

# The heritability of cluster A personality disorders assessed by both personal interview and questionnaire

KENNETH S. KENDLER<sup>1,2\*</sup>, JOHN MYERS<sup>1</sup>, SVENN TORGENSEN<sup>3</sup>,  
MICHAEL C. NEALE<sup>1,2</sup> AND TED REICHBORN-KJENNERUD<sup>4</sup>

*Departments of <sup>1</sup> Psychiatry and <sup>2</sup> Human Genetics, Medical College of Virginia of Virginia Commonwealth University, Richmond, VA, USA; <sup>3</sup> Institute of Psychology, University of Oslo, Center for Child and Adolescent Mental Health Eastern and Southern Norway and Nic Waal's Institute, Norway; <sup>4</sup> Division of Mental Health, Norwegian Institute of Public Health and the Institute of Psychiatry, University of Oslo, Norway*

## ABSTRACT

**Background.** Personality disorders (PDs) as assessed by questionnaires and personal interviews are heritable. However, we know neither how much unreliability of measurement impacts on heritability estimates nor whether the genetic and environmental risk factors assessed by these two methods are the same. We wish to know whether the same set of PD vulnerability factors are assessed by these two methods.

**Method.** A total of 3334 young adult twin pairs from the Norwegian Institute of Public Health Twin Panel (NIPHTP) completed a questionnaire containing 91 PD items. One to 6 years later, 1386 of these pairs were interviewed with the Structured Interview for DSM-IV Personality (SIDP-IV). Self-report items predicting interview results were selected by regression. Measurement models were fitted using Mx.

**Results.** In the best-fit models, the latent liabilities to paranoid personality disorder (PPD), schizoid personality disorder (SPD) and schizotypal personality disorder (STPD) were all highly heritable with no evidence of shared environmental effects. For PPD and STPD, only unique environmental effects were specific to the interview measure whereas both environmental and genetic effects were found to be specific to the questionnaire assessment. For SPD, the best-fit model contained genetic and environmental effects specific to both forms of assessment.

**Conclusions.** The latent liabilities to the cluster A PDs are highly heritable but are assessed by current methods with only moderate reliability. The personal interviews assessed the genetic risk for the latent trait with excellent specificity for PPD and STPD and good specificity for SPD. However, for all three PDs, the questionnaires were less specific, also indexing an independent set of genetic risk factors.

## INTRODUCTION

Two different approaches have been developed towards the measurement of personality disorders (PDs): structured interviews (SIs) (e.g. Loranger, 1988; First *et al.* 1995; Pfohl *et al.* 1995) and self-report questionnaires (SRQs)

(e.g. Livesley *et al.* 1992; Clark, 1993). To date, all twin studies of PDs have used one these assessment methods (Livesley *et al.* 1993; Jang *et al.* 1996; Torgersen *et al.* 2000; Kendler *et al.* 2006; Reichborn-Kjennerud *et al.* 2006). A genetically informative study that included the symptoms of PDs as assessed by both approaches could clarify the degree to which these two assessment methods index the same genetic and environmental risk factors. That is, such a study could answer the question: 'Do SIs and

\* Address for correspondence: Kenneth S. Kendler, M.D., Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University Medical School, Box 980126, 800 E. Leigh Street, Room 1-123, Richmond, VA 23298-0126, USA.  
(Email: kendler@vcu.edu)

SRQs assess the same or different PD-related vulnerability factors?'

In twin or adoption studies, the calculation of heritability from a single assessment confounds the roles of individual-specific environment and measurement error. For example, differences in the scores of a pair of monozygotic (MZ) twins on a particular PD could result from differences in their prior environmental experiences or from assessment error. Error in the assessment of PDs is a concern because PD symptoms might not be as temporally stable as once thought (McGlashan *et al.* 2005) and agreement on the level of PD symptomatology across assessment methods is typically modest (Zimmerman, 1994; Widiger & Coker, 2002). Furthermore, the heritabilities obtained recently from a population-based sample for cluster A (Kendler *et al.* 2006) and cluster C PDs (Reichborn-Kjennerud *et al.* 2006) assessed by SI were lower than those typically seen for normative personality traits (e.g. Loehlin, 1992; Lake *et al.* 2000). Such low heritabilities could arise either because PDs as assessed by SIs are reliably measured and have modest heritabilities or because they have heritabilities similar to standard personality traits but are assessed with lower reliability.

In the Norwegian Institute of Public Health Twin Panel (NIPHTP), interviews including the Structured Interview for DSM-IV Personality (SIDP-IV; Pfohl *et al.* 1995) were conducted between 1999 and 2004. The results of these interviews form the basis of our prior twin analyses of PDs in this sample (Kendler *et al.* 2006; Reichborn-Kjennerud *et al.* 2006). However, in 1998, the NIPHTP completed the Dysfunctional Personality Questionnaire, a 91-item SRQ that assessed dimensional representations of a wide range of PDs.

In this report, we incorporate into twin models assessments by both SI and SRQ of dimensional representations of the three cluster A PDs: paranoid personality disorder (PPD), schizoid personality disorder (SPD) and schizotypal personality disorder (STPD). With these models, we addressed two questions. First, accounting for unreliability of measurement by using two measures differing in both time and mode of assessment, how heritable is the liability to PPD, SPD and STPD? Second, what is the relative sensitivity and specificity with which

SI and SRQ assessments index the underlying liability to these individual PDs?

## METHOD

### Sample and assessment methods

Twins in the NIPHTP (described in detail elsewhere; Harris *et al.* 2002) were identified through the Norwegian National Medical Birth Registry, established 1 January 1967, which receives mandatory notification of all live births. The current panel began with 15 370 like- and unlike-sexed twins born 1967–1979. Two questionnaire studies have been conducted in 1992 (twins born 1967–1974) and in 1998 (twins born 1967–1979). Altogether, 12 700 twins received the second questionnaire, and 8045 responded after one reminder (response rate 63%). The sample included 3334 pairs and 1377 single responders. The second questionnaire contained the Dysfunctional Personality Questionnaire, which contained 91 items selected by one of us (S.T.) to assess DSM PD traits. These items were either selected from one of three prior established instruments (Foulds, 1965; Lazare *et al.* 1966; Conte *et al.* 1980) or developed and subsequently validated by S.T. (Torgersen, 1980 *a, b*). The Dysfunctional Personality Questionnaire was also administered in a prior large-scale Norwegian epidemiological study of PDs (Torgersen *et al.* 2001).

In the personal interview phase of this study, twins were not approached for interview until preliminary consent had been obtained from both members of the pair. Participants were recruited among 3153 complete pairs who, in the second questionnaire, agreed to participate in the interview study, and 68 pairs who were drawn directly from the NIPHTP. Of these 3221 eligible pairs, 0.8% were unwilling or unable to participate, and in 16.2% of pairs only one twin agreed to the interview. After two contacts requesting participation, 38.2% did not respond. Where only one twin agreed to the interview, the uncooperative twin either did not respond to our contacts (96.0%), had an unknown address (2.9%) or refused (1.1%). Altogether, 2794 twins (44% of those eligible) were interviewed for the assessment of PDs. Approval was received from the Norwegian Data Inspectorate and the Regional Ethical Committee, and written informed consent was obtained from all

participants after complete description of the study.

PDs were assessed by a Norwegian version of the SIDP-IV (Pfohl *et al.* 1995). Both DSM-III-R and DSM-IV versions have been used previously in major Norwegian studies (Torgersen *et al.* 2001; Helgeland *et al.* 2005). The SIDP is a comprehensive semi-structured diagnostic interview for the assessment of all DSM-IV axis II diagnoses. The instrument includes non-pejorative questions organized into topical sections to produce a natural flow in the interview.

The SIDP uses the '5-year rule', meaning that behaviors, cognitions and feelings that have been predominant for most of the past 5 years are considered to be representative of the individual's long-term personality functioning. This rule is supported by empirical evidence of high stability of normative personality traits during adulthood (McCrae & Costa, 1990). Each DSM-IV criterion is scored as 0 = absent, 1 = subthreshold, 2 = present or 3 = strongly present.

Interviewers (mostly experienced psychology students and psychiatric nurses) were trained by one psychiatrist and two psychologists with previous extensive experience with the instrument. The interviews, largely conducted face-to-face, were carried out between June 1999 and May 2004. For practical reasons, 231 interviews (8.3%) were conducted over the telephone. Members of a pair were assessed by different interviewers.

Inter-rater reliability was assessed by two raters scoring 70 audiotaped interviews. They obtained high intra-class (and polychoric) correlations for the number of endorsed criteria at the subthreshold level: PPD +0.92 (+0.94), SPD +0.81 (+0.86) and STPD +0.86 (+0.90).

Zygoty was determined by standard questionnaire items used in a discriminant analysis with results of 24 microsatellite markers available on 676 of the like-sex pairs in the sample. From these data, we estimated that, in our entire sample, zygoty misclassification rates are under 1% (Neale, 2003).

### Regression analyses

Subjects missing scores on 10% or more of their SRQ items (210 out of 8030, or 2.6%) were excluded from the analysis, producing a total sample of 7820. For those missing less than 10% (1782 twins, or 22.8% of the sample, of

whom 1001 were missing a single item and 343 two items), the missing values were imputed using IVEware (Raghunathan *et al.* 2000). Separately for each of the three cluster A PDs, we then conducted step-wise ordinal logistic regression analyses using PROC LOGISTIC in SAS (SAS Institute, 2006) in twin 1 from each twin pair ( $n = 1363$ ), attempting to predict, from responses to the Dysfunctional Personality Questionnaire, the number of criteria endorsed with a score of 1 or higher. The significance level for entry and exit into this regression analysis was 0.20. We then took these resulting items and repeated the analyses in the second twin in each pair ( $n = 1368$ ), this time using an entry and exit criteria of 0.05.

### Model fitting

We use a liability-threshold model to estimate the genetic and environmental contributions to twin resemblance for PDs. For SRQs, this liability is indexed by the number of items responded to in the positive direction. For the SI, liability is indexed by the number of DSM-IV criteria endorsed at the subthreshold level. In this paper, we refer, for convenience, to PDs but we are in fact assessing a dimensional representation of these PDs, operationalized as the number of endorsed criteria. We showed previously, using the multiple threshold model, that the number of criteria for the cluster A PDs could be regarded as differences of severity on a single normally distributed continuum of liability (Kendler *et al.* 2006).

The model-fitting used here divides the variation in liability to PDs into three classes: (i) additive genetic (A), which contributes twice as much to the correlation in MZ twins as dizygotic (DZ) twins, (ii) family or 'common' environment (C), which contributes equally to the correlation in MZ and DZ twins, and (iii) individual specific environment (E), which reflects environmental experiences not shared by both members of a twin pair and therefore contribute to differences between them in their liability to PDs.

Our model for PDs, previously referred to as a measurement model (Foley *et al.* 1998), uses simultaneously both our SRQ and SI data from our twin sample. As illustrated in Fig. 1, the model assumes that there is a true latent liability to each PD. The latent liability to the PD is

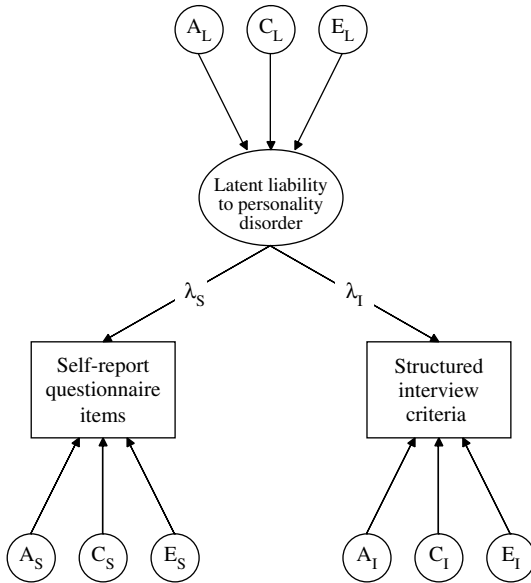


FIG. 1. The measurement model used in this report. The latent liability to personality disorder (PD) is indexed by both items from the self-report questionnaire (SRQ) and DSM-IV criteria as assessed the structured interview. The magnitude of this relationship is reflected in the paths  $\lambda_S$  and  $\lambda_I$  respectively, where S and I refer to self-report questionnaire and interviewed-based assessment. Genetic (A), shared environmental (C) and individual-specific environmental effects (E) are included in the model for the latent liability to PD (indicated by subscript L), specific to the self-report questionnaire (indicated by subscript S), and specific to the self-report structured interview (indicated by subscript I).

indexed by both items from the SRQ and DSM-IV criteria assessed by SI. The magnitude of this relationship is reflected in the paths  $\lambda_S$  and  $\lambda_I$ , where S and I refer to self-report and interviewed-based assessments respectively. Genetic (A), shared environmental (C) and individual-specific environmental effects (E) are included in the model for the latent liability to the PD (indicated by the subscript L), specific to the self-report questionnaire (indicated by the subscript S), and specific to the self-report structured interview (indicated by the subscript I).

In our measurement model (Fig. 1), if there are no shared environmental effects, the two  $\lambda$  paths are unconstrained, and genetic effects exist that are specific to one occasion of measurement; models that assign the unique genetic effect to one or other time of measurement will typically have identical fits. To avoid this confound, we added the constraint to our models:  $\lambda_I \geq \lambda_S$ . Given that the SI was specifically designed to operationalize the DSM PD criteria,

we have made the plausible assumption that the latent liability to the individual PDs would be indexed at least as well by the SI as by the SRQ.

To maximize power, we fitted models, in the software program Mx (Neale et al. 2003), to the raw data from all twins including those without a co-twin and twins who had completed the SRQ but not the SI. Alternative models are evaluated by comparing the difference in their  $\chi^2$  relative to the difference in their degrees of freedom (df), according to the principle of parsimony – models with fewer parameters are preferable if they do not provide a significantly worse fit. We operationalized this balance between explanatory power and parsimony by the use of Akaike’s Information Criterion (AIC; Akaike, 1987; Williams & Holahan, 1994), which is calculated as  $\chi^2 - 2df$ , where df equals the difference in the number of degrees of freedom between the two models being compared. The lower (or more negative) the value of the AIC, the better is the balance between explanatory power and parsimony.

**RESULTS**

**Questionnaire item selection by regression analysis**

Nine SRQ items from the Dysfunctional Personality Questionnaire were selected from our two-stage regression analyses to maximally predict the number of endorsed PPD criteria at interview. As shown in Table 1, these items had varied content. Two of them (items 126 and 127) directly reflect aspects of suspiciousness. Two items reflect general emotionality (items 99 and 116) and two reflect problems in self-concept (items 157 and 165). Items 137 and 149 reflect problems in inter-personal relationships, being either unpredictable or aggressive. The polychoric correlation between the sum of these SRQ items and the number of endorsed PPD criteria at SI ( $n=2731$ ) was +0.34.

Only six SRQ items were selected from the two-stage multiple regression analysis to predict the number of endorsed SPD criteria at interview. As shown in Table 1, four of them (items 92, 107 135 and 148) directly reflect schizoid/introverted traits. Of note, item 126, also seen in PPD, was scored negatively, so not endorsing this item predicts meeting SPD criteria. The correlation between the sum of these items and

Table 1. *Items from the self-report questionnaire selected to assess liability to paranoid, schizoid and schizotypal personality disorder*

| Item number                             | Item content  |
|---|---|
| <b>Paranoid personality disorder</b>    |   |
| -99                                     | On the average, I am calm and even-tempered   |
| 116                                     | My mood will easily change in accordance with the environment   |
| 126                                     | I easily get hurt if someone ridicules me or makes derogatory remarks   |
| 127                                     | It is difficult for me to trust people, since they very often turn their backs on me or let me down                                   |
| 137                                     | I behave in a way that people consider as unexpected or changing  |
| 149                                     | Some of the people that know me think that I am rather aggressive   |
| 153                                     | People treat me as if I were an 'object'  |
| 157                                     | I have been involved in relationships where I was unable to identify whether thoughts and emotions belonged to me or the other person |
| 165                                     | I wonder who I really am  |
| <b>Schizoid personality disorder</b>    |   |
| 92                                      | Among other people, I prefer to stand back  |
| 107                                     | I decided long ago that it was best to have little to do with others  |
| 115                                     | If things turn up suddenly and unexpectedly, I feel completely confused   |
| -126                                    | I easily get hurt if someone ridicules me or makes derogatory remarks   |
| 135                                     | I am afraid of close relationships  |
| -148                                    | Many people regard me as a lively person  |
| <b>Schizotypal personality disorder</b> |   |
| 95                                      | I act the way I feel  |
| -99                                     | On the average, I am calm and even-tempered   |
| -110                                    | I keep both feet on the ground. I stick to what is tangible rather than be lost in reverie  |
| 127                                     | It is difficult for me to trust people, since they very often turn their backs on me or let me down                                   |
| 128                                     | Somehow I feel it is hard for me to know how I should behave among other people   |
| 138                                     | I have seen or heard things that have no logical explanation  |
| 158                                     | In situations where I ought to speak out, I sometimes become speechless and unable to say a single word                               |

A minus sign (–) before the item indicates that it was reverse coded.

the number of endorsed SPD criteria at SI equaled +0.30.

Seven SRQ items from the Dysfunctional Personality Questionnaire were selected from our two-stage multiple regression analysis to predict the number of endorsed STPD criteria at interview. As shown in Table 1, the content of these items was variable and reflected traits of suspiciousness (item 127), social ill-ease (items 128 and 158), soft psychotic-like symptoms (item 138) and lack of even-temperedness (item 99) and groundedness (item 110). The polychoric correlation between the sum of these items and the number of endorsed STPD criteria at SI was +0.32.

### Model fitting

#### *Paranoid personality disorder (PPD)*

We begin by describing in detail the results in model fitting for PPD (Table 2). Model I,

the full model, allowed for qualitative and quantitative sex effects and genetic shared and individual-specific environmental effects for the latent index to PPD as well as the PPD measurements obtained by SRQ and SI. In models II and III, we omitted the qualitative and quantitative sex effects respectively, both times producing an improvement in the fit as indexed by the AIC. Working from model III, we then dropped all shared environmental and all genetic effects respectively in models IV and V. The quality of fit improved with model IV but deteriorated substantially with model V. Working from model IV, we then dropped one at a time, in models VI through VIII, additive genetic effects for the latent liability to PPD and then for the PPD measurements obtained by SRQ and SI. Of these three models, by far the best fit was obtained by model VIII, which omitted the genetic effects for the personal interview. We tried to simplify the model further by

Table 2. Model fitting results for paranoid, schizoid and schizotypal personality disorder

| Model no.          | Sex  |       | Latent |   |   | Self-report |   |   | Interview |   |   | $\chi^2$ | df | AIC     |
|--------------------|------|-------|--------|---|---|-------------|---|---|-----------|---|---|----------|----|---------|
|                    | Qual | Quant | A      | C | E | A           | C | E | A         | C | E |          |    |         |
| <b>Paranoid</b>    |      |       |        |   |   |             |   |   |           |   |   |          |    |         |
| I                  | +    | +     | +      | + | + | +           | + | + | +         | + | + | —        | —  | —       |
| II                 | —    | +     | +      | + | + | +           | + | + | +         | + | + | -0.58    | 1  | -2.58   |
| III                | —    | —     | +      | + | + | +           | + | + | +         | + | + | 5.58     | 10 | -14.42  |
| IV                 | —    | —     | +      | + | + | —           | + | + | +         | — | + | 5.57     | 13 | -20.43  |
| V                  | —    | —     | —      | + | + | —           | + | — | +         | + | + | 52.02    | 13 | 26.02   |
| VI                 | —    | —     | —      | — | + | —           | + | + | +         | — | + | 60.09    | 14 | 32.09   |
| VII                | —    | —     | +      | — | + | —           | + | + | —         | — | + | 47.92    | 14 | 19.92   |
| VIII               | —    | —     | +      | — | + | —           | + | — | —         | — | + | 5.57     | 14 | -22.43* |
| IX                 | —    | —     | +      | — | + | —           | — | + | —         | — | + | 47.92    | 15 | 17.92   |
| X                  | —    | —     | —      | — | + | +           | — | + | —         | — | + | 64.11    | 15 | 34.11   |
| <b>Schizoid</b>    |      |       |        |   |   |             |   |   |           |   |   |          |    |         |
| I                  | +    | +     | +      | + | + | +           | + | + | +         | + | + | —        | —  | —       |
| II                 | —    | +     | +      | + | + | +           | + | + | +         | + | + | 0.01     | 1  | -1.99   |
| III                | —    | —     | +      | + | + | +           | + | + | +         | + | + | 9.05     | 10 | -10.96  |
| IV                 | —    | —     | +      | — | + | —           | + | + | —         | — | + | 9.28     | 13 | -16.72* |
| V                  | —    | —     | —      | + | + | —           | + | + | —         | + | + | 38.79    | 13 | -12.79  |
| VI                 | —    | —     | —      | — | + | +           | — | + | +         | — | + | 32.33    | 14 | 4.33    |
| VII                | —    | —     | +      | — | + | —           | + | + | —         | — | + | 34.49    | 14 | 6.49    |
| VIII               | —    | —     | +      | — | + | —           | + | — | —         | — | + | 11.86    | 14 | -16.14  |
| IX                 | —    | —     | +      | — | + | —           | + | — | —         | — | + | 34.65    | 15 | 4.65    |
| X                  | —    | —     | —      | — | + | +           | — | + | —         | — | + | 43.69    | 15 | 13.69   |
| <b>Schizotypal</b> |      |       |        |   |   |             |   |   |           |   |   |          |    |         |
| I                  | +    | +     | +      | + | + | +           | + | + | +         | + | + | —        | —  | —       |
| II                 | —    | +     | +      | + | + | +           | + | + | +         | + | + | 0.09     | 1  | -1.91   |
| III                | —    | —     | +      | + | + | +           | + | + | +         | + | + | 16.79    | 10 | -3.22   |
| IV                 | —    | —     | +      | — | + | —           | + | + | —         | — | + | 17.50    | 13 | -8.50   |
| V                  | —    | —     | —      | + | + | —           | + | + | —         | + | + | 42.74    | 13 | 16.74   |
| VI                 | —    | —     | —      | — | + | —           | + | + | —         | — | + | 66.03    | 14 | 38.03   |
| VII                | —    | —     | +      | — | + | —           | + | + | —         | — | + | 34.83    | 14 | 6.83    |
| VIII               | —    | —     | +      | — | + | —           | + | — | —         | — | + | 17.59    | 14 | -10.41* |
| IX                 | —    | —     | +      | — | + | —           | — | + | —         | — | + | 34.83    | 15 | 4.83    |
| X                  | —    | —     | —      | — | + | +           | — | + | —         | — | + | 73.09    | 15 | 43.09   |

Qual, Qualitative sex effects; Quant, quantitative sex effects; df, degrees of freedom; AIC, Akaike's Information Criterion (Akaike, 1987).  
 \* Best-fit model.

dropping genetic effects for the SRQ (model IX) and the latent liability (model X) but the fit deteriorated substantially in both cases, indicating that model VIII was our best fit.

Parameter estimates (and 95% confidence intervals) for this best-fit model for PPD are shown in Fig. 2. Five results are noteworthy. First, as indicated in the model fitting, shared environment plays no apparent role in the etiology of PPD. Second, the latent liability to PPD, as indexed by SRQ and SI, is fairly heritable, with an estimated heritability of 66%. Third, the latent liability to PPD is indexed equally well by the SRQ ( $\lambda_S = +0.58$ ) and the SI ( $\lambda_I = +0.58$ ). Fourth, the only factors that impact specifically on SI are individual-specific environmental in nature. That is, the model suggests that a set of factors unique to each member of a twin pair, such as measurement error and

environmental experiences, acts specifically on the SI assessment of PPD. Fifth, by contrast, we found evidence for substantial genetic effects specific to the SRQ assessment of PPD. Using the rules of path analysis, we can calculate that the total heritability of our SRQ PPD-related items is 43%. Of that total, 51% comes from genes that index the latent liability to PPD (and are thus shared with the SI assessed criteria) and 49% are unique to the SRQ measures. These results indicate that the SRQ and SI assessments of PPD traits are not genetically equivalent. The SRQ measures index a broad array of genetic risk factors, only some of which are shared by the SI measures.

*Schizoid personality disorder (SPD)*

The model fitting results for SPD, depicted in Table 2, differed in one important way from

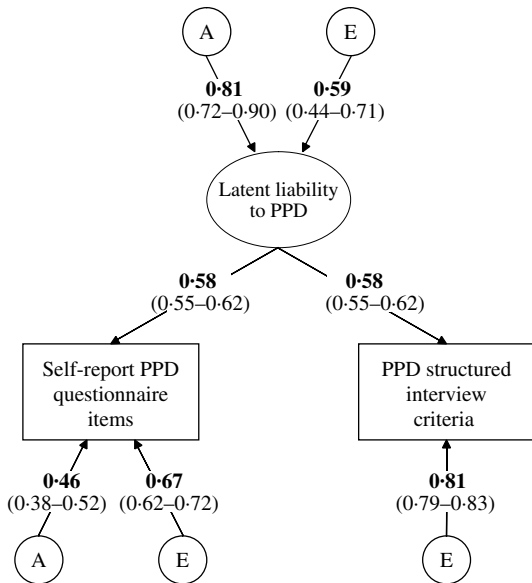


FIG. 2. Parameter estimates in bold and 95% confidence intervals in parentheses from the best-fit model IV (see Table 2) for paranoid personality disorder (PPD). Path estimates must be squared to equal the proportion of variance in the dependent variable accounted for by the independent variable. A stands for additive genetic and E for individual-specific environmental effects.

those obtained for PPD. Model IV, which contains specific genetic effects for both the SI and SRQ, fit slightly better than model VIII, the best-fit model for PPD. Because the difference in AIC values between these two models (0.58 units) is too small to permit us to choose between them with confidence, we present parameter estimates for both in Figs 3a and 3b respectively. Model IV predicts a heritability of the latent liability to SPD of 55%. This liability is equally well indexed by the SRQ ( $\lambda_S = +0.55$ ) and the SI ( $\lambda_I = +0.55$ ). Genes specific to each method of measurement contribute a substantially greater proportion of variance to the SRQ ( $0.44^2 = 19\%$ ) than to the SI ( $0.30^2 = 9\%$ ). Expressed in another way, the proportion of total genetic effects on the SRQ and SI measures of SPD that derive from factors unique to that form of measurement equals 53% and 35% respectively. The parameter estimates from model VIII (Fig. 3b) are similar. The latent liability to SPD is estimated to be slightly more heritable (59%) and genetic specific effects are seen only for the SRQ. The proportion of genetic effects on the SRQ measures of SPD due to genetic effects specific to that form of assessment is 50%.

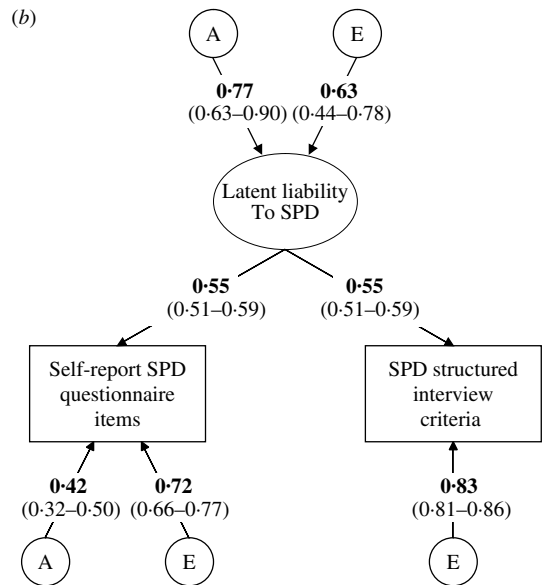
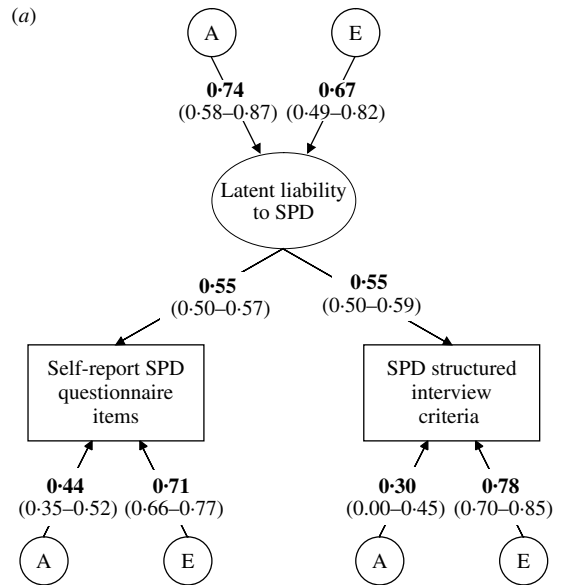


FIG. 3. Parameter estimates in bold and 95% confidence intervals in parentheses from (a) the best-fit model IV and (b) the second-best-fit model VIII (see Table 2) for schizoid personality disorder (SPD). Path estimates must be squared to equal the proportion of variance in the dependent variable accounted for by the independent variable. A stands for additive genetic and E for individual-specific environmental effects.

Schizotypal personality disorder (STPD)

Model fitting results for STPD were identical to those seen for PPD producing model VIII as the

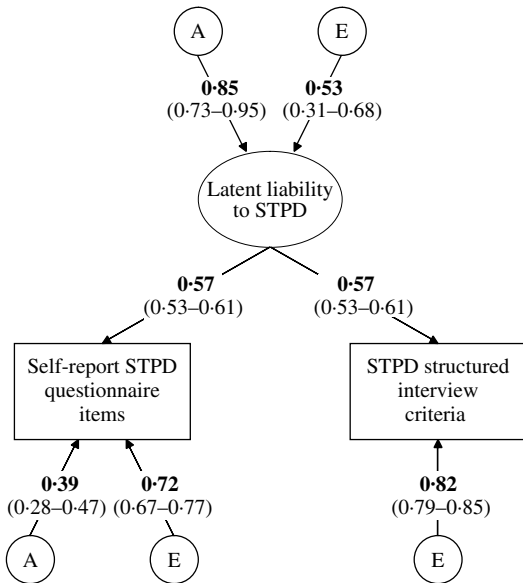


FIG. 4. Parameter estimates in bold and 95% confidence intervals in parentheses from the best-fit model VIII (see Table 2) for schizotypal personality disorder (STPD). Path estimates must be squared to equal the proportion of variance in the dependent variable accounted for by the independent variable. A stands for additive genetic and E for individual-specific environmental effects.

best-fit model. The parameter estimates of this model for STPD are shown in Fig. 4. As seen with PPD, the best-fit model contained no evidence for shared environmental effects for STPD. The heritability of the latent liability to STPD is estimated at 72%, modestly higher than that seen for PPD and SPD. As with the other two PDs, the latent liability to STPD was equally well indexed by the SRQ ( $\lambda_S = +0.57$ ) and the SI ( $\lambda_I = +0.57$ ). As with PPD, but not SPD, in the best-fit model for STPD, the only factors that impact specifically on the SI are individual-specific environmental in nature, but substantial genetic effects specific to the SRQ assessment were seen. Of the total heritability of our SRQ STPD-related items, 61% comes from genes that index the latent liability to STPD and 39% are unique to the SRQ measures.

**DISCUSSION**

In our previous analyses of the cluster A PDs as assessed at SI, we were surprised at the heritability estimates, ranging from 21% to 28% (Kendler et al. 2006), which were substantially

lower than commonly seen for normative personality (e.g. Loehlin, 1992; Lake et al. 2000). Were these results an indication that the cluster A constructs as operationalized by SI were less genetic than more standard personality traits or could they have arisen because of problems of measurement? We obtained a clear answer to this crucial question in the present analyses. When examined using a measurement model that corrects for unreliability of assessment, the heritability of PPD, SPD and STPD was fairly high, ranging from 55% to 72%. Taking into account imperfections in measurement, the etiologic role of genetic factors appears to be at least as strong in these pathologic dimensions of personality function as they are in the better studied normative traits.

These analyses also addressed a subtler but equally important question: what is the relative sensitivity and specificity with which SI and SRQ assessments index the underlying liability to the individual PDs? For all three PDs, the best-fit model estimated the two lambda paths ( $\lambda_S$  and  $\lambda_I$ ) to be equal, meaning that the SI and SRQ assessments were equally sensitive at detecting the latent liability to PPD, SPD and STPD. However, they were not equally specific. For two of the three cluster A PDs, the best-fit model contained no genetic risk factors unique to the SI assessment. For the third PD (SPD), the impact of these unique genetic factors on the SI was modest, and in a second model that fit nearly as well, was absent altogether. By contrast, for all three of these PDs, the best-fit model contained a unique set of genetic risk factors, which impacted substantially on the self-report questionnaire data, that were not reflected in the interview-based measures. Thus, in our results, the SI had greater specificity in indexing the genetic risk for the latent liability to the individual PDs than did the SRQ.

There are two plausible explanations for these findings. First, the items in our SRQ did not map in any one-to-one manner with the DSM criteria for the cluster A PDs. Therefore, the questionnaire-specific genetic effects might result from item content unique to the SRQ. If this is the case, then these results are probably specific to the Dysfunctional Personality Questionnaire and would probably not generalize to other SRQ measurements of PD. Second, the genetic effects specific to the SRQ might



arise because of something more fundamental to the assessment method; that is, SRQ responses might tap a set of genetically influenced traits distinct from those that impact on SI assessments.

We do not have data to directly address the relative plausibility of these two explanations. We can, however, indirectly evaluate them in three different ways. First, as the SRQ items used in these analyses were selected for their ability to predict results of the SIDP interview, perhaps these relationships reflect only idiosyncratic features of our own data. However, the Dysfunctional Personality Questionnaire was also used along within the SIDP interview in an earlier epidemiological study in Norway (Torgersen *et al.* 2001). In 2053 adults living in and around Oslo, the polychoric correlations between the SRQ items depicted in Table 1 and the count of DSM criteria assigned on personal interview with the SID-P were +0.35, +0.39 and +0.33 for PPD, SPD and STPD respectively. These correlations were only slightly higher than those observed in our twin sample and this difference could easily have arisen because these two measures were completed within a week of one another in the epidemiological study and at least a year apart in our twin sample.

Second, the modest correlations between our SI and SRQ assessments for the three cluster A PDs, averaging +0.32, might suggest that these two measures were indeed tapping different traits. However, these results are congruent with the prior evidence for modest levels of agreement between different assessment approaches to PDs when the item content is similar (Zimmerman, 1994). Widiger & Coker (2002) reviewed 18 studies reporting correlations in dimensional PD ratings between SRQ and SI assessments that, for the cluster A disorders, averaged +0.33. Five of these studies used the Personality Disorder Questionnaire (Hyler, 1994), which contained items written to represent each of the DSM PD diagnostic criteria. These studies obtained correlations in the dimensional scores between the SRQ and SI measures of PPD, SPD and STPD, which averaged +0.35, hardly greater than that found in our sample given that these studies typically obtained SRQ and SI measurements at the same time while these measures were separated by a year or more in our sample.

Third, we explored how our results might have varied had we used different SRQ items – picking PPD as an example because its DSM-IV criteria set are relatively homogeneous in content. K.S.K. reviewed all SRQ items and picked seven items typifying the features of PPD. Three of these were among those selected by the regression analyses. Using these items to form a new SRQ assessment of PPD, model VIII was again the best-fit model and it produced parameter estimates very similar to those seen in Fig. 2. For example, using both the ‘clinician-selected’ and regression-derived items for PPD, 49% of the total genetic variance for the SRQ measures come from sources unique to that mode of assessment. The specific genetic factors obtained for the SRQ assessments, at least for PPD, are not simply a result of the apparently variable content of the selected items.

The genetic effects specific to the SRQ that we found for all three cluster A PDs in this sample certainly could reflect idiosyncratic features of our SRQ items and/or our sample. However, our results suggest that the second hypothesized explanation, that these unique genetic factors might reflect something fundamental about how PDs are assessed by SRQ *versus* SI, deserves serious consideration. We hope that future research will further clarify this important question.

### Limitations

These results should be interpreted in the context of four potentially significant methodological limitations. First, we examined dimensional representations of PPD, SPD and STPD rather than diagnoses. This made sense given that our goal was to compare the performance of SRQ and SI evaluations of PDs. This approach was also partially dictated by the rarity of the fully syndromal diagnoses in our sample and meant that much of the information in our analyses about the SI assessment of the cluster A PDs came from individuals who endorsed one or a few criteria. However, many researchers have argued that the PDs are best conceptualized as dimensional constructs rather than dichotomous diagnostic entities (Oldham & Skodol, 2000; Skodol *et al.* 2005; Widiger & Samuel, 2005; Widiger & Simonsen, 2005; Cramer *et al.* 2006). Second, our sample only contained one assessment of PDs by SI and one

by SRQ separated by at least a year. Therefore, we could not determine the degree to which the correlation between our two assessments were the result of method variance, underlying unreliability of assessment and/or true change in PD symptoms over the time separating the two assessments. Third, our sample included only young adult Norwegians. Our results cannot be assumed to extrapolate to other cultural and ethnic groups. Fourth, considerable attrition was observed in this sample from the original birth registry through three waves of contact. Detailed analyses of the predictors of non-response across waves suggest that psychopathology was unrelated to cooperation although retention in the sample was strongly predicted by female sex, monozygosity, older age and higher educational status (Harris *et al.* unpublished observations). In addition, as outlined previously (Kendler *et al.* 2006), cooperation at the stage of personal interview was not predicted by scores for any cluster A FPD dimension as assessed by SRQ. While we cannot be certain that our sample was representative with respect to cluster A psychopathology, these findings suggest that a significant bias is unlikely.

## ACKNOWLEDGMENTS

This work was supported in part by grant MH-068643 from the National Institutes of Health, and by grants from The Norwegian Research Council, The Norwegian Foundation for Health and Rehabilitation, The Norwegian Council for Mental Health, and The European Commission under the program 'Quality of Life and Management of the Living Resources' of the 5th Framework Program (no. QLG2-CT-2002-01254).

## DECLARATION OF INTEREST

None.

## REFERENCES

- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika* **52**, 317–332.
- Clark, L. A. (1993). *Schedule for Nonadaptive and Adaptive Personality (SNAP): Manual for Administration, Scoring, and Interpretation*. University of Minnesota Press: Minneapolis.
- Conte, R. C., Plutchik, R., Karasu, T. B. & Jerrett, I. (1980). A self-report borderline scale. Discriminative validity and preliminary norms. *Journal of Nervous and Mental Disease* **168**, 428–435.
- Cramer, V., Torgersen, S. & Kringlen, E. (2006). Personality disorders and quality of life. A population study. *Comprehensive Psychiatry* **47**, 178–184.
- First, M. B., Spitzer, R. L., Gibbon, M. & Williams, J. B. W. (1995). *Structured Clinical Interview for DSM-IV Axis II Personality Disorders – Patient (SCID-I/P) (Version 2.0)*. New York State Psychiatric Institute: New York.
- Foley, D. L., Neale, M. C. & Kendler, K. S. (1998). Reliability of a lifetime history of major depression: implications for heritability and co-morbidity. *Psychological Medicine* **28**, 857–870.
- Foulds, G. A. (1965). *Personality and Personal Illness*. Tavistock: London.
- Harris, J. R., Magnus, P. & Tambs, K. (2002). The Norwegian Institute of Public Health Twin Panel: a description of the sample and program of research. *Twin Research* **5**, 415–423.
- Helgeland, M. L., Kjelsberg, E. & Torgersen, S. (2005). Continuities between emotional and disruptive behavior disorders in adolescence and personality disorders in adulthood. *American Journal of Psychiatry* **162**, 1941–1947.
- Hyler, S. E. (1994). *Personality Diagnostic Questionnaire-4 (PDQ-4) Unpublished Test*. New York State Psychiatric Institute: New York.
- Jang, K. L., Livesley, W. J., Vernon, P. A. & Jackson, D. N. (1996). Heritability of personality disorder traits: a twin study. *Acta Psychiatrica Scandinavica* **94**, 438–444.
- Kendler, K. S., Czajkowski, N., Tambs, K., Torgersen, S., Aggen, S. H., Neale, M. C. & Reichborn-Kjennerud, T. (2006). Dimensional representation of DSM-IV Cluster A personality disorders in a population-based sample of Norwegian twins: a multivariate study. *Psychological Medicine* doi:10.1017/S0033291706008609.
- Lake, R. L., Eaves, L. J., Maes, H. H., Heath, A. C. & Martin, N. G. (2000). Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. *Behavior Genetics* **30**, 223–233.
- Lazare, A., Klerman, G. L. & Armor, D. J. (1966). Oral, obsessive, and hysterical personality patterns. An investigation of psychoanalytic concepts by means of factor analysis. *Archives of General Psychiatry* **14**, 624–630.
- Livesley, W. J., Jackson, D. N. & Schroeder, M. L. (1992). Factorial structure of traits delineating personality disorders in clinical and general population samples. *Journal of Abnormal Psychology* **101**, 432–440.
- Livesley, W. J., Jang, K. L., Jackson, D. N. & Vernon, P. A. (1993). Genetic and environmental contributions to dimensions of personality disorder. *American Journal of Psychiatry* **150**, 1826–1831.
- Loehlin, J. C. (1992). *Genes and Environment in Personality Development*. Sage Publications: Newbury Park, CA.
- Loranger, A. W. (1988). *Personality Disorder Examination (PSE)*. DV Communications: Yonkers, NY.
- McCrae, R. R. & Costa Jr., P. T. (1990). *Personality in Adulthood*. Guilford Press: New York, NY.
- McGlashan, T. H., Grilo, C. M., Sanislow, C. A., Ralevski, E., Morey, L. C., Gunderson, J. G., Skodol, A. E., Shea, M. T., Zanarini, M. C., Bender, D., Stout, R. L., Yen, S. & Pagano, M. (2005). Two-year prevalence and stability of individual DSM-IV criteria for schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders: toward a hybrid model of axis II disorders. *American Journal of Psychiatry* **162**, 883–889.
- Neale, M. C. (2003). A finite mixture distribution model for data collected from twins. *Twin Research* **6**, 235–239.
- Neale, M. C., Boker, S. M., Xie, G. & Maes, H. H. (2003). *Mx: Statistical Modeling*. Department of Psychiatry, Virginia Commonwealth University Medical School: Box 980126, Richmond, VA 23298.
- Oldham, J. M. & Skodol, A. E. (2000). Charting the future of axis II. *Journal of Personality Disorders* **14**, 17–29.
- Pfohl, B., Blum, N. & Zimmerman, M. (1995). *Structured Interview for DSM-IV Personality (SIDP-IV)*. University of Iowa, Department of Psychiatry: Iowa City.

- Raghunathan, T. E., Solenberger, P. & Van Hoewyk, J. (2000). *IVeare: Imputation and Variance Estimation Software: Installation Instructions and User Guide*. Survey Research Center, Institute for Social Research, University of Michigan: Ann Arbor, MI.
- Reichborn-Kjennerud, T., Czajkowski, N., Neale, M. C., Orstavik, R. E., Torgersen, S., Tambs, K., Roysamb, E., Harris, J. R. & Kendler, K. S. (2006). Genetic and environmental influences on dimensional representations of DSM-IV Cluster C personality disorders: a population-based multivariate twin study. *Psychological Medicine*. Published online: 30 November 2006. doi:10.1017/S0033291706009548.
- SAS Institute (2006). *SAS OnlineDoc 9.1.3*. SAS Institute Inc.: Cary, NC.
- Skodol, A. E., Oldham, J. M., Bender, D. S., Dyck, I. R., Stout, R. L., Morey, L. C., Shea, M. T., Zanarini, M. C., Sanislow, C. A., Grilo, C. M., McGlashan, T. H. & Gunderson, J. G. (2005). Dimensional representations of DSM-IV personality disorders: relationships to functional impairment. *American Journal of Psychiatry* **162**, 1919–1925.
- Torgersen, S. (1980a). Hereditary-environmental differentiation of general neurotic, obsessive, and impulsive hysterical personality traits. *Acta Geneticae Medicae et Gemellologiae* **29**, 193–207.
- Torgersen, S. (1980b). The oral, obsessive, and hysterical personality syndromes. A study of hereditary and environmental factors by means of the twin method. *Archives of General Psychiatry* **37**, 1272–1277.
- Torgersen, S., Kringlen, E. & Cramer, V. (2001). The prevalence of personality disorders in a community sample. *Archives of General Psychiatry* **58**, 590–596.
- Torgersen, S., Lygren, S., Oien, P. A., Skre, I., Onstad, S., Edvardsen, J., Tambs, K. & Kringlen, E. (2000). A twin study of personality disorders. *Comprehensive Psychiatry* **41**, 416–425.
- Widiger, T. A. & Coker, L. A. (2002). Assessing personality disorders. In *Clinical Personality Assessment: Practical Approaches* (ed. J. N. Butcher), pp. 407–434. Oxford University Press: New York.
- Widiger, T. A. & Samuel, D. B. (2005). Diagnostic categories or dimensions: a question for the Diagnostic and Statistical Manual of Mental Disorders – fifth edition. *Journal of Abnormal Psychology* **114**, 494–504.
- Widiger, T. A. & Simonsen, E. (2005). Alternative dimensional models of personality disorder: finding a common ground. *Journal of Personality Disorders* **19**, 110–130.
- Williams, L. & Holahan, P. (1994). Parsimony-based fit indices for multiple-indicator models: do they work? *Structural Equation Modeling* **1**, 161–189.
- Zimmerman, M. (1994). Diagnosing personality disorders. A review of issues and research methods. *Archives of General Psychiatry* **51**, 225–245.