A Psychometric Evaluation of the Defense Style Questionnaire-40 in a UK Forensic Patient Population

James Tapp\textsuperscript{a}, Louise Cottle\textsuperscript{b}, Michelle Christmas\textsuperscript{a}, Rose Stratton\textsuperscript{a}, Theresa A. Gannon\textsuperscript{b} and Estelle Moore\textsuperscript{a}

Affiliations

\textsuperscript{a} Centralised Groupwork Service, Broadmoor Hospital, Crowthorne, Berkshire, RG45 7EG

\textsuperscript{b} CORE-FP, School of Psychology, University of Kent, Canterbury, Kent, CT2 7NP
Abstract

Psychological defense mechanisms have been considered important personality processes in the onset, maintenance, and recovery of mental disorders. More recently their application to understanding presenting problems and as potential outcome indicators for forensic mental health patients has been recommended. However, to date there have been no investigations into the reliability and factor structure of defense mechanism assessments for this population. The current study investigated the factor structure and internal consistency and test re-test reliability of the Defense Style Questionnaire-40 for 160 adult male UK forensic patients. The three-factor model of defences proposed by the DSQ-40 developers was not confirmed in the study sample. Reliability indices of the three factors indicated that the Immature factor was the most ‘acceptable’ in terms of internal consistency. Test-retest reliability coefficients ranged from .70 to .91. A revised three-factor structure that closely corresponds to the original validation study is recommended following an exploratory factor analysis. The findings are compared with previous reliability and factor analytic evaluations of the DSQ-40, and recommendations for its use with forensic patients are discussed.

Keywords: forensic patients, defense style questionnaire, assessment, reliability, confirmatory factor analysis.
A Psychometric Evaluation of the Defense Style Questionnaire-40 in a UK Forensic Patient Population

Defense mechanisms are defined as automatic, often involuntary psychological processes that occur out of a person’s awareness, protecting them from anxieties and other internal or external stressors (Bond & Vaillant, 1986; Cramer, 1991; Freud, 1926). Initially originating in psychodynamic theory, defenses were considered functional to protecting one’s sense of self by offering relief (through the defense of repression) from instinctive drives, in particular sexual and aggressive impulses, which would challenge a person’s internalized standards (Baumeister, Dale & Sommer, 1998). Based on this function, defenses were viewed as important adaptive processes in coping with difficult emotions. However, when occurring excessively or ineffectively, it has been argued that these mechanisms can distort how reality is made sense of and negatively impact upon healthy psychological adjustment and relationships with others (Bond, 2004; Cramer, 1987; Johnson, Bornstein, & Krukonis, 1992).

The early concepts of defense mechanisms came under empirical scrutiny, and the existence of such processes was questioned on the basis of competing explanations (see Holmes, 1968, 1978). However, the perceived clinical validity and utility of psychological defenses reported by practitioners, and theoretical interest from other fields of enquiry beyond psychoanalysis supported their continued investigation in the field of personality assessment (Cooper, 1998; Cramer, 2000; Fenichel, 1945). To date, defense mechanisms have been investigated across a broad range of fields in psychology including: the aetiology and maintenance of, and recovery from, mental disorders (Andrews et al., 1993; Bond, 2004; Vaillant, 1971, 1994); physical well-being (MacGregor et al., 2003); social functioning (Vaillant,
1977) and childhood and adolescent development (Cramer, 1997). To give an example for the latter field, investigations into strong defense use in adolescents have demonstrated reduced levels of moral judgement. For a useful review of the ways in which defense mechanisms have been related to psychosocial problems see Cramer (2000).

One arena in which psychological defenses have received less attention, but are arguably relevant given existing research, is in forensic mental health. Forensic mental health patients are often described as presenting with complex, co-morbid mental disorders, including personality disorder, that are connected with high-risk behaviours, such as interpersonal violence (Taylor, Grounds & Snowden, 1993). Individuals in forensic mental health settings are also likely to have experienced childhood abuse, neglect, and/ or familial disruption (e.g., bereavement; domestic violence; Taylor, 1998), which are considered to impact on personality development and relationship functioning in adult life (Nickel & Egle, 2006). For the task of understanding and treating the wide range of needs reported for this particular patient group, there are a number of potential contributions that the concept of defense mechanisms can make to formulating both intra and interpersonal difficulties (Vaillant, 1992).

A challenge that forensic patients face in their contact with services is encountering the (self and public) stigma of having a psychiatric disorder and an offender identity (Menditto, 2002). Identification with these labels as being part of one’s personality perhaps expectedly can give rise to distress and shame (Corrigan, 2004). To defend against these, patients may conceal or deny distress, symptom experience, or risk (Cramer, 2000) and may even decline to participate in psychological therapy (Bowins, 2004; Thygesen, Drapeau, Trijsburg, Lecours, & de
Roten, 2008). On this basis it has been recommended that patient defenses are considered in determining suitability and responsivity for treatment approaches (Vaillant, 1994).

In connection with treatment planning, the assessment of defense mechanisms of forensic patients may also indicate important outcomes for this patient group that have tended to be overlooked (Huband, Duggan, McCarthy, Mason & Rathbone, 2014). These may relate to risk reduction, where the deployment of defense mechanisms has led to dangerous behaviours such as violence (Apter et al., 1989; Bateman, 1996) and mental health recovery, as previously referenced (Bond, 2004). An awareness of defence mechanisms of forensic patients may also be useful for working therapeutically with individuals who often have a history of transgressing peer and professional boundaries (Horvath, 2001; Moore, 2012). The challenge of identifying and reflecting on interpersonal exchanges in forensic settings that might be influenced by defensive processes is important in maintaining safe alliances, limiting the risk of boundary violations, and promoting an understanding of the use of defenses in relationships, including therapeutic alliances.

Given these potential implications for forensic practice an important objective is to determine how best to assess defense mechanisms. An initial challenge in this task is that the types and functions of defense mechanisms have been continually debated and revised since first introduced. For example, 17 defenses were initially described by Freud (1926), with subsequent developments, aimed at providing increasingly investigable and valid concepts (Freud, 1968; Kernberg, 1967; Vaillant, 1976; Klein, 1975). Whilst debate and refinement has and will likely continue, a common set of defenses can be outlined across competing positions that are also captured in the existing assessments of defense mechanisms (see Table 1). A further
challenge in the assessment of defenses lies in their experience as being outside of a person’s awareness, which has potential implications, certainly for self-reporting (Bond, 2004; Davidson & MacGregor, 1998). Defense researchers have argued that whilst the activation or function of defenses may be difficult to be aware of, resulting behaviour from them may be observable, and individuals may also be able to reflect on attitudes or thoughts and behaviours when experiencing emotional stress, and therefore identify defensive processes (Andrews et al., 1993; Bond, 1986; Cramer, 1991; Vaillant, 1994). This it has been argued is the difference between defense mechanisms and defense (coping) styles, with the latter being a conscious and intentional effort by a person to overcome uncomfortable emotions (Cramer, 1998). It is at this intentional adaptation level that the current study focuses, in particular how defense styles might be reliably and validly assessed.

Insert table 1 about here

There are a number of existing self-report assessments of defense mechanisms including the Defense Mechanism Inventory (Gleser & Ihilevich, 1969), the Defense Mechanism Rating Scale (Perry & Cooper, 1986), the Coping and Defending Scales (Joffe & Naditch, 1977), the Life-Style Index (Plutchik, Kellerman, & Conte, 1979) and the Defense Style Questionnaire (DSQ; Bond, Gardner, Christian & Sigal, 1983). These have been variably tested in terms of reliability parameters and construct validity and their application to clinical populations (see Davidson & MacGregor, 1998 for a review). The one cited as the most widely researched, most consistent with psychiatric nosology and most extensively used with clinical samples, is the Defense Style Questionnaire (Bond, 2004; Vaillant, 1986).
The 40 item version of the DSQ (DSQ-40; Andrews et al., 1993) is one of the latest iterations of the assessment following revisions of earlier versions (Andrews, Pollock & Stewart, 1989; Bond et al., 1983; Bond et al., 1989; Vaillant, Bond & Vaillant, 1986) for reasons of: unreliable discrimination between patient and community samples, unclearly phrased items, inadequate explanation of item selection, fatigue effects, and unequal item representations for defenses (Saint-Martin, Valls, Rousseau, Callahan & Chabrol, 2013; Thygesen et al., 2008). The DSQ-40 assesses 20 different defenses that are hierarchically categorized as three key defense styles: immature, mature and neurotic. These three styles have been argued to be more robust than previous DSQ factor solutions, which were based on sample sizes inadequate to warrant factor analysis methods (Thygesen et al., 2008).

Since the development of the DSQ-40, investigations into its factor structure and reliability (internal consistency and test re-test) in various settings have produced mixed findings, whereby differing factor structure solutions have been proposed, ranging from 3 to 6 factors, and internal consistency coefficients for the three factors proposed by the developers have varied from .51 to .85 (Andrews et al., 1993; Lopez & Gormley, 2002; Ruutu et al., 2006; Watson & Sinha, 1998). Based on these findings, and the absence of psychometric investigations in forensic mental health settings it is difficult to extrapolate whether the DSQ-40 would be a reliable and valid assessment for forensic mental health patients. The current study therefore investigated the factor structure; internal consistency and test re-test reliability of the DSQ-40 with forensic patients detained in a UK secure hospital setting.
Method

Design

A cross sectional correlational study design was applied to investigate the factor structure and internal consistency of the DSQ-40 among a sample of forensic mental health patients. A further prospective design was applied to investigate the test re-test reliability of the assessment.

Participants

Participants were 160 adult (>18yrs) male forensic mental health patients detained under the Mental Health Act 2007 (HMSO, 2007) in a British high security hospital. Inclusion criteria for the study were: capacity to consent to participate in research and English speaking. Exclusion criteria were: lacking capacity for consent to participate in research; non-English speaking and individual risk to others too high. The mean age of participants was 42 years (s.d. 11.54). Participants represented the broader demographic of the study site, which included being predominantly of White Caucasian ethnicity, (59.4%); with an index offence of violence (29.4%); and a primary diagnosis of schizophrenia (55.6%). Full demographic features of the participants are outlined in Table 2.

Insert table 2 about here

Procedure

One hundred and sixty prospective participants were initially identified from correspondence with clinical care teams at the study site. Of these 135 were
approached in the time frame of the study, and provided with an information sheet and consent form. Fifty-three opted in to the study and completed the DSQ-40 at time 1. After an 8-week interval participants were re-approached to complete a further DSQ-40; one participant was withdrawn from stage 2 of the study over mental capacity concerns, but his data were retained from stage 1 as it was anonymised. Nine declined to continue to participate but gave consent for time 1 data to be used leaving a time 2 sample of 42. To increase the sample size for time 1 data, further DSQ-40 data (N = 107) were collated from existing routinely collected assessments that were completed with patients at the study site on entry to a clinical service. All reported procedures were reviewed and approved by the West London Mental Health Research and Development Consortium and a National Health Service (NHS) Research Ethics Committee.

The Defense Style Questionnaire 40-item (DSQ-40; Andrews, Singh & Bond, 1993).

The DSQ-40 assesses 20 different defenses (see Table 1) that are categorised into one of three factors: mature (4 defenses; 8 items), immature (12 defenses; 24 items), or neurotic (4 defenses; 8 items). Respondents rate individual agreement on 40 statements that correspond to the defense styles, agreement is rated on a 9 point Likert scale, with labels of agreement confined to each end of the scale, 1 = “strongly disagree” and 9 = “strongly agree”, with no descriptor in between. A score for each of the 20 different defenses is calculated by averaging the two questionnaire items that correspond to the defense.
Statistical Analyses

To manage missing data, Little’s Missing Completely at Random (MCAR) test was conducted to determine the most appropriate method for missing data imputation. Nine item responses were missing from the data. Little’s MCAR test indicated the data were missing at random ($x^2 = 295.57$, df = 312, $p = .74$), and the Expectation Maximisation method was used to replace missing values. To control for potential bias from this particular method of imputation; factor, internal consistency and test re-test reliability analyses were conducted on data with and without missing values, and both sets of findings are reported for comparison purposes.

A confirmatory factor analysis, using the Maximum Likelihood extraction method, was conducted to determine the goodness of fit for the three-factor: mature, immature and neurotic, structure of the DSQ-40, as reported by Andrews et al.’s (1993). The analysis was conducted on the item correlation matrix. The rationale for this procedure was informed by recent evidence that has demonstrated the item pairs that represent 20 individual defenses of the three defense styles are not confirmed item parcels in terms of unidimensionality and could lead to biased estimates of parameters and model fit if factor analysed (Wilkinson & Ritchie, 2015). The factor model was tested using the lavaan package for latent variable modelling (Rosseel, 2012), which includes confirmatory factor analysis within the statistical programme R (R Development Core Team, 2011). Goodness of fit for the factor model was evaluated using a range of reported fit indices: chi-square ($x^2$/ df = $p > 0.5 / < 2 x$ df; Comparative Fit Index (CFI) $\geq 0.95$; Root Mean Square Error of Approximation (RMSEA) $\leq 0.05$; Standardised Root Mean Square Residual (SRMR) $\leq 0.05$. Indices were selected based on recommended criteria to reduce the risk of accepting a poor model fit (Hayduk, 1996; Hu & Bentler, 1999).
To investigate the internal consistency of the three defense styles, indices of reliability, Average Variance Extracted (AVE) and Composite Reliability (CR), were reported using confirmatory factor analysis item loadings. Fornell & Larcker (1981) cut-off criterion were applied: AVE, > 0.7 ‘very good’, 0.5 ‘acceptable’; CR, > 0.7 ‘acceptable’. For comparison and replication purposes Cronbach’s alpha coefficients were also calculated for the three defense styles using the Statistical Programme for the Social Sciences version 21 (SPSS v.21; IBM, 2011). Whilst a range of criterion to indicate an acceptable alpha level for scales have been proposed in the psychometric literature, limitations of setting criterion levels have been indicated (Clark & Watson, 1995; Kline, 2000). The current study therefore reports the specific alpha levels for the three factors and discusses these in the context of proposed criterion and alongside existing research findings for the DSQ-40’s reliability. The same procedure was proposed for an exploratory factor analysis.

Structural equation modelling using the Analysis of Moment Structures (AMOS; IBM, 2011) was used to assess the test-retest reliability of the three factors over an eight-week period. No formal criterion of acceptability was set, but the current findings are discussed in the context of existing reliability findings and proposed criterion.

In the event that goodness of fit indices criteria for the confirmatory factor analysis were not met, an exploratory factor analysis was proposed. Based on existing factor structure investigations and discussions of the DSQ-40 and the theoretical position of the proposed three factors, an item level principal components analysis using varimax rotation was planned a priori (Wilkinson & Ritchie, 2015).
Findings

Means and standard deviations for the 20 different reported defenses for 160 forensic patients are reported in table 3. When compared to DSQ-40 normative data from a community sample the reported defenses were broadly comparable (Andrews et al., 1993). Of note, participants reported a higher level of agreement with statements that represented the defense of projection.

Insert table 3 about here

Confirmatory Factor Analysis of the DSQ-40

Goodness of fit indices reported for the confirmatory factor analysis were as follows: $\chi^2/df = p < 0.5$, df = 736; CFI = 0.52; RMSEA = 0.08 (90% CI 0.07-0.08); SRMR = 0.10. Findings were comparable when retaining missing data ($\chi^2/df = p < 0.5$, df = 698; CFI = 0.59; RMSEA = 0.07; SRMR = 0.09).

Reliability analyses of the DSQ-40

The AVE, CR and Cronbach’s alpha reliability indices for the three factors of the DSQ-40 are reported in table 4. The AVE for all three factors was below the ‘acceptable’ criterion, ranging from 0.18 to 0.21. One of the CR indices, for the immature factor, met ‘acceptable’ levels; the remaining factors did not. The immature factor had the highest reported alpha coefficient at .85. A minor improvement to the scale was indicated by the removal of items 4 (“I am able to find good reasons for everything I do”) and 16 (“there are always good reasons when things don't work out for me”), both of which represent rationalisation. The mature factor of the DSQ had a lower reported alpha of .65, with no indication of improved internal consistency with
the removal of any items. The neurotic factor had the lowest reported alpha level of .53, with a marked increase in internal consistency (α=.62) with the removal of item 7 (“if someone mugged me and stole my money I’d rather he be helped than punished”). The reported coefficients when missing data was retained were broadly consistent (mature: α=.65; neurotic: α=.53; immature: α=.85), and the recommended items to remove were identical.

**Insert table 4 about here**

Test re-test reliability coefficients are reported in figures 1-3. The immature factor yielded the highest correlation over time (r = .91), and a small mean change (.02) with goodness of fit indices as follows: $x^2 (273, 133) = 417.65$, $p < .001$, RMSEA = .06, CFI = .81. The mature factor had a lower reported correlation (r = .76), with a small mean change (-.15), and goodness of fit indices as: $x^2 (26, 133) = 30.95$, $p = .23$, RMSEA = .04, CFI = .95. The correlation for the neurotic factor was .70, with a small mean change (.01) and goodness of fit indices as: $x^2 (25, 133) = 32.59$, $p = .14$, RMSEA = .05, CFI = .91. As with the previous analyses, findings were consistent when missing data were retained.

**Insert figures 1-3 about here**

**Exploratory Factor Analysis**

Principal components analysis of the 40 DSQ items initially identified five DSQ items with ‘low’ (<0.3) inter-item correlations (items, 1, 7, 16, 24, 35). Factor solutions were investigated with these items included and excluded to establish the
most parsimonious factor structure. From the preliminary analysis, factor extraction based on eigenvalue criterion (>1; Kaiser, 1960) indicated an 11-factor solution (10 with low inter item correlations excluded). The number of components explaining the percentage of variance greater than 5% was three (including and excluding low inter item correlations). Scree plots for 40 item and 35 item components indicated six and seven factor solutions, respectively. Three, six and seven factor solutions were therefore examined for interpretability. The most parsimonious solution, assuming the existing DSQ concepts was a three-factor solution with revised item loadings (see table 5). Of note, item 24 did not load on any of the three factors and is therefore excluded from the solution. Whilst the six and seven factor solutions explained greater variance (44 and 48% variance respectively), the conceptual coherence and interpretability of the factors was not evident (see Appendix A).

Insert table 5 about here

Discussion

Limited research has examined forensic patients’ psychological defenses despite the advantages that understanding such defences may hold for assessment and treatment purposes. More fundamentally, there have been no psychometric evaluations of the use of defense style assessments with this population. The current study presents the first psychometric evaluation of the self-report DSQ-40 for an adult (>18 years) male forensic inpatient population. On initial inspection, forensic inpatients rated agreement with a wide range of defense styles. This included adaptive defenses to help regulate negative feelings (e.g., use of humour and suppression), and
defenses that may be effective in the short term, but more problematic over time (e.g., projection and denial). Of note, the sample did not report many marked differences in the use of defenses when compared to a community sample, despite existing evidence of associations between maladaptive defense styles and mental disorders (e.g., Bond, 1986; Watson, 2002). One explanation for the finding could correspond to existing reviews of the DSQ that suggest a poor level of discrimination between clinical and non-clinical populations (Thygesen et al., 2008). Another explanation for the finding, which has been demonstrated in other studies comparing assessment ratings between forensic inpatients and community samples using the DSQ-40 (Andrews et al., 1993; Thygesen et al., 2008), is that the former group will be conscious about responses to assessments and how findings may impact on their detention or treatment (Rogers & Bender, 2003). A clinical implication to inform the level of confidence in self-report would be to consider triangulating additional assessments of impression management when assessing defense styles, or consider observer rated defense assessments (e.g., Defense Mechanisms Rating Sale, DMRS; Perry, 1990). Age could also account for the apparent lack of distinction; given the participants in the current study were, on average, older than participants in the Andrews et al. (1993) evaluation of the DSQ-40, and defenses have been proposed to change with age, in particular mature defenses increasing, and immature and neurotic defenses decreasing over time (Cramer, 1991). A further explanation is that respondents have a biased perspective of their own functioning and (either consciously or unconsciously) over or under emphasise individual abilities or behaviours; which has been evidenced in comparable settings (e.g. Milton et al., 2005). Finally, participants were predominantly on assertive rehabilitation wards at the study site where there is an expectation for patients to actively participate in their treatment to improve mental health and reduce
risk using the “Recovery” approach, and where the focus is on preparing them for a lesser secure environment, the anticipated next ‘step’ in their care pathway. This might suggest, albeit tentatively, that these individuals were more stable in mental state and functioning, and able to manage anxieties without relying on or resorting to unhealthy defenses.

The confirmatory factor analysis of the three-factor structure of the DSQ-40 was not empirically supported, based on the reported goodness of fit indices. This finding would therefore suggest caution when applying the three factors of the DSQ-40 to forensic inpatients. This is further emphasised with the reported reliability indices of the mature and neurotic factors which would not meet criterion recommended for ‘adequate’ internal consistency (e.g., Fornell & Larcker, 1981; Nunnally, 1978). This finding was consistent with reliability indices (Cronbach’s alpha) reported for these two factors in non-forensic inpatient samples (Andrews et al., 1993; Lopez & Gormley, 2002; Watson & Sinha, 1998). No improvements from item changes were indicated for the mature factor, and the removal of only one item, representing reaction formation, from the neurotic factor, would only marginally improve the scale. Item revision has been recommended in existing evaluations of the DSQ-40 (i.e. Watson & Sinha, 1998), but based on the current findings, this would not markedly improve the internal consistency of the assessment. The immature factor yielded improved reliability indices (with the exception of AVE), which is also consistent with previously referenced studies (e.g., Andrews et al., 1993). A possible artefact of this difference to the other factors is the higher number of items (n = 24), which may inflate internal consistency (Tavakol & Dennick, 2011). Items that could be revised to further improve its reliability, again albeit marginally, both represented
the rationalisation defense. One explanation offered is that the wording of these items is insufficient to support the construct validity of this defense. There are no negative connotations to either statement, yet conceptual references to rationalisation are typically in the context of being at fault or feeling regret for one’s actions (Vaillant, 1986). The test re-test reliability for the three factors over an eight-week period ranged from .70 to .91, which would suggest the factors are stable when assessed over time, albeit of variable internal consistency. Andrews et al., (1993) reported similar coefficients over a four-week period (n = 89; Mature = .75; Neurotic = .78; Immature = .85).

A further exploratory factor analysis of the study data did indicate a possible three-factor solution that predominantly corresponded to the three defense styles proposed by Andrews et al., (1993), but with eight (two representing neurotic defenses, and six immature) of the 40 items loading onto different factors. When examining item content there is conceptual scope for immature defenses to represent nervous (neurotic) psychological processes and the reverse is also applicable (see table 5). As has been previously discussed, items 16 appears to be representing a mature, rather than an immature defense response. If taking a parsimonious approach to the factor solutions, the revisions to the three defenses in terms of item content are arguably closest to replicating the original DSQ-40 validation study. However, continued development and testing of theoretical models of defense mechanisms is advised, as has been in view of other psychometric reviews of the DSQ-40, to be confident in its reliability and validity (Wilkinson & Ritchie, 2015). This should include investigations of item face validity, particularly if assuming defenses lie outside of a person’s awareness; as these have been limited to date (Chabrol, et al., 2005).
A number of limitations of the existing study and future research directions warrant discussion alongside these findings. The study sample size would not qualify as meeting the ‘typical’ criterion for conducting a confirmatory or exploratory factor analysis (Kline, 2013). However, this minimum sample size rule of thumb may be invalid with other, more robust, criteria such as goodness of fit (for the former) being advised (MacCallum, Widaman, Preacher & Hong, 2001). Further same sample size limitations apply to the internal consistency and test re-test findings, meaning further evaluations of the DSQ-40 in this type of settings are required. Finally, the use of the DSQ-40 as opposed to a version of the DSQ that includes a ‘lie-scale’ (e.g. DSQ-88; Bond, 1986) may also have been a limitation in terms of accounting for any impression management.

Conclusions

Based on the reported properties of the DSQ-40 certain recommendations for its continued use with male forensic inpatients are put forward. Use of the original three (higher-order) factor structure may be problematic, given the lack of fit for items to factors, and a more specific focus on the individual defenses or individual item responses may be more meaningful in the context of both case formulation and outcome evaluation. If the factors are applied, then certain revisions to items may be advisable, to at a minimum, improve their coherence. It may also be advisable to either triangulate clinician rated assessments of defenses or include assessments of impression management, to increase confidence in self-reported defense styles.
References


Retrieved from: http://www.ijpsy.com


