A Psychometric Validation Analysis of Eysenck’s Neuroticism and Extraversion Scales in a Sample of First Time Depressed Patients

Stine Bjerrum Moeller*, Per Bech, Lars Kessing, Erik Lykke Mortensen, Stephen F. Austin and Jens Otto Drachmann Bukh

Mental Health Centre North of Zealand, Hilleroed, Denmark

Abstract

Eysenck and Eysenck identified the two-factor structure of personality, namely neuroticism and extraversion which has been widely used in clinical psychiatry, and generated much research on the psychometric properties of the scales. Using a classical psychometric approach the neuroticism and extraversion scales have shown robust psychometric properties. The present study used both classical psychometric and item response theory (IRT) analyses to evaluate the neuroticism and extraversion scales and improve scalability of the instrument. A first time depressed sample completed the Eysenck Personality Questionnaire (EPQ) and Hamilton Depression Rating Scale (Ham-D 17-item) at baseline and at follow-up five years later.

For the neuroticism scale a subscale containing depression-related symptoms, a subscale containing anxiety-related symptoms, and a subscale containing symptoms related to interpersonal sensitivity were identified. For the extraversion scale a shorter and psychometrically more robust version was identified together with a short introversion scale. Clinically discriminant validity was analysed using correlations. The correlation between depression (Ham-D-17) and neuroticism-anxiety was below clinical significance, while the correlation of depression (Ham-D-17) was above clinical significance for both interpersonal sensitivity and neuroticism-depression. Clinically discriminant validity was modest for the extraversion/ introversion subscales however the introversion subscale approached clinical significant correlation with depression. The identified subscales of the EPQ, neuroticism-anxiety, interpersonal sensitivity, neuroticism-depression, extraversion and introversion are psychometrically valid measures of type and severity of the stress response. Using these subscales, it is possible to perform a quick and psychometrically valid evaluation of the type and severity of the stress response. The subscales may be useful in predicting types of psychopathology and in identifying underlying vulnerability that could serve as specific treatment targets in preventing as well as treating depression and anxiety.

Keywords: Eysenck personality questionnaire; First time depressed patients; Personality; Traits; Psychometric validation; Uni-dimensional measures; Mokken analysis

Introduction

Eysenck and Eysenck identified the two-factor structure of personality, namely neuroticism and extraversion [1], and later added a third, namely psychoticism [2]. The development of Eysenck’s two factor model was based on Wundt’s attempt [3] to converge the melancholic and phlegmatic temperaments into the internalizing personality factor of neuroticism, and the sanguine and choleric temperaments into the externalizing personality factor of extraversion. The Eysenck Personality Questionnaire (EPQ) [4] has been widely used in clinical psychiatry, and generated much research on the psychometric properties of the scales [5]. Especially the neuroticism and extraversion scales have shown robust psychometric properties using a classical psychometric approach [6,7].

The present analysis focuses on the psychometric validation procedure using both a classical psychometric model and an item response theory (IRT) model to evaluate the neuroticism and extraversion scales with the purpose of providing brief and uni-dimensional measures of neuroticism and extraversion. Brief measures save time in a busy clinical practice. The advantage of uni-dimensional measures is a more unambiguous association with other constructs and improved scalability of the instrument [5]. The clinical implications are potentially more accurate measurement of targets for treatment, resulting in better treatment outcomes.

Neuroticism refers to a tendency to experience negative affect. When Eysenck conceptualised neuroticism it included several components [3] which is why it has been identified as a common non-specific vulnerability factor in psychopathology [8-10]. Depression and anxiety are highly co-morbid disorders [11] and recent transdiagnostic research has focused on examining to what extent transdiagnostic concepts and processes are operating across disorders [12]. Neuroticism is one such transdiagnostic concept representing the underlying trait that predisposes the individual to react with all kinds of distress when a stressful life event occurs. One way to interpret results from large population studies on neuroticism and the response to life stressors is that the stress response seems to manifests itself on a continuum. Covering the mild range of distress are symptoms related to anxiety, covering the moderate range of distress are symptoms related to interpersonal sensitivity, and covering the severe range of distress are symptoms related to depression [13-15].

Extraversion is also conceptualised as reflecting varied features, such as sociability, liveliness, energy and positive emotionalism, while introversion is the opposite, referring to a tendency to withdraw socially. The extraversion/introversion construct has repeatedly

*Corresponding author: Stine Bjerrum Moeller, Psychologist, Phd., Mental Health Centre North of Zealand, Hilleroed, Denmark, Tel: 45 38 64 30 62; E-mail: Stine.Bjerrum.Moeller@regionh.dk

Received August 19, 2015; Accepted September 23, 2015; Published September 26, 2015


Copyright: © 2015 Moeller SB, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
shown an association with depression, with low extraversion and high introversion in the depressed state, but also in the euthymic phase [16]. According to Eysenck [3] introversion is a personality trait associated with dysthymia. An important distinction here is between social withdrawal due to neurotic feelings of inadequacy, worry and fear of failure in social situations, and social withdrawal due to true introversion because the person prefers to be alone. The first is a more anxious response, while the latter is the true introverted trait.

In a previous study of chronic idiopathic non-malignant pain disorder patients, we found psychometric support for a 23 item principal component of the neuroticism scale of the EPQ with a subscale covering anxiety and a subscale covering depression. We also found support for a shortened extraversion scale in EPQ [17]. This previous study indicated that the full neuroticism and extraversion scales have more items than required and that subscales can be successfully identified, potentially improving psychometric qualities and clinical utility. The present study further tests the validity of the EPQ by repeating the validation of these uni-dimensional subscales of neuroticism and extraversion in a different sample of first-time depressed patients.

Using depressed patient for this psychometric validity study as opposed to a general population is essential because the clinical utility is on measuring relevant personality traits among patients suffering from an affective disorder. As such, the validity must be tested in such a sample. We used first-time depressed patients to avoid the impact of earlier affective episodes on personality traits, and we assessed the patients at a relatively neutral point in the depressive episode to limit the influence of depressive symptoms.

Specifically, we aimed to answer the following research questions:

1. Are the subscales: neuroticism-anxiety, neuroticism-depression, neuroticism-interpersonal sensitivity, extraversion and introversion psychometrically uni-dimensional?
2. Is the hierarchical nature of the neuroticism construct supported by the Hamilton Depression Scale (HAM-D-17), showing the strongest positive correlation with neuroticism-depression, the second strongest correlation with neuroticism-interpersonal sensitivity, and the weakest positive correlation with neuroticism-anxiety?
3. Is extraversion negatively correlated with depression while introversion is positively correlated with depression (Ham-D-17)?

**Method**

**The sample**

The present study is part of a larger study of gene-environment interaction in first episode depression and predictors of the course of affective disorders. We have previously described the baseline study in details [18]. A total of 399 Danish patients aged 18-70 years with diagnosis of a single depressive episode (ICD-10) [19] that had been recently discharged from their first ever admission or out-patient contact to a psychiatric hospital in eastern Denmark (Zealand) were sampled consecutively in a 2-year period from 2005-2007 via the Danish Psychiatric Central Research (DPCR) Register [20]. ICD-10 diagnoses were established for the episode leading to psychiatric hospital care and for any lifetime episodes by the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) [21]. Patients diagnosed with non-affective psychiatric disorders (N=39), bipolar disorder (N=13), or recurrent depression (N=44) were excluded, leaving a total of 301 patients for further analyses. In order to evaluate the longitudinal course of illness, participants were re-assessed from 2011-2013. The mean follow-up time was 5.8 years (S.D. 0.3 years; range 5.2-7.7 years). As the present paper focuses on providing robust evidence for the psychometric properties of trait-based concepts, the stability of the measures had high priority. Therefore our analyses are based on a very conservative approach, including only complete datasets at baseline as well as at follow-up, leaving a total of 150 patients to be included in the analyses of the present paper.

**Measures**

At baseline the participants completed the Eysenck Personality Questionnaire (EPQ) (Eysenck, Eysenck 1975), and the severity of depressive and anxiety symptoms was assessed with the 17-item Hamilton Depression Rating Scale (HAM-D-17). At the time of follow-up, the participants completed the EPQ again, and current depressive symptoms were evaluated by HAM-D 17-item scale.

**Psychometric validation analysis**

In an attempt to define all useful common variance in the EPQ neuroticism and extraversion subscales, a principal component analysis (PCA) was performed. It was expected that the first principal component would be a general factor because all the items were selected as having positive inter-correlation. The second or the third components were expected to be bi-directional. This part of the analysis tested the extent to which the anxiety items and the depression items were separated in the PCA by different factor loadings, i.e., negative versus positive loadings. If this was found to be the case, no rotation procedure would be necessary [4,6,22]. To test the continnum hypothesis of the the stress response, we identified the moderate range by selecting 5 clinically relevant items covering interpersonal sensitivity among the depression items of the neuroticism subscale.

A PCA on the extraversion scale was expected to show loadings with different signs for the three introversion items from those of the extraversion items within the first principal component. Cronbach’s coefficient alpha [23] was used to test the hypothesis that the full extraversion scale contains too many items that are merely variants of one single aspect of the dimension of extraversion. The 7 items with the highest loadings in the first component were expected to prove psychometrically sound in covering the entire range of the extraversion construct.

The non-parametric Mokken analysis was used to evaluate the measurement aspects of the EPQ subscales for scalability. The Mokken model [24] is based on the Guttman cumulative rating scale principle [5,25]: that scorings on lower prevalence manifest items must be preceded by scorings on high prevalence items. The scalability is evaluated by use of Loewinger’s coefficient of homogeneity [5] but the rank order of prevalence is evaluated by the mean scores of the manifest items. The items with higher mean scores are considered as more prevalent. Conducting a Mokken analysis [24,26], we tested the hypothesis that the subscales: neuroticism-anxiety, neuroticism-depression, neuroticism interpersonal sensitivity, extraversion, and introversion are uni-dimensional, and that the total summed item subscore is a sufficient statistic. According to Mokken [24] or van Schuur [27] coefficients of homogeneity from 0.20 to 0.29 belong in a questionable zone as to constituting a cumulative scale. Coefficients of homogeneity from 0.30 to 0.39 are just acceptable for scalability while a coefficient of homogeneity of 0.40 or higher clearly indicates scalability, i.e. that the items are additive and their summed total score...
a sufficient statistic. The Mokken analysis was performed using the MSP program [26].

Finally, the clinical validity was tested using correlational statistics. As a sign of clinical significance, we used a correlation coefficient of 0.40 corresponding to the recommendations by Mokken [24].

Results

Out of the 301 participants that were included in the gene-environment interaction study, we had missing item scores at either baseline or follow-up for 112 of the participants (42.7%), resulting in the inclusion of 150 patients in the present analyses. Comparing the included 150 participants for the present analyses with the 112 participants with missing items revealed significantly more male participants in the group with complete answers to EPQ at baseline and follow-up (36.0 % vs. 22.3 %, p=0.02). No differences in mean age at baseline (40.9 years vs. 38.5 years, p=0.2), severity of depressive symptoms (mean HAM-17 scores) at baseline (9.3 vs. 9.3, p=0.9) or at follow-up (5.5 vs. 5.7, p=0.8) was found between the two groups.

The principal components analysis of the EPQ neuroticism dimension identified five components with an eigenvalue above 1 at baseline as well as follow-up. At baseline, the first principal component had an eigenvalue of 8.26, and the second an eigenvalue of 1.79. Together these two factors accounted for 44% of the variance. At follow-up, similar results emerged with an eigenvalue of 7.97 for the first principal component, and 1.86 for the second component. Together these two factors accounted for 43% of the variance. All loadings were positive, indicating a general factor, and were comparable from baseline to follow-up, indicating stability of the construct.

The second component was clinically meaningful. It revealed itself as a bi-directional component, with items with negative loadings mainly covering the depression aspect of neuroticism, and items with positive loadings covering the anxiety aspect of neuroticism. Overall the results for baseline and follow-up were comparable with a few exceptions: items 72 ("Do you worry too long after an embarrassing experience?") and 88 ("Are you touchy about some things?") did not load negatively as expected at base line and item 66 ("Do you worry a lot about your looks?") did not load negatively as expected at follow-up. The lacking negative loadings at baseline for items 72, 80, 88 were not important as these items were selected to be included in the interpersonal sensitivity subscale of neuroticism. Because item 66 did not load negatively at follow up it was not selected to be included in the depression subscale of neuroticism.

Regarding the positive loadings of the anxiety subscale, item 54 ("Do you suffer from sleeplessness?") unexpectedly did not load positively at follow up, and was thus not included in the anxiety subscale. The loadings are displayed in Table 1.

The principal component analyses of the EPQ extraversion dimension identified five components with an eigenvalue above 1 at baseline as well as follow-up. At baseline, the first principal component had an eigenvalue of 7.10, accounting for 34% of the variance. At follow-up, similar results emerged with an eigenvalue of 6.93, accounting for

<table>
<thead>
<tr>
<th>No</th>
<th>Affiliation</th>
<th>Symptoms</th>
<th>Positive loadings baseline</th>
<th>Negative loadings baseline</th>
<th>Positive loadings follow-up</th>
<th>Negative loadings follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>A</td>
<td>Would you call yourself a nervou person?</td>
<td>0.56</td>
<td>0.56</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>41</td>
<td>A</td>
<td>Would you call yourself tense or &quot;highly-strung&quot;?</td>
<td>0.51</td>
<td>0.51</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>75</td>
<td>A</td>
<td>Do you suffer from &quot;nerves&quot;?</td>
<td>0.51</td>
<td>0.51</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>34</td>
<td>A</td>
<td>Are you a worrier?</td>
<td>0.38</td>
<td>0.38</td>
<td>0.52</td>
<td>0.52</td>
</tr>
<tr>
<td>47</td>
<td>A</td>
<td>Do you worry about your health?</td>
<td>0.28</td>
<td>0.28</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>54</td>
<td>A</td>
<td>Do you suffer from sleeplessness?</td>
<td>0.24</td>
<td>0.24</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>38</td>
<td>A</td>
<td>Do you worry about awful things that might happen?</td>
<td>0.12</td>
<td>0.12</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>80</td>
<td>IS</td>
<td>Are you easily hurt when people find fault with you or the work you do?</td>
<td>0.08</td>
<td>0.08</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>72</td>
<td>IS</td>
<td>Do you worry too long after an embarrassing experience?</td>
<td>0.05</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>88</td>
<td>IS</td>
<td>Are you touchy about some things?</td>
<td>0.03</td>
<td>0.03</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. Letter signifies the affiliation to depression (D) or interpersonal sensitivity (IS) subscale. Note. Items are ordered with the highest loadings at baseline ranked first. Letter (A) signifies the affiliation to the anxiety subscale.

Table 1: Loadings of the second principal component of the EPQ neuroticism at baseline and follow-up.
33% of the variance. The second and the third components were not clinically meaningful. Within the first component the introversion items showed the expected negative loadings. Results for baseline and follow-up were comparable. The 7 extraversion items with the highest loadings at baseline within the first component were selected to cover the entire dimension of extraversion (see appendix for subscale items). The selected items overlapped by and large with the items from Bech et al. [17] that were considered clinically valid in covering the entire dimension of extraversion. However, with three items (item 32 “Do you have many friends?”; 60 “Do you like doing things in which you have to act quickly?”; and 82 “Do you like plenty of bustle and excitement around you?”) loading below those selected in the present sample, the subscale was not entirely the same as in the previous study. Testing the reliability and the uni-dimensionality, Cronbach’s alphas and Loevinger’s coefficient of homogeneity of the neuroticism scale (N23), neuroticism-depression (N10), neuroticism-anxiety (N6), neuroticism interpersonal sensitivity (N5), extraversion (E21), and extraversion (E7) and introversion (I3) were computed. Results from the analyses revealed values going from acceptable to excellent and signifying uni-dimensionality. In particular, shortening the extraversion scale resulted in a much better psychometric validity. The results are available in Table 2.

Finally, the clinical validity of these uni-dimensional subscales of the neuroticism construct was tested using correlational statistics. We used the follow up data set because at follow up the influence of the depressive episode would be smallest and because the variance of the score distribution would be largest here, which is optimal for test of correlations [28]. Thus the variance of the distribution would be largest because some participants would be completely remitted from depression, others would still have residual symptoms, and some would still be above cut-off point for clinical depression. Neuroticism-depression (N10) showed the expected strongest positive correlation with depression measured with the HAM-D-17, while neuroticism-interpersonal sensitivity (N5) showed the expected second strongest, and neuroticism-anxiety (N6) the weakest positive correlation with depression. The correlation of depression (HAM-D-17) and neuroticism-anxiety (N6) was below clinical significance, while the correlation of depression (HAM-D-17) was above clinical significance for both interpersonal sensitivity (N5) and neuroticism-depression (N10). These results support the validity of the subscales. Results are displayed in Table 3. As for extraversion and introversion, the correlations were in the expected direction, however without reaching clinical significance. Introverson (I3) approached a clinically significant correlation with depression (Ham-D-17) (r=0.37).

### Discussion and Conclusion

Psychometrically uni-dimensional subscales of the EPQ neuroticism and extraversion

In the present study on first-time depressed patients, we have demonstrated that Eysenck’s neuroticism scale covers a broad construct that contains identifiable subscales. We identified a ten-item subscale with depression-related symptoms (N10), a six-item subscale containing anxiety-related symptoms (N6), and a five-item subscale containing symptoms related to interpersonal sensitivity (N5). The results provide strong evidence that the subscales have stronger psychometric properties than the full scale and are uni-dimensional. The 5 items selected from the neuroticism scale to reflect interpersonal sensitivity showed the expected moderate correlation to depression indicating that interpersonal sensitivity covers the middle range of the stress response. This, however, should be interpreted with caution, and further investigation is needed to draw a more definite conclusion.

The exclusion of item 54 “Do you suffer from sleeplessness?” in the anxiety subscale was necessary as sleeplessness is not specific to anxiety. In addition, this item reflects a state of depression/anxiety more than a trait-like disposition to experience particular symptoms. It is important to differentiate between the underlying vulnerability trait and the active illness in order to avoid confounding.

We also demonstrated that the Eysenck extraversion scale shows a psychometrically better performance without losing clinical validity when shortened. The extraversion subscale showed increased homogeneity measured by the Loevinger coefficient of homogeneity when separated into 7 extraversion items and 3 introversion items. The Cronbach’s alpha for this scale was lower than those reported in other studies [29]. This may indicate principal problems in the conceptualization of the extraversion construct which reflects varied features, such as sociability, impulsivity, risk taking, energy and positive emotionality. The introversion scale, however, contains too few items.

The answer to the first research question is that the subscales: neuroticism-anxiety, neuroticism-depression, neuroticism-interpersonal sensitivity, extraversion, and introversion are psychometrically uni-dimensional.

### Hierarchical nature of the neuroticism construct

Because the subscale neuroticism-anxiety (N6) showed the weakest correlation with depression (r=0.37), neuroticism-interpersonal sensitivity (N5) an intermediate correlation to depression (r=0.44), and neuroticism-depression (N10) the strongest correlation to depression (r=0.51), the results supported the proposed hierarchical nature of neuroticism. These results should be interpreted very cautiously because the differences between the correlations are not significant (although approaching significance for the difference between the correlation for neuroticism-anxiety and depression (r=0.37) and for neuroticism-depression and depression (r=0.51) (z=1.5, p=0.068).
Further studies testing the predictive validity of these subscales are warranted. Regarding the clinical validity, while depression and anxiety relate closely or even overlap, it has also proven relevant to differentiate between the underlying constructs responsible for the different symptomatic expressions. One example of an underlying concept that has specificity when investigated in detail is repetitive negative thinking, including rumination and worry [30]. Neuroticism is a parallel construct to repetitive negative thinking and in the present study we identified distinct subscales within the neuroticism scale with a subscale related more to the anxious response to stress and another subscale related more to the depressive response to stress. As differences among the underlying constructs involved (e.g. neuroticism) in depression and anxiety are subtle, we advise to be careful about wording when constructing items. An example of this is item 12 ("Do you often worry about things you should not have done or said?") which includes the phrase “worry” in regard to thinking about an earlier situation. As worry is a future-oriented concept related to anxiety [31], we would advise to use the phrase “pondering about” or “ruminating” when referring to repetitive thinking about past events. Securing more precisely worded items would increase the discriminant validity of scales and subscales. The identification of a depression subscale and an anxiety subscale of the neuroticism scale may facilitate differentiating between vulnerability to anxiety or depression. Future studies are needed to investigate that potential relationship further.

Lastly, the stronger correlations at follow-up compared to baseline indicated difficulty in assessing personality traits close to an affective episode. This indicates at least some state dependence of the measurement of personality traits which the clinician should be aware of when conducting personality assessment of patients. It is therefore recommended to evaluate personality traits independently of a current affective episode.

In conclusion, the answer to the second research question is that the hierarchical nature of the neuroticism construct was supported.

The affiliation between the extraversion/introversion subscales and depression

Although only approaching clinical significance, the correlation between extraversion and depression showed the expected negative association, indicating that extraversion functions as a protective trait against depression [10]. The negative correlation with depression was modest at baseline and higher at follow-up where it was in the same range as in our previous study [17]. This could point to difficulties in measuring extraversion shortly after an affective episode. Introversion reflects the tendency to social withdrawal due to a preference to being alone. In response to stress this personality trait will increase the risk of withdrawal and avoidance behaviour that may lead to decreased positive reinforcement and the triggering of a depressive episode. Additionally, because social withdrawal due to neurotic feelings of inadequacy, worry and fear of failure in social situations is a prominent response in episodes of anxiety and depression, more research is needed to disentangle introversion as a trait from the symptomatic avoidance response of anxiety and depression.

The answer to the third research question is that extraversion is negatively correlated with depression while introversion is positively correlated with depression.

Study limitations and strengths

Limitations of the study are that evaluation of causal relations requires a longitudinal design beginning before the first episode of a distress disorder. Neuroticism and extraversion as vulnerability factors are not ideally studied after the first episode of depression, however including participants before their first depressive episode would require a very large sample. We also acknowledge that the sample cannot be generalized; we looked at a specific follow up period not representative of all people that achieve remission from depression. If a 2-year period had been looked at instead of 5 years the relationship between the variables might have been different. A strength of the present study is the longitudinal design in which first-time depressed patients were followed, as the influence of repeated affective episodes, the “scar effect”, is limited, while at the same time the population is of specific relevance for the constructs of interest (e.g depression and personality traits). However, as we did not consecutively follow up the participants in the period between baseline assessment and follow up assessment, participants may have varied in the degree to which they experienced depressive symptoms in the period between baseline and follow up.

Studies using these subscales based on the EPQ should be replicated and further validity studies in different samples are required. Ideally, the personality traits of the EPQ should be measured prior to any affective episode in a longitudinal design and participants should be consecutively monitored during the study period to control for the influence of affective symptoms.

In conclusion, the present study identified uni-dimensional subscales within the Eysenck’s neuroticism scale and extraversion scale that we believe have clinical importance. Using these subscales, it is possible to perform a psychometrically valid evaluation of the type and severity of the stress response. These subscales may be useful in predicting types of psychopathology and in identifying underlying vulnerability that could serve as specific treatment targets in preventing as well as treating depression and anxiety. For example, an individual presenting a profile reflecting the anxious component of neuroticism may benefit from a treatment focused specifically on anxiety to treat or prevent future affective episodes.

References


