other treatments. The effectiveness of mentalization-based treatment has also been demonstrated in patients diagnosed with antisocial personality disorder (McGauley, Yakeley, Williams & Bateman, 2011), psychotic disorders (Brent, 2009), substance abuse disorders (Söderström & Skårderud, 2009), and for reducing self-harm in adolescents (Rossouw & Fonagy, 2012). It is also used for family conflicts and for reducing aggression between adolescents at school (Asen & Fonagy, 2012; Fonagy & Luyten, 2009).

Despite the use of mentalizing in all sorts of treatments and the positive treatment outcomes of mentalization programs or modules, mentalization itself is a difficult concept to assess. Current assessment of mentalization is time-consuming, cost-expensive and unidimensional (Choi-Kain & Gunderson, 2008; Kemps & Kooiman, 2015; Luyten, Fonagy, Lowyck & Vermote, 2012). For example, the first instrument developed for measuring mentalization was the Reflective Functioning Scale (RFS; Fonagy, Target, Steele & Steele, 1997), which can be used in addition to semi-structured interviews such as the Adult Attachment Interview (AAI; George, Kaplan & Main, 1985). However, research that supports the validity of the RFS has not been replicated since 1998 and the test-retest reliability is not yet established (Cho-Kain & Gunderson, 2008; Luyten et al., 2012). Hence, the Reflective Functioning Rating Scale (RFRS; Meehan, Levy, Reynoso, Hill & Clarkin, 2009) was developed which was less time-consuming and more multidimensional. Yet, according to Kemps and Kooiman (2015), the RFRS showed insufficient psychometric qualities, in particular with respect to the third scale (i.e., non-mentalized behavior). Another instrument, the 'Reflective Self Functioning Scale' (RSFS; Fonagy et al., 1997), had good psychometric properties but showed inconsistent results and was also relatively time-consuming (Hausberg et al., 2012).

It can be concluded that there is an urgent need for a new cost-efficient, less time-consuming and multidimensional questionnaire for the assessment of mentalization ability. In response to this need, the self-report Mentalization Questionnaire (MZQ; Hausberg et al., 2012) was developed. Self-report questionnaires such as the MZQ are an advantage as opposed to the instruments mentioned above, because they are easy and fast to complete, involve minimal clinician time and generally less administration time (Ouwersloot, Brink, Diekstra & Hoogduin, 1994; Foa, Cashman, Jaycox, Perry & Kevin, 1997). Hausberg et al. (2012) were the first researchers who studied the psychometric properties of the MZQ, and

they found acceptable reliability and sufficient validity outcomes in a mental disorder patient population. However, they also mentioned some limitations of their study; the factor-solution found in their study was not recommended to use for further validation and the discreteness of the MZQ for measuring mentalizing ability was to be further investigated. This study aims to address these limitations by further validating the MZQ by means of three purposes, presented below.

The first purpose of this study concerns the first limitation found in the study of Hausberg et al. (2012). They found a four-factor solution of the MZQ: 'Refusing self-reflection', 'Emotional awareness', 'Psychic equivalence mode' and 'Regulation of affect', these four factors are more fully described in the method section. Although these four subscales had satisfactory internal consistencies, they did not recommend to use this four-factor structure for further validation. Hausberg et al. (2012) found different reliability coefficients between the samples that might be explained due to significant differences in sample characteristics (e.g. differences in distribution of diagnosis, suicide attempts, and self-harming behavior). Therefore, they advised to re-examine the factors in a larger heterogeneous sample. Hence, the first purpose of this study, exploring the factor solution of the MZQ in the general population, and to see if this matches the factor-solution found in the study of Hausberg et al. (2012).

The second purpose of this study concerns the second limitation in the study of Hausberg et al. (2012). They stated that examining the divergent validity is an important next step, because it is not yet clear if the MZQ is a discrete measure of the ability to mentalize. Mentalization was found to partially overlap with several concepts such as theory of mind, the ability to reflect, empathy, alexithymia, mindfulness, and affect regulation (e.g. Gallagher & Firth, 2003; Cho-Kain & Gunderson, 2008; Allen & Fonagy, 2006; Swart, Kortekaas & Aleman, 2009; Fonagy, Gergely, Jurist & Target, 2002). Mechanisms underlying the theory

of mind are the same mechanisms that underlie mentalizing ability (Gallagher & Firth, 2003). However, the concept mentalization is more related to affect and understanding interpersonal relationships between oneself and others. Another concept that is found to be related to the ability to mentalize is empathy (Hooker, Verosky, Germine, Knight, & D'Esposito, 2008; Allen, & Fonagy, 2006; Schnell, Bluschke, Konradt, & Walter, 2010; Singer et al., 2004). People with high empathy can transpose themselves in another person/situation, experience feelings of sympathy and concern for others, and experience self-oriented feelings of fear and discomfort at witnessing negative experiences of others (Davis, 1980). For example, Hooker et al. (2008) found that predicting an emotional response requires using internal affective representations, the same representations that are used in mentalizing. Greater use of these internal affective representations in trying to understand emotional experience of others has been associated to perceiving more empathy (Hooker et al., 2008). The study of Langdon and Coltheart (2001) found that especially perspective taking (a part of empathy) is related to mentalization ability. Another study found that the cortical network that involves mentalization is also activated when cognitive empathy is experienced (Schnell et al., 2010). Singer et al. (2004) found that when we want to understand someone else's emotional reaction to pain, we use the same representations for our ability to mentalize and our ability to empathize. However, some studies found that empathy and mentalizing are different concepts (e.g., Achim, Ouellet, Roy & Jackson, 2010; Cho-Kain & Gunderson, 2008). Achim et al. (2010) did not find a strong association between empathy and mentalizing. Moreover, Cho-Kain and Gunderson (2008) stated that although both mentalization and empathy involve imagining mental states in others, empathy is more other-oriented, whereas mentalization is both self and other-oriented. Mindfulness is also a construct that is similar to mentalization (Cho-Kain & Gunderson, 2008). Mindfulness means having a receptive attention to and awareness of our thoughts, feelings, bodily sensations, and surrounding environment (Brown

& Ryan, 2003). Mentalization and mindfulness both include seeing mental states as temporary and subjective and they both enhance emotion regulation (Wallin, 2007). Cho-Kain & Gunderson (2008) stated that mindfulness seems to be a subdivision of the mentalization framework. The difference between these two, according to Cho-kain and Gunderson (2008) is that "mindfulness aims at acceptance of internal experience, whereas mentalization emphasizes the construction of representation and meaning related to these experiences" (pp. 5). A concept that is negatively associated to the ability to mentalize is alexithymia (e.g., Swart et al., 2009; Moriguchi et al., 2006). People with alexithymia have trouble identifying and describing their emotions and tend to minimize emotional experience and focus their attention externally (Moriguchi et al., 2006; Bagby, Parker, & Taylor; 1994a). Alexithymia is associated with an impairment for higher order mentalizing which in turn is associated with an inability to take perspective of others (Swart et al., 2009; Moriguchi et al., 2006). Moriguchi et al. (2006) also stated that the skills used in both mentalizing and alexithymia are inter-related. In sum, there are different concepts that are similar to mentalization, however also differences exists between these concepts and mentalization. Therefore, a second aim of this study is testing the convergent and divergent validity of the MZQ, because it can establish if the MZQ is an appropriate, distinct questionnaire to assess mentalization or that it is measuring the same as one of the concepts mentioned above.

The last purpose of this study is also in line with the second limitation mentioned in the study of Hausberg et al. (2012); investigate if the MZQ has a unique value in the explanation of relevant behavior such as antisocial behavior relative to the already existing questionnaires, e.g., the Davis Empathy Scale (DES; Davis 1980), a measurement tool for empathy. For example, it is known that a lack of empathy and a lack of mentalization ability both are related to displaying more antisocial behavior (e.g., Marshall & Marshall, 2011; Miller & Eisenberg, 1988; McGauley et al., 2011; Leichsenring, Kunst & Hoyer 2003; Taylor

& Signal, 2005). It is also known that mindfulness decreases the probability of engaging in antisocial behaviors (Singh et al., 2007). Moreover, having alexithymia is also related to displaying more antisocial behavior (Manninen et al., 2011). However, mentalization, empathy, mindfulness and alexithymia are seemingly different concepts (e.g. Achim et al., 2010; Cho-Kain & Gunderson, 2008; Wallin, 2007; Moriguchi et al., 2006). Therefore, it can be expected that the MZQ has a different value in the explanation of antisocial behavior as opposed to the other concepts. If this is not the case, then the MZQ is not a distinct measure of mentalization, as it captures one of the other concepts.

In conclusion, there are three aims for this study to investigate. First, explore the factor solution of the MZQ in a general population, and to see if this matches the four-factor solution found by Hausberg et al. (2012). The hypothesis is that the factor solution will differ from the four-factor solution found in Hausberg et al. (2012), because the sample characteristics differ significantly (e.g., mental disorder patient population versus general population). Second is to test the convergent and divergent validity of the MZQ in relation to (partly) similar concepts. Out of previous research (e.g., Singer et al., 2004; Cho-Kain & Gunderson, 2008; Swart et al., 2009), it is to be expected that higher scores on the MZQ, which is measuring mentalization ability, is related to higher scores on the questionnaires measuring empathy and mindfulness, and to lower scores on the questionnaire measuring alexithymia. However, because of the multidimensional character of the MZQ it is also expected that the MZQ is distinctive in measuring mentalization relative to the other questionnaires measuring similar constructs. Third is to test whether mentalization uniquely explains antisocial behavior in addition to the measures of empathy, mindfulness and alexithymia. It is to be expected that mentalization has a unique role in the explanation of antisocial behavior, because empathy and mindfulness for example appears to be a component of the mentalizing framework (Cho-Kain & Gunderson, 2008).

Method

Participants and procedure

The participants that at least completed the MZQ (*N* = 173) included 74 men (42,8%), 87 woman (50,3%) and 12 missing (6,9%). Average age was 40,6 years (SD = 14,5). The youngest participant was 19 years and the oldest was 78 years. The majority of the participants was highly educated (Higher professional education [HBO] & University education [WO] = 64,1%), was operating in the Health Care business (32,0%), had an income above 2000 a month (60,1%) and had a Dutch ethnicity (98,3%).

Data was collected as part of the master thesis forensic psychology at Tilburg University. Data were collected by two Master students of psychology in the period of February 2016 till April 2016. The participants were recruited online using the Facebook and email of the two Master students. Participants were able to choose when and where they wanted to fill in the questionnaires. Before the start of the questionnaires, there was an introduction that mentioned the goal of the study and that all information would be kept confidential and anonymous. At the bottom of the introduction they were able to proceed to the next page and thereby gave permission for participating. At the end of the questionnaires, there was an opportunity to fill in an email, making it therefore possible to give back the results of this study to the participant. Participation was entirely voluntary and did not yield any advantages for the participants.

The questionnaire comprises a composition of items about demographics (i.e., gender, age, education level, income, ethnic background, and profession) and existing tests in the field of personality characteristics (Dirty Dozen [DD]; Davis Empathy Scale [DES]; Mindful Attention Awareness Scale [MAAS]; Multidimensional Personality Questionnaire [MPQ]; Mentalization Questionnaire [MZQ]; Toronto Alexithymia Scale – 20 [TAS-20]), control strategies (Resource Control Strategy Inventory [RCSI]), and aggression (Subtypes of

Antisocial Behavior [STAB]). For this study the questionnaires MZQ, DES, MAAS, TAS-20 and STAB were used.

Mentalization Questionnaire (MZQ). The MZQ (Hausberg et al., 2012) is a 15-item self-reported questionnaire that measures the construct mentalization. The underlying theory of this questionnaire is originative from current literature on psychopathology and mentalization (Bateman & Fonagy, 2004; Bergmann-Mausfeld, 2006; Fonagy et al., 2002; Stein, 2003). Some items of the MZQ were derived from the German reflective functioning manual (Daudert, 2002). All items are controlled for formulation and plausibility by an expert in psychological diagnostics and experts in the field of MBT. As mentioned in the introduction section, the MZQ exists of four subscales: 'Refusing self-reflection' (e.g., "Most of the time it is better not to feel anything."), 'Emotional awareness' (e.g., "Sometimes I only become aware of my feelings in retrospect."), 'Psychic equivalence mode' (e.g., "Often I feel threatened by the idea that someone could criticize or offend me."), and 'Regulation of affect' (e.g., Often I can't control my feelings."). The 'refusing self-reflection' scale (4 items) embodies avoidance of thinking about inner mental states or a categorical rejection of one's own feelings combined with fear of being overwhelmed by them. Higher scores indicate less ability to reflect on one's own feelings. The 'emotional awareness' scale (4 items) captures a lack of perceiving and differentiating one's own inner states. Higher scores indicate less identification of feelings and less sense of belonging. Regulation of affect is therefore not possible and feelings are experienced as very diffuse. In the 'psychic equivalence mode' scale (4 items) it is assessed to what extent inner mental states and outer reality are assimilated. Higher scores indicate that imagining a critical situation is experienced emotionally the same as an actual experienced harm. The 'regulation of affect' scale (3 items) captures the inability to modulate affect. Higher scores indicate possible feelings of helplessness and make people feel more threatened by their own feelings. Answers can be given on a five-point scale

ranging from 0 (totally disagree) to 4 (totally agree). Total scores can vary between 0 and 60 with higher scores indicating less mentalizing ability. Specific cut-off scores are not available for this instrument (Hausberg et al., 2012).

In previous research (i.e., Hausberg et al. 2012), the MZQ had a Cronbach's α for the total scale of .81, which can be considered as good (COTAN; Evers, Lucassen, Meijer & Sijtsma, 2010). Internal consistencies for the subscales ranged from .54 (Regulation of affect) to .72 (Psychic equivalence mode). These alphas have been considered insufficient to satisfactory according to the COTAN criteria (Evers et al., 2010). Test-retest coefficients can be considered sufficient for the total scale (r =.76) and insufficient for the subscales ($.60 \geq r \leq .68$). Criterion-related validity has found to be good (Hausberg et al., 2012). The questionnaire's convergent validity was good, but the divergent validity was insufficient (Hausberg et al., 2012). Cut-off scores and norms are not yet available (Hausberg et al., 2012).

The Davis Empathy Scale (DES). The DES (Davis, 1980) is a self-report questionnaire that measures different aspects of empathy. The difference with other empathy questionnaires, is that the DES holds individual differences in account (Davis, 1980). The DES consists of four subscales each containing seven items: 'Fantasy' (e.g., "I really get involved with the feelings of the characters in a novel."), 'Perspective Taking' (e.g., "Before criticizing somebody, I try to imagine how I would feel if I were in their place."), 'Empathic Concern' (e.g., "I often have tender, concerned feelings for people less fortunate than me."), and 'Personal Distress' (e.g., "I tend to lose control during emergencies."). The 'fantasy' scale (FS), measures the extent to which someone can convert oneself into imaginary situations (e.g., books, movies, daydreams). The 'perspective-taking' scale (PT), also measures the ability to transpose oneself in another person/situation, but the difference with the FS scale is that the PT scale reflects the "real life" and not imaginary situations. The other two subscales

highlight the individual differences in emotional responses to observed emotionality in others. Whereas on the one hand the 'empathic concern' scale (EC), is assessing the degree to which the participant experiences feelings of warmth, compassion and concern for the observed individual and the 'personal distress' scale (PD), on the other hand, measures the individual's own feelings of fear, apprehension and discomfort at witnessing negative experiences of others. Answers can be given on a five-point scale anchored by 0 (does not describe me well) and 4 (describes me very well). Total scores can vary between 0 and 112 and a higher score indicate more perceived empathy. Cut-off scores are not available for this instrument (Davis, 1980).

Internal reliability can be considered as sufficient (α ranging from .70 to .78) according to the COTAN criteria (Evers et al., 2010). Inter-correlations range from *r*=-.29 (Perspective taking-Personal distress) to *r*=.33 (Empathic concern-Personal Distress). Test-retest reliability can be considered satisfactory to good (α ranging from .61 to .81). Evidence has been found for sex differences on all four subscales (Davis, 1980).

Mindful Attention Awareness Scale (MAAS). The MAAS (Brown & Ryan, 2003) is a self-report questionnaire which assesses individual differences in the frequency of mindful states over time. The questionnaire has a total of 15 items and has a single-factor solution. The MAAS includes items like: "I could be experiencing some emotion and not be conscious of it until some time later" and "I forget a person's name almost as soon as I've been told it for the first time". Answers can be given on a 6-point Likert scale ranging from 1 (almost always) to 6 (almost never), where high scores reflect more awareness of and receptive to inner experiences, and more mindfulness of behavior. Total scores vary between 15 to 90, and specific cut-off scores are not available for this instrument (Simmons & Lehman, 2012).

Internal consistency of responses are considered good (α ranging from .82 to .87) according to the COTAN criteria (Evers et al., 2010). Item-total correlations varied from .25

to .72. The convergent and divergent validity can be considered as good (Brown & Ryan, 2003). Test-retest score agreement did not significantly differ (t (59) =.11) and intra-class correlation was .81, indicating good reliability.

Toronto Alexithymia Scale (TAS-20). The TAS-20 (Bagby et al., 1994a) is a 20-item self-report questionnaire that consists of three subscales: difficulty describing feelings (7 items), difficulty identifying feelings (5 items) and externally orientated thinking (8 items). The 'difficulty describing feelings' scale includes items like: "It is difficult for me to find the right words for my feelings" and "I am able to describe my feelings easily", the 'difficulty identifying feelings' scale includes items such as: "I am often confused about what emotion I am feeling" and "I don't know what's going on inside me", and the 'externally orientated thinking' scale includes items like: "I prefer to analyze problems rather than just describe them" and "Being in touch with emotions is essential". Answers can be given on a five-point scale ranging from 1 (Strongly disagree) to 5 (Strongly agree). Total scores range from 20 to 100. The TAS-20 uses the cutoff scoring: equal to or less than 51 = non-alexithymia, between 52 to 60 = possible alexithymia, and equal to or greater than 61 = alexithymia.

This scale is used in clinical and non-clinical populations (Bagby et al., 1994a).There is a satisfactory to good internal consistency ($\alpha = .81$) and this also counts for the three subscales: (difficulty describing feelings $\alpha = .78$, difficulty identifying feelings $\alpha = .75$, and externally orientated thinking $\alpha = .66$). Test-retest gives a Cronbach's α of .77 which can be considered as sufficient (COTAN; Evers et al., 2010). According to Bagby, Parker and Taylor (1994b) the TAS-20 has adequate levels of convergent and concurrent validity and also modest support for the divergent validity.

Subtypes of Antisocial Behavior Questionnaire (STAB). The STAB (Burt & Donnellan, 2009) is a 32-item self-report questionnaire for the assessment of antisocial behavior. Burt and Donnellan (2009) found that antisocial behavior exists out of three

subtypes, each with different developmental pathways, demographic patterns, correlates, and etiological foundations. Hence, the three subscales of the STAB: physical aggression (10 items), rule-breaking (11 items) and social aggression (11 items). The physical aggression scale comprises items like: “Felt like hitting people” and “Cursed or yelled at someone”. The rule-breaking scale is concerned with violating rules and is not necessarily directed towards specific others. This scale consists of items like: “Had trouble keeping a job” and “Broke into a store, mall or warehouse”. The social aggression scale includes both nonverbal and verbal forms of social exclusion and consists of items like: “Blamed others” and “Intentionally damaged someone’s reputation”. Answers can be given with a five-point scoring system ranging from 1 (never) to 5 (nearly all the time). Total scores range from 32 to 160, where higher scores indicate more display of antisocial behavior.

The factorial validity of this questionnaire has been confirmed in three different samples (i.e., college students, community adults and adjudicated adults) (Burt & Donnellan, 2009). There is an acceptable internal consistency, αs ranged from .84 to .91 for physical aggression, from .71 to .87 for rule-breaking, and from .83 to .90 for social aggression. Averaged inter-item correlations were r =.37 for physical aggression, r =.38 for rule-breaking, and r =.36 for social aggression. Burt and Donnellan (2009) also found consistent support for the criterion-related validity.

Statistical analysis

The following questionnaires were analyzed: MZQ, DES, MAAS, TAS-20, and STAB. Statistical analysis was performed with IBM-SPSS version 24.0. First, I recoded some of the items that were formulated in the reverse fashion (i.e., of the DES and TAS-20). To simplify interpretation between questionnaires, I recoded all items of the MZQ, making higher scores indicate more mentalizing ability. After trimming outliers to two standard deviations from the mean, I examined all questionnaires for normal distribution. Regarding

reliability, I interpreted Cronbach's alpha. Thereafter, an exploratory factor analysis (EFA) was conducted to assess whether the items of the MZQ could be grouped into meaningful factors. In advance I looked at the Kaiser–Meyer–Olkin (KMO) and Bartlett's tests. In order to reliably use factor analysis for data analyses, the KMO test should be greater than 0.60 (Tabachnick & Fidell, 2007) and Bartlett's test of sphericity has to be significant. For the EFA I used three methods. The first test that was used to assess the optimal number of factors was Cattell's scree test (1966). Based on looking at the elbow in the scree-plot and at the eigenvalues (Zhu & Ghodsi, 2006), the number of factors should be identified. Next, the Velicer's (1976) Minimum Average Partial (MAP) test was executed. This test focuses on "the relative amounts of systematic and unsystematic variance remaining in a correlation matrix after extractions of increasing number of components" (O'Connor, 2000, p. 5). Finally, parallel analysis was performed. Within this analysis, eigenvalues of the original data are compared with eigenvalues from 1000 random datasets (O'Conner, 2000). To assess the number of items that could be retained in the final factor solution, we looked at the communalities. Communalities are the proportion of each variables variance that can be explained for by all factors, whereas communalities lower than .10 indicate low common variances between items (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Subsequently, I conducted descriptive analyses (i.e., means, standard deviations, and range) of all questionnaires and their subscales.

Next, for examining the divergent and convergent validity, mean total scores were calculated to investigate the relationship between the MZQ and other questionnaires (i.e., DES, MAAS, TAS-20) by looking at the Pearson's correlation coefficients. Mean scores of the subscales of the DES and TAS-20 were also calculated and compared with the MZQ.

Subsequently, four multiple regression analysis were conducted with the total STAB and with the three subscales of the STAB as dependent variables. Using the enter-method,

independent variables were simultaneously added into the regression analysis, resulting in three models. The first model included control variables such as age, gender, and intelligence (i.e. education level), because previous research showed that they also might have influence on the dependent variable (Bartusch, Lynam, Moffitt, & Silva, 1997; Barriga, Morrison, Liau, & Gibbs, 2001; Kandel et al., 1988). Bartusch et al. (1997) found that the number of people who display antisocial behavior decreases with age. Barriga et al. (2001) found that males display more antisocial behavior than females. Kandel et al., (1988) found that higher IQ results in displaying less antisocial behavior. For the gender variable I created a dummy variable coded: 0 = 'female' and 1 = 'male'. The second model added mentalization and the third model added empathy, mindfulness and alexithymia. The effect-size R^2 change was used to see if the model significantly improved relative to the previous model. The betas were used to see which variable was a significant predictor. A significance level of p=.05 was applied.

Results

Assumptions

All predictors were continue or categorical and all dependent variables were continue. The item "Broke into a store" of the STAB showed no variance and therefore was not used for further analysis. All measures were normally distributed (Skewness ranged from -.047 to .747; Kurtosis ranged from -.496 to .024).

All questionnaires and their subscales had sufficient to good reliability according to the COTAN criteria (Evers et al., 2010). Mentalization had a Cronbach's α of .812 and mindfulness had a Cronbach's α of .861. Empathy had a Cronbach's α of .847 and the subscales were ranging from .663 to .767. Alexithymia had an alpha of .836 with the subscales varying from .616 to .826. Antisocial behavior had a Cronbach's alpha of .872 with the subscales ranging between .727 and .837.

Regarding the multiple regression analyses, we examined a few additional assumptions. Tests for multicollinearity indicated that a very low level of multicollinearity was present between the MZQ and the other questionnaires (VIF ranged from 1.018 to 1.983; Tolerance ranged from .504 to .982). In addition, the values of the Durbin-Watson test (ranging from 1.811 to 1.994) indicated no correlation between the residuals. Finally, the residuals were found to be normally distributed.

Exploratory Factor analysis

The Kaiser–Meyer–Olkin (KMO) measure was .791. The Bartlett's test of sphericity was significant, indicating that the correlation matrix was not identical (i.e., $\chi^2 = 588.724$, $df = 105$, $p<.001$). Both tests suggested that it is feasible to conduct an exploratory factor analysis.

Cattell's scree test (1966) identified a one-factor model based on the eigenvalues and the elbow in the scree-plot. The revised Velicer's MAP test (Velicer, Eaton, & Fava, 2000) also identified one factor. Finally, the parallel analysis, after simulating 1000 random datasets

revealed an optimal factor solution of one. Hence, we computed the one-factor model for the whole sample. All items positively correlated with that factor, with factor loadings varying from .605 to .483. Two items had factor loadings lower than .230, the items “Explanations of others are of little assistance in understanding my feelings” and “Sometimes I only become aware of my feelings in retrospect”. These items also had low common variances (<.10). However, removal of these two items did not show significant changes in the reliability of the MZQ, and were therefore not excluded from further analysis.

Descriptive statistics

In table 1, I report means, standard deviations, and ranges of all questionnaires including their subscales.

Table 1

Descriptive statistics of all study variables

Variables	N	Mean	Standard deviation	Range
Mentalization	173	3.598	.636	2.000 – 5.000
Empathy	153	3.175	.457	2.140 – 4.320
Fantasy	152	3.116	.769	1.430 – 5.000
Empathic Concern	151	3.644	.580	2.000 – 4.860
Perspective taking	149	3.548	.656	1.570 – 5.000
Personal Distress	149	2.395	.565	1.290 – 4.000
Mindfulness	163	3.912	.630	2.560 – 5.270
Alexithymia	149	2.210	.530	1.000 – 3.600
Difficulty describing feelings	145	2.395	.857	1.000 – 5.000
Difficulty identifying feelings	148	1.921	.700	1.000 – 4.140
Externally oriented thinking	146	2.349	.552	1.250 – 3.750
Antisocial behavior	147	1.351	.254	1.000 – 1.970
Physical aggression	143	1.442	.356	1.000 – 2.820
Social aggression	142	1.573	.469	1.000 – 3.400
Rule-breaking behavior	142	1.093	.237	1.000 – 3.090

Correlations

To examine the convergent and divergent validity of the MZQ, Pearson correlations with the DES, MAAS and TAS-20 were examined. As expected, bootstrapped correlation analyses showed large significant correlations with the MZQ and measures of mindfulness and alexithymia (mindfulness r = .418, p<.001; alexithymia r = -.584, p<.001). Individuals who report more mindfulness and report less alexithymia, report more mentalizing ability. However, no significant correlation was found between mentalization and empathy. When

examining the subscales of empathy, positive correlations were found with the perspective taking scale (r = .294, p<.001), and negative correlations were found with the personal distress scale (r = -.312, p<.001).

Regression

Four hierarchical multiple regression analyses were conducted to examine whether mentalization is a significant predictor in antisocial behavior, in addition to the measures of empathy, mindfulness, and alexithymia. All four multiple regressions consisted out of three models. The control variables age, gender and education level were entered at model 1 of the regression. Mentalization was entered at model 2 and empathy, mindfulness and alexithymia were entered at model 3. The first multiple regression is a hierarchical multiple regression with antisocial behavior as the dependent variable. Results were presented in table 2.

Table 2

Summary of Hierarchical Regression Analysis for Variables predicting Antisocial Behavior (N=132)

	Model 1			Model 2			Model 3		
Variable	*B*	*SE B*	*β*	*B*	*SE B*	*β*	*B*	*SE B*	*β*
Age	-.009	.001	-.497***	-.009	.001	-.497***	-.008	.001	-.448***
Gender	.067	.041	.132	.051	.039	.099	.042	.041	.083
Education	-.040	.015	-.205**	-.029	.014	-.148*	-.026	.015	-.136
Mentalization				-.118	.032	-.276***	-.082	.043	-.191
Empathy							-.025	.044	-.044
Mindfulness							-.059	.036	-.147
Alexithymia							.014	.046	.029
ΔR^2				.071***			.016		

*Note: R^2 was .262 for model 1 (p<.001). *p<.05, **p<.01, ***p<.001.*

The hierarchical multiple regression revealed that in Model 1, age and education level contributed significantly to the regression model, F (3,129) = 15.266, p<.001, and accounted for 26.2% of the variation in antisocial behavior. Both had a significant negative effect, indicating that an increase in age and education level leads to displaying less antisocial behavior (p<.001 for age and p=.008 for education level). Introducing mentalization explained an additional 7.1% of variation in antisocial behavior and this change in R^2 was significant F(1.128) = 13.671, p<.001. Mentalization had a significant negative effect, indicating that an increase in mentalization leads to displaying less antisocial behavior (p<.001). Age and education level were also significant in model 2, indicating that they had a negative contribution on the relationship between mentalization and antisocial behavior (p<.001 for age and p=.048). Adding empathy, mindfulness, and alexithymia to the regression model explained no significant additional variation in antisocial behavior (p=.388). When all independent variables were included in model 3 of the regression model, only age was a significant predictor of antisocial behavior (p<.001).

The second multiple regression is a three stage hierarchical multiple regression with physical aggression as the dependent variable. Results of this multiple regression were presented in table 3.

Table 3

Summary of Hierarchical Regression Analysis for Variables predicting Physical Aggression (N=128)

	Model 1			Model 2			Model 3		
Variable	*B*	*SE B*	*β*	*B*	*SE B*	*β*	*B*	*SE B*	*β*
Age	-.010	.002	-.410***	-.010	.002	-.406***	-.010	.002	-.388***
Gender	.042	.061	.058	.022	.060	.030	.005	.063	.007
Education	-.062	.023	-.221**	-.050	.023	-.181*	-.051	.023	-.184*
Mentalization				-.120	.049	-.201*	-.114	.066	-.191
Empathy							-.069	.067	-.088
Mindfulness							-.070	.055	-.124
Alexithymia							-.056	.072	-.082
ΔR^2				.038*			.017		

*Note: R^2=.197 for model 1 (p<.001). *p<.05, **p<.01, ***p<.001.*

The hierarchical multiple regression revealed that in Model 1, age and education level contributed significantly to the regression model, F (3,125) = 10.210, p<.001. Age and education level together accounted for 19.7% of the variation in physical aggression. They both had a significant negative effect, indicating that an increase in age and education level leads to displaying less physical aggression (p<.001 for age and p=.007) Introducing mentalization explained an additional 3.8% of variation in physical aggression and this change in R^2 was significant, F(1.124) = 6.145, p=.015. Mentalization had a significant negative effect, suggesting that an increase in mentalization leads to displaying less physical aggression (p=.015). Age and education level were also significant in model 2, indicating that they contributed negatively on the relationship between mentalization and physical aggression (p<.001 for age and p=.027 for education level). Adding empathy, mindfulness, and

alexithymia to the regression model explained no significant additional variation in physical aggression (p=.429). When all independent variables were included in model 3 of the regression model, only age and education level were significant predictors of physical aggression (p<.001 for age and p=.031 for education level).

The third multiple regression is a three stage hierarchical multiple regression with social aggression as the dependent variable. Results were presented in table 4.

Table 4

Summary of Hierarchical Regression Analysis for Variables predicting Social Aggression (N=127)

	Model 1			Model 2			Model 3		
Variable	*B*	*SE B*	*β*	*B*	*SE B*	*β*	*B*	*SE B*	*β*
Age	-.016	.003	-.478***	-.016	.003	-.482***	-.015	.003	-.441***
Gender	.166	.080	.174*	.149	.078	.157	.135	.082	.141
Education	-.036	.029	-.100	-.016	.029	-.044	-.007	.031	-.019
Mentalization				-.193	.064	-.241**	-.118	.087	-.148
Empathy							-.044	.086	-.043
Mindfulness							-.070	.071	-.094
Alexithymia							.079	.094	.087
ΔR^2				.054**			.012		

*Note: R^2=.216 for model 1 (p<.001). *p<.05, **p<.01, ***p<.001.*

The hierarchical multiple regression revealed that in Model 1, age and gender contributed significantly to the regression model, F (3,124) = 11.394, p<.001, and accounted for 21.6% of the variation in social aggression. Both had a significant effect, meaning that an increase in age and being male leads to displaying less social aggression (p<.001 for age and p=.041 for gender). Introducing mentalization explained an additional 5.4% of variation in

social aggression and this change in R^2 was significant F(1.123) = 9.112, p=.003. Mentalization had a significant negative effect, indicating that an increase in mentalization leads to displaying less social aggression (p=.003). Age was also significant in model 2, indicating that it contributed negatively on the relationship between mentalization and social aggression (p<.001). Adding empathy, mindfulness, and alexithymia to the regression model explained no significant additional variation in social aggression (p=.588). When all independent variables were included in model 3 of the regression model, only age was a significant predictor of social aggression (p<.001).

The fourth multiple regression is a three stage hierarchical multiple regression with rule-breaking behavior as the dependent variable. Results of this multiple regression were presented in table 5.

Table 5

Summary of Hierarchical Regression Analysis for Variables predicting Rule-Breaking Behavior (N=127)

	Model 1			Model 2			Model 3		
Variable	*B*	*SE B*	*β*	*B*	*SE B*	*β*	*B*	*SE B*	*β*
Age	-.004	.002	-.245**	-.004	.002	-.247**	-.003	.002	-.178
Gender	.059	.044	.120	.052	.044	.106	.066	.045	.133
Education	-.043	.016	-.224**	-.036	.017	-.191*	-.031	.017	-.156
Mentalization				-.060	.036	-.146	.003	.048	.007
Empathy							.071	.049	.130
Mindfulness							-.031	.039	-.080
Alexithymia							.090	.051	.196
ΔR^2				.020			.041		

*Note: R^2=.108 for model 1 (p=.003). *p<.05, **p<.01, ***p<.001.*

The hierarchical multiple regression revealed that in Model 1, age and education level contributed significantly to the regression model, $F(3,124) = 4.979$, $p=.003$. Age and education level accounted for 10.8% of the variation in rule-breaking behavior. Both had a significant negative effect ($p=.007$ for age and $p=.010$ for education level), suggesting that an increase in age and education level leads to displaying less rule-breaking behavior. Introducing mentalization into the model explained no additional variation in rule-breaking behavior ($p=.097$).

Discussion

The current study further validated the MZQ, a measure for mentalization ability that was originally developed by Hausberg et al. (2012). At the beginning of this study, I had three expectations. I expected to find a different factor solution as opposed to the four-factor solution that was found in the study of Hausberg et al. (2012). I also expected that the MZQ was an appropriate and distinct questionnaire for the assessment of mentalization. In addition to this assumption, I expected that the MZQ had good convergent and divergent validity, and that mentalization was a unique predictor in the explanation of antisocial behavior, in addition to empathy, mindfulness and alexithymia.

In line with the first assumption, exploratory factor analyses suggested that mentalization could be captured with one underlying factor. This is in contrast to the four-factors found by Hausberg et al. (2012). An explanation of this finding could be addressed to the differences in sample characteristics. Hausberg et al. (2012) used inpatients in a psychiatric hospital and a psychosomatic clinic, whereas I used a large heterogeneous sample retrieved from the general population. Hausberg et al. (2012) questioned their external validity and the reliability of their factor structure. These differences in sample characteristics are therefore a possible explanation for finding a different factor structure. Another possible explanation is that Hausberg et al. (2012) used a different method for examining the factor structure. He conducted principal components analysis and used the-eigenvalues-greater-than-one rule proposed by Kaiser (1960). I used the Cattell's scree test (1966), the revised Velicer's MAP test (Velicer, Eaton, & Fava, 2000), and parallel analysis. These relatively newer methods represent a more strict way for interpreting eigenvalues as opposed to the method of Kaiser (1960), and is therefore an explanation why I only found one-factor. A third explanation of this finding is that items of the MZQ were derived from current literature on psychopathology and mentalization (Bateman & Fonagy, 2004; Bergmann-Mausfeld, 2006;

Fonagy et al., 2002; Stein, 2003). These items describe a lack in mentalization ability based on what mostly borderline patients perceive as a lack of mentalization ability (Hausberg et al., 2012). This may be different of what the general population perceive as a lack of mentalization ability.

The second assumption only partly corresponds with my expectations. As expected, reporting more mindfulness and less alexithymia was associated with reporting more mentalizing ability. This is in line with previous studies (Wallin, 2007; Swart et al., 2009; Moriguchi et al., 2006). However, in this study, reporting more empathy was not associated with reporting more mentalizing ability. This is in contradiction with the findings in some previous studies (Hooker et al., 2008; Schnell et al., 2010; Singer et al., 2004). A possible explanation for this discrepancy is that Hooker at al. (2008) used a False belief task instead of a self-report questionnaire for the assessment of mentalization. Self-report questionnaires can often underreport or overreport the truth (Donaldson & Grant-Vallone, 2002), executing a task has no such self-report bias. Another explanation is that Schnell et al. (2010) only examined the relationship between cognitive empathy and mentalization. Cognitive empathy consists of the subscales ‘fantasy’ and ‘perspective taking’ (Davis, 1983), therefore the subscales ‘empathic concern’ and ‘personal distress’ (i.e. affective empathy) were not included. Singer et al. (2004) only used the empathic concern scale in his study. In this study, I examined all subscales and found positive correlations with ‘perspective taking’ and negative correlations with ‘personal distress’. It seems to be that the MZQ is therefore more related to cognitive empathy than to affective empathy. The MZQ possibly measures in a greater extent the skill to adopt the point of view of others as opposed to the skill to connect this point of view of others to own emotions and feelings.

The third assumption was also in line with my expectations. Mentalization is a unique predictor in the explanation of antisocial behavior. It was also a unique predictor in the

explanation of physical aggression and social aggression, however not in rule-breaking behavior. This is not surprising as rule-breaking behavior has nothing to do with reading the mental states of other people (Burt & Donnellan, 2009). An explanation why empathy and mindfulness explained no unique variance in antisocial behavior may be due to the fact that both concepts seems to be a subdivision of the mentalizing framework (Cho-Kain & Gunderson, 2008). Alexithymia also had no unique variance in the explanation of antisocial behavior. A possible explanation is that alexithymia is mainly related to perspective taking (Swart et al., 2009), and in this study mentalization was related to the 'perspective taking' subscale of empathy. If alexithymia is the same as perspective taking, than it is no surprise that alexithymia had no unique value in the explanation of antisocial behavior in addition to measures of mentalization and empathy. Notwithstanding the fact that mentalization found to be a unique predictor in the explanation of antisocial behavior, physical aggression, and social aggression, age however seems to be the best predictor.

Findings of this study should be interpreted while keeping a few limitations in mind. First, the current study deviated on two aspects from the target population for which the MZQ (Hausberg et al., 2012) was initially designed. First, the MZQ is typically used as a measure of treatment effects in an inpatient mental disorder population. Second, the items of the MZQ were formulated based on case studies of borderline personality disordered patients (Bateman & Fonagy, 2004; Bergmann-Mausfeld, 2006; Fonagy et al., 2002; Stein, 2003). In contrast, this sample consisted out of non-patient males and females retrieved from the general population, which is not the MZQ's intended target population. However, the MZQ also has proven to be a reliable instrument (α=.812) to use in this population. Another limitation is that self-report questionnaires increases the chance of social desirable answers (van de Mortel, 2008). This tendency for people to present a favorable image of themselves is probably even more present if participants were recruited in the researchers own environment, as is the case

in this study. Future research should therefore control for social desirable responding, as this could have influenced this research. Another limitation is that the participants were only measured at one moment in time, whereas the study of Hausberg et al. (2012) had three measure moments. "Repeated measures provides a more powerful test of the null hypothesis because there are more sources of variability that can be extracted from error" (Girden, 1992, pp. 13). Future research should have more than one measure moment.

This research also has strong points. This large and heterogeneous sample (N = 173) is an accurate reflection of the general population. Stevens (1992) stated that it is a rule of thumb to have at least 15 participants per predictor for reliable estimates of regression analyses. In this study with seven predictors, I adhered to this rule. Furthermore, the study indicates that the MZQ is valid in both woman and men retrieved from the general population. Moreover, the MZQ is associated with antisocial behavior even in low base rate situations where such associations can be difficult to find.

For future research, it would be interesting to further examine the validity of the MZQ to increase the external validity, because comparison between the study of Hausberg et al. (2012) and this study is complicated. The relationship between mentalization and empathy is another interesting finding that needs further exploration. In this study, it seems to be that mentalization is more related to cognitive empathy than to affective empathy. Someone can have the skill to adopt the point of view of others, while not experiencing the feelings and emotions that should be connected to that point of view of others. Maybe it is possible that when people learn how to mentalize, they can fake empathy. For example, if people with antisocial personality disorder or with psychopathic traits develop more mentalizing ability, they can also learn new ways of how to manipulate someone. Moreover, Dolan and Fullam (2004) found that mentalizing ability may have an adaptive function in maintaining a criminal lifestyle. If this is the case, then mentalization based treatment should be applied with caution,

especially in the forensic care. Next, it should be noted that my study is the first to test the validity of the MZQ in a large heterogeneous general population. It is therefore difficult to provide solid recommendations for practical implications as more studies are needed to replicate these findings.

In sum, our findings suggest that the MZQ has one underlying factor in predicting mentalizing ability. It is shown to be a discrete measure for the assessment of mentalization, specifically with respect to reading mental states (e.g. point of views) of others. It has a unique role in the explanation of antisocial behavior in addition to the measures of empathy, mindfulness and alexithymia, even in situations with low base rates. The fact that the MZQ shows to be a reliable and valid measurement of mentalization in this population, therefore indicates a very strong and distinctive instrument.

References

Achim, A. M., Ouellet, R., Roy, M. A., & Jackson, P. L. (2010). Assessment of empathy in first-episode psychosis and meta-analytic comparison with previous studies in schizophrenia. *Psychiatry Research, 190,* 3-8. doi:10.1016/j.psychres.2010.10.030

Allen, J. G., & Fonagy, P. (2006). *The handbook of mentalization-based treatment.* West Sussex, England: John Wiley & Sons.

Asen, E., & Fonagy, P. (2012). Mentalization-based therapeutic interventions for families. *Journal of Family Therapy, 34,* 347–370. doi: 10.1111/j.1467-6427.2011.00552.x

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994a). The twenty-item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38,* 23-32.

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994b). The twenty-item Toronto Alexithymia Scale—II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research, 38,* 33-40. doi:10.1016/0022-3999(94)90006-X

Bartusch, D. R. J., Lynam, D. R., Moffitt, T. E., & Silva, P. A. (1997). Is age important? Testing a general versus a developmental theory of antisocial behavior. *Criminology, 35,* 13-48. doi: 10.1111/j.1745-9125.1997.tb00869.x

Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *The American Journal of Psychiatry,* 156, 1563-1569.

Bateman, A., & Fonagy, P. (2004). *Psychotherapy for Borderline Personality Disorder: Mentalization Based Treatment.* Oxford: Oxford University Press.

Bateman, A., & Fonagy, P. (2008). 8-Year Follow-Up of Patients Treated for Borderline Personality Disorder: Mentalization-Based Treatment Versus Treatment as Usual. *The American Journal of Psychiatry, 165,* 631-638. doi: 10.1176/appi.ajp.2007.07040636

Bateman, A., & Fonagy, P. (2010). Mentalization based treatment for borderline personality disorder. *World Psychiatry, 9,* 11-15. doi: 10.1002/j.2051-5545.2010.tb00255.x

Bateman, A., & Tyrer, P. (2004). Psychological treatment for personality disorders. *Advances in Psychiatric Treatment, 10,* 378-388. doi: 10.1192/apt.10.5.378

Bergmann-Mausfeld, G. (2006). Pathologische Passung, Mentalisierung und negative therapeutische Reaktion. *Forum der Psychoanalyse, 22,* 249-267. doi: 10.1007/s00451-006-0289-2

Brent, B. (2009). Mentalization-Based Psychodynamic Psychotherapy for Psychosis. *Journal of clinical psychology, 65,* 803-814. doi: 10.1002/jclp.20615

Brown, K.W. & Ryan, R.M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84,* 822-848.

Burt, S. A., & Donnellan, M. B. (2009). Development and validation of the Sub-Types of Antisocial Behavior questionnaire (STAB). *Aggressive Behavior, 35,* 376–398.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245–276. http://dx.doi.org/10.1207/s15327906mbr0102_10

Choi-Kain, L. W., & Gunderson, J. G. (2008). Mentalization: Ontogeny, Assessment, and Application in the Treatment of Borderline Personality Disorder. *The American journal of psychiatry, 165,* 1127-1135. doi: 10.1176/appi.ajp.2008.07081360

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297-334.

Daudert, E. (2002). Die Reflective Self Functioning Scale. In B. Strauss, H. Kaechele, & A. Buchheim (Eds.), *Klinische Bindungsforschung. Theorien, Methoden, Ergebnisse (pp. 54-67).* Stuttgart: Schattauer.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10,* 85.

Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44,* 113-126. doi: 10.1037/0022-3514.44.1.113

Dolan, M., & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological Medicine, 6,* 1093-1102. http://dx.doi.org/10.1017/S0033291704002028

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology, 17,* 245-260. doi:10.1023/A:1019637632584

Evers, A., Lucassen, W., Meijer, R., & Sijtsma, K. (2010). *COTAN Beoordelingssysteem voor de kwaliteit van tests.* (revised version, may 2009). Retrieved from http://www.psynip.nl/website-openbaar-documenten-nip-algemeen/beoordelingssysteem.pdf

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4,* 272–299. doi: 10.1037/1082-989X.4.3.272

Foa, E. B., Cashman, L., Jaycox, L., & Perry, K. (1997). The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. *Psychological assessment, 9,* 445-451. doi: 10.1037/1040-3590.9.4.445

Fonagy, P. (1989). On tolerating mental states: Theory of mind in borderline patients. *Bulletin of the Anna Freud Centre, 12,* 91-115.

Fonagy, P., & Bateman, A. W. (2006). Mechanisms of change in mentalization-based treatment of BPD. *Journal of Clinical Psychology , 62,* 411-430. doi: 10.1002/jclp.20241

Fonagy, P., Gergely, G., Jurist, E., & Target, M. (2002). *Affect Regulation, Mentalization and the Development of the Self.* New York: Other Press LLC.

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. doi:10.1017/S0954579409990198

Fonagy, P., Luyten, P., & Strathearn, L. (2011). Borderline personality disorder, mentalization, and the neurobiology of attachment. *Infant Mental Health Journal, 32,* 47-69. doi: 10.1002/imhj.20283

Fonagy, P., Target, M., Steele, M., & Steele, H. (1997). *Reflective-functioning manual: For application to Adult Attachement Interviews.* London: University College London.

Fonagy, P., Twemlow, S. W., Vernberg, E. M., Mize Nelson, J., Dill, E. J., & Little, T. D. (2009). A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression. *Journal of Child Psychology and Psychiatry, 50,* 607–616. doi: 10.1111/j.1469-7610.2008.02025.x

Fossati, A., Acquarini, E., Feeney, J. A., Borroni, S., Grazioli, F., Giarolli, L. E., Franciosi, G., & Maffei, C. (2009). Alexithymia and attachment insecurities in impulsive aggression. *Attachment & Human Development, 11,* 165-182. doi: 10.1080/14616730802625235

Frith, C. D. (2006). The Neural Basis of Mentalizing. *Neuron, 50,* 531–534. doi: 10.1016/j.neuron.2006.05.001

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences, 7,* 77-83. doi:10.1016/S1364-6613(02)00025-6

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

Girden, E. (1992). *ANOVA: Repeated Measures.* California: Sage publications.

Hausberg, M. C., Schulz, H., Piegler, T., Happach, C. G., Klöpper, M., Brûtt, A. L., Sammet, I., & Andreas, S. (2012). Is a self-rated instrument appropriate to assess mentalization in patients with mental disorders? Development and first validation of the Mentalization Questionnaire (MZQ). *Psychotherapy Research,* 1-11. doi:10.1080/10503307.2012.709325

Hooker, C. I., Verosky, S. C., Germine, L. T., Knight, R. T., & D'Esposito, M. (2008). Mentalizing about emotion and its relationship to empathy. *Social Cognitive & Affective Neuroscience, 3,* 204-217. doi: 10.1093/scan/nsn019

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141-151. http://dx.doi.org/10.1177/001316446002000116

Kandel, E., Mednick, S. A., Kirkegaard-Sorensen, L., Hutchings, B., Knop, J., Rosenberg, R., Schulsinger F. (1988). IQ as a protective factor for subjects at high risk for antisocial behavior. *Journal of Consulting and Clinical Psychology, 56,* 224-226. http://dx.doi.org/10.1037/0022-006X.56.2.224

Kemps, S. J. L. M., & Kooiman, C. G. (2015). Het meten van mentaliseren; een poging met de Reflective Functioning Rating Scale. *Tijdschrift voor Psychiatrie, 57,* 645-655.

Kernberg, O. F., Diamond, D., Yeomans, F. E., Clarkin, J. F., Levy, K. N., Jurist, E. L., Slade, A., & Bergner, S. (2008). Mentalization and attachment in borderline patients in transference focused psychotherapy. *Mind to mind: Infant research, neuroscience, and psychoanalysis,* (pp. 167-201). New York, NY, US: Other Press, ix, 454 pp.

Langdon, R., & Coltheart, M. (2001). Visual perspective-taking and schizotypy: evidence for a simulation-based account of mentalizing in normal adults. *Cognition, 82,* 1-26. doi:10.1016/S0010-0277(01)00139-1

Leichsenring, F., Kunst, H., & Hoyer, J. (2003). Borderline personality organization in violent offenders: Correlations of identity diffusion and primitive defense mechanisms with antisocial features, neuroticism, and interpersonal problems. *Bulletin of the Menninger Clinic, 67,* 314-327. doi: 10.1521/bumc.67.4.314.26983

Luyten, P., Fonagy, P., Lowyck, B., & Vermote, R. (2012). The assessment of mentalization. In A. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice (pp. 43-65).* Washington, DC: American Psychiatric Association.

Manninen, M., Therman, S., Suvisaari, J., Ebeling, H., Moilanen, I., Huttunen, M., & Joukamaa, M. (2011). Alexithymia Is Common Among Adolescents with Severe Disruptive Behavior. *Journal of Nervous & Mental disease, 199,* 506-509. doi: 10.1097/NMD.0b013e3182214281

Marshall, L. E., & Marshall, W. L. (2011). Empathy and antisocial behavior. *Journal of Forensic Psychiatry & Psychology, 22,* 742-759. doi: 10.1080/14789949.2011.617544

McGauley G., Yakeley J., Williams A., & Bateman, A. (2011). Attachment, mentalization, and antisocial personality disorder: the possible contribution of mentalization-based treatment. *European Journal of Psychotherapy & Counselling, 13,* 371-393. doi: 10.1080/13642537.2011.629118

Meehan, K. B., Levy, K. N., Reynoso, J. S., Hill, L. L., & Clarkin, J. F. (2009). Measuring reflective function with a multidimensional rating scale: comparison with scoring reflective function on the AAI. *Journal of the American Psychoanalytic Association, 57,* 208-213. doi: 10.1177/00030651090570011008

Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological bulletin, 103,* 324-344. doi: 10.1037/0033-2909.103.3.324

Moriguchi, Y., Ohnishi, T., Lane, R. D., Maeda, M., Mori, T., Nemoto, K., Matsuda, H., & Komaki, G. (2006). Impaired self-awareness and theory of mind: An fMRI study of mentalizing in alexithymia. *NeuroImage, 32,* 1472-1482. doi:10.1016/j.neuroimage.2006.04.186a

Mortel, T. F. van de (2008). Faking it: social desirability response bias in self-report research. *Australian Journal of Advanced Nursing, 25*, 40-48.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32,* 396–402. http://dx.doi.org/10.3758/BF03200807

Ouwersloot, G., Brink, W. van den, Diekstra, R. F. W., & Hoogduin, C. A. L. (1994). Diagnostiek van persoonlijkheidsstoornissen. Een evaluatie van Nederlandstalig instrumentarium. *Tijdschrift voor psychiatrie, 8,* 1-14. Retrieved from http://www.tijdschriftvoorpsychiatrie.nl/assets/articles/articles_1055pdf.pdf

Rossouw, T. I., & Fonagy, P. (2012). Mentalization-Based Treatment for Self-Harm in Adolescents: A Randomized Controlled Trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 51,* 1304-1313. doi:10.1016/j.jaac.2012.09.018

Schnell, K., Bluschke, S., Konradt, B., & Walter, H. (2011). Functional relations of empathy and mentalizing: An fMRI study on the neural basis of cognitive empathy. *Neuroimage, 54,* 1743-1754. doi: 10.1016/j.neuroimage.2010.08.024

Simmons, C. A., & Lehman, P. (2012). *Tools for Strengths-Based Assessment and Evaluation.* New York: Springer Publishing Company.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science, 303,* 1157-1161. doi: 10.1126/science.1093535

Singh, N. N., Lancioni, G. E., Joy, S. D. S., Winton, A. S. W., Sabaawi, M., Wahler, R. G., & Singh, J. (2007). Adolescents With Conduct Disorder Can Be Mindful of Their Aggressive Behavior. *Journal of Emotional and Behavioral Disorders, 15,* 56-63. doi: 10.1177/10634266070150010601

Söderström, K., & Skårderud, F. (2009). MINDING THE BABY Mentalization-based treatment in families with parental substance use disorder: Theoretical framework. *Nordic Psychology, 61,* 47-65. doi: 10.1027/1901-2276.61.3.47

Stein, H. (2003). Mentalization with reservations. *Bulletin of the Menninger Clinic, 67,* 143-149. doi: 10.1521/bumc.67.2.143.23446

Stevens, J. (1992). *Applied multivariate statistics for the social sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Swart, M., Kortekaas, R., & Aleman, A. (2009). Dealing with Feelings: Characterization of Trait Alexithymia on Emotion Regulation Strategies and Cognitive-Emotional Processing. *PLoS ONE, 4(6):* e5751. doi:10.1371/journal.pone.0005751

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics (5th ed.).* Boston, MA: Pearson Education.

Taubner, S., Kessler, H., Buchheim, A., Kächele, H., & Staun, L. (2011). The Role of Mentalization in the Psychoanalytic Treatment of Chronic Depression. *Psychiatry: Interpersonal and Biological Processes, 74*, 49-57. doi: 10.1521/psyc.2011.74.1.49

Taylor, N., & Signal, T. D. (2005). Empathy and attitudes to animals. *Anthrozoös, 18,* 18-27. doi: 10.2752/089279305785594342

Vandereycken, W., Hoogduin, C. A. L., & Emmelkamp, P. M. G. (2011). *Handboek psychopathologie deel 2 klinische praktijk.* (3rd ed.). Houten, Nederland: Bohn Stafleu van Loghum.

Velicer, W. F. (1976). Determining the number of components form the matrix of partial correlations. *Psychometrika, 41,* 321–327. http://dx.doi.org/10.1007/BF02293557

Velicer, W. F., Eaton, C. A., and Fava, J. L. (2000). *Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. Pp. 41-71 in R. D. Goffin and E. Helmes, eds., Problems and solutions in human assessment.* Boston: Kluwer.

Wallin, D. J. (2007). *Attachment in psychotherapy.* New York: Guilford Press.

Zhu, M., & Ghodsi, A. (2006). Automatic dimensionality selection from the scree plot via the use of profile likelihood. *Computational Statistics & Data Analysis, 51,* 918-930. doi:10.1016/j.csda.2005.09.010