Phenotypic and Genetic Analyses of a Short Measure of Psychosis-proneness in a Large-scale Australian Twin Study

David A. Hay1, Nicholas G. Martin2, Debra Foley3, Susan A. Treloar2, Katherine M. Kirk2, and Andrew C. Heath4

1School of Psychology, Curtin University of Technology and Mental Health Research Institute of Victoria
2Queensland Institute of Medical Research and Joint Genetics Program, The University of Queensland
3Department of Psychiatry, Medical College of Virginia
4Department of Psychiatry, Washington University School of Medicine

Previous genetic analyses of psychosis proneness have been limited by their small sample size. For the purposes of large-scale screening, a 12-item questionnaire was developed through a two-stage process of reduction from the full Chapman and Chapman scales. 3685 individuals (including 1438 complete twin pairs) aged 18-25 years and enrolled in the volunteer Australian Twin Registry returned a mail questionnaire which included this psychosis proneness scale and the Eysenck Personality Questionnaire. Despite the brevity of the questionnaire, item and factor analysis identified four unambiguous and essentially uncorrelated scales. There were (1) Perceptual Aberration – Magical Ideation; (2) Hypomania – Impulsivity/Nonconformity; (3) Social Anhedonia and (4) Physical Anhedonia. Model-fitting analyses showed additive genetic and specific environmental factors were sufficient for three of the four scales, with the Social Anhedonia scale requiring also a parameter for genetic dominance. There was no evidence for the previously hypothesised sex differences in the genetic determination of psychosis-proneness. The potential value of multivariate genetic analysis to examine the relationship between these four scales and dimensions of personality is discussed. The growing body of longitudinal evidence on psychosis-proneness suggests the value of incorporating this brief measure into developmental twin studies.

The concept of psychosis-proneness attracts ever more attention, despite the ongoing debate as to whether responses to such self-report questionnaires should be viewed as subclinical manifestations of psychopathology (Hewitt & Claridge, 1989) or as dimensions of personality (Muntaner et al., 1988). This is part of the wider debate over the dimensional or categorical view of mental illness (Claridge, 1987, 1994; Kendler, 1985). While many would accept the argument that the personality trait of neuroticism may partly underlie anxiety and depression at least at the genetic level (Eaves et al., 1989), such a possibility of an underlying behavioral continuum remains less intuitive for the psychoses and their more unusual behavioral manifestations, despite all the work since the initial development of the concept of schizotypia (Meehl, 1990).

Because of the connotations of the label “psychosis”, it is important to decide whether psychosis proneness is one of the markers for psychosis (Moldin & Erlenmeyer-Kinsberg, 1994) or simply a measure of a trait aspect of personality. Many would feel uncomfortable with the assertion that psychosis-proneness “implies a risk or diathesis for psychosis, even though s/he may never decompensate into clinical psychosis” (Chapman, Chapman & Kwapi, 1994). There is confusion over whether psychosis proneness scores identify genetic liability as indexed by both the higher rates of personality disorder and dimensions of psychosis proneness among relatives of those with schizophrenia (Clementz et al., 1991; Lenzenweger & Loranger, 1989) or are a temporal marker of progression to psychosis within the one individual. Longitudinal studies show that few people with formally diagnosed personality disorders and even fewer of those with high scores on psychosis proneness go on to develop psychosis. Chapman et al. (1994) followed-up 508 of the 534 college students (out of 7800) who ten years earlier had scored at least 1.96 SD above the mean on their psychosis proneness measures. They found 12 had developed psychosis, four of these being schizophrenic. Only two of a control group from the same cohort had developed psychosis. The high-risk group reported more psychotic relatives and had higher scores on dimensions of the Personality Disorder Examination (PD) but no higher rates of formally diagnosed personality disorders. While this does imply some prediction of psychosis, there are two problems beyond the low specificity (496 of the identified 508 high scorers did not become psychotic). Firstly the rates of psychosis are low and may reflect the nature of the sample selected for high academic achievement. Also their measures were not specific to the diagnosis of schizophrenia with four schizophrenics, one delusional disorder, three bipolar disorder, one major depression and three cases of atypical psychosis being found among the twelve. Verdoux et al. (1999) found increased depression in psychosis-prone subjects in their follow-up study of subjects with no history of mood disorder, while research aimed at replicating...
Chapman et al’s ten year follow-up study found that college students scoring highly on a measure of psychosis-proneness (Magical Ideation Scale: Eckblad & Chapman, 1983) exceeded a control group on severity of psychotic-like experiences, ratings of schizotypal, paranoid and borderline personality disorder symptoms, and rates of mood and substance use disorders (Kwapil et al., 1997).

The concept of psychosis proneness would be clarified by a fuller understanding of the factor structure of the scales and their ability to reflect this heterogeneity of outcomes, which are inconsistent with any unitary view of proneness. The main measures of proneness are reviewed by Claridge (1994) and Lenzenweger (1994), while Claridge and Beech (1995) examine their factor analyses. There have been some other schizotypy scales (Raine, 1991; Rust, 1998; Venables et al., 1990), but most of the work to date has used the scales developed by Chapman and Chapman and by Claridge. Starting initially from Meehl’s schizotypy and importantly also from Cattell’s pseudoneuropsychotic schizophrenia (reviewed in Chapman & Chapman, 1985; and Chapman et al., 1982), Chapman and Chapman developed multiple scales where Perceptual Aberration (PER), Magical Ideation (MAG), Social Anhedonia (SAN) and Physical Anhedonia (PAN) have received the most widespread use. These are also scales of Impulsivity–Nonconformity (IMP) and of Hypomania (HYP) (Eckblad & Chapman, 1986). Starting from a different basis, Claridge (1987) modelled his Psychotic Traits Questionnaire (STQ) on DSM-III Schizotypal and Borderline symptoms to come up with the Schizotypal Personality Scale (STA) and Borderline Personality Scale (BTA) respectively.

Simplifying the results of the various factor analyses, there are generally at least two factors. The first reflects the cognitive-perceptual aspects of psychosis-proneness, and has been called “positive symptoms” (Claridge & Hewitt, 1987), “schizophrenia” (Venables et al., 1990) and “schizotypy” (Kelley & Coursey, 1992). This consistent factor indicates that measures such as PER and MAG are highly correlated and it has become common practice to combine these as the PER-MAG scale. The second factor of anhedonia has been described (Lipp et al., 1994) as showing “perfect” separation from the first factor, a result which has been repeatedly observed (Clementz et al., 1991). Less resolved is the question of whether SAN and PAN are independent components of anhedonia. If the first factor can be termed “positive symptoms” (Claridge & Hewitt, 1987), then the parallel to schizophrenia may be continued by positing a second “negative symptoms” factor (Harvey & Walker, 1987) in which case some overlap would be expected (Chapman, Chapman & Raulin, 1978). The correlation of the anhedonia scales, SAN and PAN, is generally low (Kendler & Hewitt, 1992) and often negligible (Kelley & Coursey, 1992; Venables et al., 1990).

When factor analyses have included additional instruments, most often the Eysenck Personality Questionnaire (Eysenck et al., 1985) or at least its Psychoticism (P) scale, more complex patterns of covariance are reported in terms of the overlap of personality and psychosis proneness components. Raine and Allbutt (1989) found SAN loaded on the same factor as Eysenck’s P, while in one of the few studies using the Hypomania (HYP) scale (Bentall et al., 1989), HYP loaded with P on a factor of disinhibition. Kendler and Hewitt (1992) identified a “Nonconformity” factor which included P; PAN and Chapman’s IMP; while SA was part of an independent factor with Eysenck’s Extroversion (E) loading in the opposite direction. Claridge and Hewitt (1987) reported Claridge’s STQ correlated from 0.6 to 0.7 (depending on the exact analysis) with Neuroticism (N) and 0.2 to 0.4 with P. The most complete proposal of how N, P and E relate to the dimensions of psychosis-proneness (Muntaner et al., 1988) has the positive symptoms relating to N and the negative (anhedonia) ones inversely to E. Psychoticism (P) loads on Claridge’s STB (Borderline Personality Disorder) and presumably also on IMP. Such a result is consistent with a ten year longitudinal study of P which showed no higher rates of psychosis in these individuals or their relatives, but rather higher rates of antisocial behavior and higher scores on dimensions of schizotypal and paranoid personality disorder (Kwapil et al., 1997).

An alternative way to identify any heterogeneity in the components of psychosis proneness and personality is by genetic analysis, examining the extent to which the factors differ in their genetic and environmental determinants and whether there is a different covariance structure than is obvious in the phenotypic analysis. Twin studies have proved invaluable in going beyond phenotypic relationships to identify the underlying structure of a whole range of behaviors (Kendler, 1993). The question arises of whether this approach can be applied to psychosis proneness. While twin studies to date have been too small to address this question, they have been important in another way in that they are based on twin registries and represent some of the rare attempts at the use of psychosis proneness measures outside college students.

In the first study, Claridge and Hewitt (1987) administered the STA, STB and the EPQ by mail to 108 MZ and 102 DZ pairs in the Birmingham Family Study Registry. A model including additive genetic effects and specific environment fitted the STA data. For the shorter STB questionnaire, the same model fitted best, although a common and specific environmental model could not be excluded. Females scored higher on STA than males and it was suggested that schizotypy may be under greater genetic control in males than females. However the sample sizes were too small for an adequate analysis.

The same issue arose in a second mail survey of over 400 twins in the Virginia Twin Registry using ten different psychosis proneness scales (Harvey & Walker, 1987). An additive genetic and specific environment model fitted all measures except Chapman’s Perceptual Aberration and another Perceptual Aberration scale derived from the STA, where a genetic component was not needed since a model with only common and specific environment effects was adequate. However when they derived factor scores as discussed earlier, the outcome of genetic analysis was uniform across (1) positive trait schizotypy; (2) nonconformity and (3) extroversion or social schizotypy (comprising social anhedonia plus paranoid ideation from the STA). In all cases an additive genetic plus specific environment model fitted, with the estimates of the genetic effect ranging from
52% to 69%. As the authors point out, their use of abbreviated scales with the associated impact on reliability may mean these are underestimates of the genetic effect. There were significant sex differences on six scales with the male mean being higher on five of these, the opposite result to that reported from the Birmingham study. Preliminary analyses based on the modest sample size suggested the only sex differences in genetic control may be on nonconformity, which is unrelated to the STA where sex-differences were postulated in the previous study.

Two further twin studies have addressed more specific aspects of psychosis proneness. Kendler et al. (1991) used a combination of structured interviews, self-report measures, attentional tasks and eye tracking with a small sample (29 twin pairs). They identified independent factors of positive and negative symptom schizotypies from the structured interview and parallel independent measures of positive trait schizotypy and trait anhedonia from the self-report measures derived from Claridge and from Chapman and Chapman. The interview and self-report positive schizotypies correlated strongly (0.83) whereas the 0.45 correlation of the negative measures was much less, reflecting the wider range of behaviors tapped in the structured interview than in the self-report. While little emphasis can be placed on genetic estimates derived from such small numbers, the trait anhedonia values stand out with an MZ correlation of 0.97 and a DZ value of 0.18. Such a disparity is inconsistent with most genetic models and warrant replication with a more adequate sample size.

In contrast, the final twin study (Kendler et al., 1987) assessed a very large number of twins (3810 pairs in the Australian Twin Registry) but on a very modest measure, a four-item suspiciousness scale derived from the Eyseck Personality Questionnaire (EPQ). While the rationale was the relationship of suspiciousness to paranoid personality disorder, there was a modest correlation with EPQ Neuroticism, which is a general index of nonpsychotic predisposition. When it came to genetic modelling, there were some gender issues paralleling the problems with the two large-scale twin schizotypy surveys reviewed earlier. It was difficult to discriminate between a simple genetic model and a complex environmental model which postulated different patterns of common environment for females and males.

Unfortunately all the substantial genetic studies of psychosis-proneness to date have been based on twin studies and have not been replicated with other approaches to genetic analysis such as family or adoption studies. There certainly have been family studies of psychosis proneness (Clementz et al., 1991; Katsans et al., 1990; Lenzenweger & Loranger, 1989; Moldin et al., 1990), but the focus has been more on the mean and distributions of scores of relatives at risk for psychosis and not on the familial correlations needed to estimate genetic effects. The lack of comparable nuclear family data is unfortunate given that there is a specific component of twin similarity for schizophrenia over and above that predicted on a genetic and environmental model (McGue et al., 1983) and because there may be some discrepancies between twin and adoption approaches to related personality measures such as those on the EPQ (Eaves et al., 1989). In one of the few family studies, Grove et al. (1991) studied 61 first degree relatives of schizophrenics on a large battery of structured interview, questionnaire and laboratory measures. While the sample size is so small that large standard errors attach to the estimates of heritability, it is more significant that the relatives did not differ from controls on many measures including Perceptual Aberration (PER) and MMPI schizotypy questions. (They were significantly higher on Physical Anhedonia (PAN) and on an interview measure of schizotypal personality). This inability of many measures to identify those at genetic risk may indicate limits to their validity. However Grove et al point out there may be a confounding factor with relatives “faking good” because they know they are under scrutiny. Obviously this is not an issue in the twin studies where subjects are targeted because they are twins and not because of any genetic vulnerability. An associated study involving a multi-center program (Clementz et al., 1991) found no consistent evidence of genetic effects on the PER scale, but again a high sib-sib correlation (0.5) in PAN and argued anhedonia measures are more indicative of any genetic predisposition to psychosis than the cognitive-perceptual aspects of PER and related measures.

Presented here is the first report of a large-scale twin study of psychosis-proneness and its correlates. This paper deals only with validation of the scale construction and initial genetic analysis. The modern methods of biometrical genetic demand large numbers of subjects especially for multivariate analyses. In many cases this has led to the use of mail questionnaires where the length must be constrained if a reasonable response rate is to be achieved. A key question is the extent to which the number of items can be reduced while still identifying and replicating the factor structure commonly found with the full-length questionnaires. Can the same pattern of genetic and environmental determinants be identified? In addition, are there independent scales or is there significant overlap between all these scales, supporting the concept of a general “psychosis”?

Method
Subjects
Twins were drawn from the Australian National Health and Medical Research Council Twin Register (ATR). The ATR is a volunteer register begun in 1978 and with about 30,000 pairs of all types and all ages enrolled and in various stages of active contact; we estimate that this represents 10–20% of living twins in Australia. In 1989, an extensive Health and Lifestyle Questionnaire (HLQ) was mailed to 4269 pairs born 1964–1971 (i.e., aged 18–25). Most of these twins had been recruited when at school some ten years earlier. It was not surprising, therefore, that despite extensive follow-up efforts, we were unable to re-establish contact with 1000 pairs. Those twins who could be located but who failed to return a completed questionnaire were followed-up by telephone up to five times. In the course of the phone contact they were asked to complete an abbreviated telephone interview to obtain basic demographic information. Both members of 2294 pairs (70% of contactable pairs) completed a questionnaire or abbreviated phone interview, plus a further 474 single twins, making an
individual cooperation rate of 5074/6122 (83%) of those with whom contact was established.

Development of the Mini-psychosis Proneness Scale (PPQ)

Given the other areas required to be included in the questionnaire package, it was possible to include only 12 psychosis-proneness items. These were chosen in the following two-step procedure using university students of similar age to the intended questionnaire recipients (18–25 years). These were selected from the Schools of Behavioral Science and Biological Sciences at LaTrobe University. Although done in class time, completion was voluntary and did not count towards course requirements.

Stage 1 involved a 307 item questionnaire comprising the complete Chapman and Chapman Scales, namely:

- Perceptual Aberration — PER (35 items) focussing on distortions in the perceptions of one’s body and senses.
- Magical Ideation — MAG (29 items) focussing on the tendency to hypothesise unusual or bizarre patterns of causality. As the scale was being planned for a general population, item 11 (“At times, I have felt that a professor’s lecture was meant especially for me”) was excluded.
- Hypomania — HYP (48 items) focussing on mild representations of the grandiose ideas and excessive energy associated with mania.
- Impulsivity/Nonconformity — IMP (51 items) focussing on impulsivity, antisocial behavior and insensitivity to others. Item 22 “In school, I sometimes got in trouble for cutting-up” was changed to the Australian “playing up”.
- Social Anhedonia — SAN (40 items) focussing on lack of capacity to experience pleasure in social contact.
- Physical Anhedonia — PAN (61 items) focussing on lack of capacity to experience pleasure in response to physical stimuli.
- Infrequency (13 items) which is a checklist of very unusual behaviors. It functions analogously to the EPQ Lie scale but would also identify those so disturbed their questionnaire response would have to be considered invalid.

In addition the questionnaire included the 30 item General Health Questionnaire, a measure of recent and current wellbeing and psychopathology (Burvill & Kruiman, 1983). Random number tables were used to allocate items to positions on the questionnaire to ensure an adequate spread of the items from each scale throughout.

Excluding those with a score of 3 or more on the Infrequency Scale, there were 173 first and second year undergraduate students who completed the entire questionnaire, all of whom were naïve to the rationale of the questionnaire and its items. The levels of intrascale homogeneity were consistent with those reported in earlier studies (Chapman & Chapman, 1985) with coefficient alphas ranging between 0.82 and 0.89, except for Infrequency where as anticipated from the nature of the scale it was very low at 0.18.

After factor analysis with Varimax rotation, the number of items was reduced to a total of 48 across all the scales (except Infrequency) using the following procedure. Two raters independently identified potentially useful items based on the following criteria, in descending order of importance:

1. High item-total correlations.
2. High frequency in both sexes to avoid problems in modelling sex differences in inheritance with items rarely answered by one sex (Claridge & Hewitt, 1987). Items checked by less than 10% or more than 90% of the respondents were also excluded.
3. Low loading on factors other than the one claimed by Chapman and Chapman (1985).
4. Limited face validity because of the possibility of “faking good” (Venables et al., 1990).
5. Reasonable level of general acceptance.

All of these are obvious except the last. Some of the questions in the Chapman and Chapman scales are unusual, especially in the Perceptual Aberration and Magical Ideation scales. While they may be accepted by the Psychology students on whom so much of the work with these scales has been done, members of the general population may be much less accepting of what they regard as “strange” questions and fail to complete the questionnaire. This may have contributed to the lower compliance in the Kendler et al. study (1991) where only 44/58 individual twins completed this questionnaire at home, despite all having been sufficiently motivated to spend over three hours completing face-to-face psychiatric interviews and neuropsychological assessment.

Obviously there is some element of subjectivity with this process of item reduction but there was little disagreement between raters. In the case of the few disputes, the data set was used to generate correlations of different potential short versions and the full-scale instrument in order to make final decisions over particular items. This 48 item questionnaire was then given to a further 150 Psychology students in a 60 item questionnaire in which the 12 item GHQ was also embedded. The same item reduction procedure was repeated when the even smaller final 12 item version had to be prepared. The final questionnaire is shown in Table 1, together with the frequencies of response both for the twin sample described below and for the Chapman and Chapman samples of college students. Chapman and Chapman (1985) provide references to the publications in which the basic frequency data on their various scales are provided. Their sample is predominantly white and aged 18–19 years of age, very similar to that utilised here.

While the 12 items given in Table 1 are derived from six Chapman and Chapman scales, the analyses discussed in the Introduction would suggest there are at most four logical groupings, combining the Perceptual Aberration and Magical Ideation items to form the usual (Pope & Kwapil, 2000) PER-MAG (Score 1) and combining the Hypomania and Impulsivity/NonConformity scales to form one HYP-IMP scale (Score 2). Social Anhedonia
(SAN) and Physical Anhedonia (PAN) are left as separate scales (Score 3 and 4 respectively), given the data on their independence. The validation of this hypothesised structure is the first aspect of the Results section.

**Instruments**

The 12 PPQ items (Table 1) were randomised amongst 54 items from a short form of Cloninger’s Tridimensional Personality Questionnaire (TPQ) (Heath et al., 1994) which includes scales for Harm Avoidance, Reward Dependence and Novelty Seeking, none of which are analysed here. Elsewhere in the HLQ questionnaire was another personality section which included the Eysenck Personality Questionnaire (Revised) Short Form (EPQ-R-S) with 12 items each for the Extraversion, Psychoticism, Neuroticism and Lie (Social Desirability) scales (Eysenck et al., 1985). The personality items were not included in the telephone interview for twins who failed to return the mail questionnaire and a maximum of 3685 twins (60%) answered the psychosis-proneness (PPQ) items. However, after allowing for listwise deletion of missing responses to PPQ items, there were 1438 pairs and a further 576 singles with complete answers to all PPQ items, representing 44% pairwise and 54% individual completion by contactable twins.

**Zygosity**

Zygosity of twins was established from their response to standard items in the HLQ about physical similarity and being mistaken for each other. Such items have been shown by ourselves and others to be at least 95% accurate when judged against genotyping results (Kastiel & Eaves, 1976; Kendler, 1993; Martin & Martin, 1975). We improved on this by selecting for further investigation any pair whose answers were not completely consistent (within or between co-twins) with either mono- or dizygosity. These pairs were telephoned to detect the source of any confusion and about 80% were readily resolved on the phone. Those still equivocal were asked to send us photographs taken at several stages of their lives and most were then assigned with little hesitation by the project staff leaving but a few genuinely doubtful cases. Where possible, blood was subsequently obtained for genotyping these few uncertain pairs. In the course of a twin asthma study (Duffy et al., 1994) we have recently genotyped 313 same-sex pairs, half from this

### Table 1

<table>
<thead>
<tr>
<th>Item Key</th>
<th>Item Description</th>
<th>Female Count</th>
<th>Male Count</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perceptual Aberration (PER)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.(PER17)</td>
<td>Sometimes part of my body seems smaller than it really is</td>
<td>15</td>
<td>23</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>2.(PER31)</td>
<td>Sometimes I feel like everything around me is tilting</td>
<td>27</td>
<td>22</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td><strong>Magical Ideation (MAG)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.(MAG8)</td>
<td>I have wondered whether the spirits of the dead can influence the living</td>
<td>40</td>
<td>34</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>2.(MAG10)</td>
<td>I have felt that I might cause something to happen just by thinking too much about it</td>
<td>38</td>
<td>33</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td><strong>Hypomania (HYP)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.(HYP7)</td>
<td>In unfamiliar surroundings, I am sometimes so assertive and sociable, that I surprise myself</td>
<td>42</td>
<td>43</td>
<td>ns</td>
</tr>
<tr>
<td>2.(HYP18)</td>
<td>I often have moods where I feel so energetic and optimistic that I feel I could outperform almost anyone or anything</td>
<td>51</td>
<td>62</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>Impulsivity/Non conformity (IMP)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.(IMP6)</td>
<td>It would embarrass me a lot to have to spend a night in jail</td>
<td>20</td>
<td>34</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>2.(IMP21)</td>
<td>I usually find myself doing things “on impulse”</td>
<td>51</td>
<td>54</td>
<td>ns</td>
</tr>
<tr>
<td><strong>Social Anhedonia (SAN)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.(SAN8)</td>
<td>Although there are things that I enjoy doing myself, I usually seem to have more fun when I do things with other people</td>
<td>19</td>
<td>18</td>
<td>ns</td>
</tr>
<tr>
<td>2.(SAN17)</td>
<td>I prefer hobbies and leisure activities that do not involve other people</td>
<td>23</td>
<td>24</td>
<td>ns</td>
</tr>
<tr>
<td><strong>Physical Anhedonia (PAN)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.(PAN20)</td>
<td>Trying new foods is something I have always enjoyed</td>
<td>30</td>
<td>35</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>2.(PAN42)</td>
<td>I seldom care to sing in the shower</td>
<td>59</td>
<td>60</td>
<td>ns</td>
</tr>
</tbody>
</table>
cohort, whose zygosity had been assigned using the above procedures. Using at least eight independent highly polymorphic short tandem repeat markers (P Dz|Conc < 10⁻³), only five errors were found (1.6%); all were MZ pairs who had previously been called DZ (four female, one male).

### Statistical methods

Factor and other standard analyses were by SAS 6.09 (SAS Institute, 1993). The resemblance of the scores of Twin 1 and Twin 2 is a potential complication in phenotypic analyses which treats the entire sample as independent individuals, but is also a means of providing closely matched replicates of the factor structure. Factor analyses carried out separately for Twin 1 and Twin 2 were essentially identical to each other and to that on the pooled sample, both on the factor loadings of individual items and on the variance accounted for by each factor. Thus only the pooled data are presented here. Polychoric correlations were calculated using PRELIS 2.1 (Jöreskog & Sörbom, 1993a) and structural modelling was implemented using LISREL 8.1 (Jöreskog & Sörbom, 1993b). The standard procedure of first fitting an ACE model (that is, one which hypothesises Additive genetic, Common family environment and specific individual Environment influences and then estimating the effects of each using a chi-square test of goodness of fit) was used (Neale & Cardon, 1992).

### Results

#### Characteristics of the Sample

The distribution of the entire sample on the four putative scales is shown in Table 2. Because items were chosen deliberately in terms of high frequency of self-report in the student surveys, there is not the extremely skewed distribution which generally characterises scales such as the EPQ Psychoticism (Claridge & Hewitt, 1987).

The analyses presented here are based on the entire sample but only after considering five potential confounding variables:

- **Age.** As expected with the deliberately restricted age range of this sample, there were few statistically significant correlations of any of the four scores with age, even given the very large samples. The two significant ones were in the female twins where age correlated slightly with Scale 1 = -0.072 (p = 0.0009) and Scale 2 = -0.097 (p = 0.0001). Such small correlations will have minimal impact on estimates of genetic and environmental parameters and no age corrections were made.

- **Zygosity.** There were no zygosity effects on any of the scales significant at even the p = 0.05 criterion.

- **Birth order.** There were no significant differences between the Twin 1 and Twin 2 scores on any of the scales.

- **Responder bias.** As participation in this voluntary questionnaire may reflect personality variables (Neale and Eaves, 1993), scores on the four scales were compared by t tests between individuals where both members of the pair replied and individuals where the cotwin did not respond. No differences even approached significance, implying the absence of any major responder bias.

- **Gender.** Claridge and Hewitt (1987) and Kendler and Hewitt (1992) speculated that sex differences may exist in genetic expression on various psychosis proneness scales. The appropriate t tests on the scores are given in Table 2. The slightly higher score for females on the PER-MAG (Score 1) parallels the result reported by Claridge and Hewitt with the STA scale. But inspection of the individual items (Table 1) shows that the effect is not consistent. Females scored higher on three of the four and significantly lower on the other. While there is a major difference on HYP-IMP (Score 2), once again it is clear from Table 1 that the effect is not consistent across different items. Analysis of the factor structure (Table 3) confirms the value of ignoring sex differences and pooling data.

### Analyses of the Phenotypic Factor Structure

Table 3 summarises the principal components analysis with varimax rotation used to identify the factor structure among the 12 items. This analysis pools the responses of both sexes. Given the sex differences (Table 1), the factor analysis was initially done separately for each sex. There were few differences, the only one of any magnitude being that the item with the largest sex difference (IMP 1 “it would embarrass me a lot to have to spend a night in jail”) loaded in males on Factor 4 as well as on Factor 2. The effect was not sufficient to warrant carrying out all the analyses separately for females and males.

Rather than simply interpreting the item loadings, an understanding of the factors follows most easily from the lower part of Table 3 which presents the correlations between the factor scores for individuals and their scores on the four a priori scales. The correlation of each scale only with the predicted factor emphasises the independent four

---

**Table 2**

<table>
<thead>
<tr>
<th>Scale 1 (PER-MAG)</th>
<th>Scale 2 (HYP-IMP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>Female</td>
</tr>
<tr>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Mean (SD) 1.20 (1.12) 1.11 (1.09) 1.64 (1.13) 1.95 (1.13)

\[ t = 2.36, \ p = 0.02 \]
\[ t = 7.89, \ p < 0.0001 \]

<table>
<thead>
<tr>
<th>Scale 3 (SAN)</th>
<th>Scale 4 (PAN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>Female</td>
</tr>
<tr>
<td>0</td>
<td>68</td>
</tr>
<tr>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
</tbody>
</table>

Mean (SD) 0.42 (0.67) 0.42 (0.66) 0.88 (0.68) 0.94 (0.72)

\[ t = 0.26, \ p = 0.79 \]
\[ t = 2.07, \ p = 0.04 \]
dimensions covered by these four questions and their relationship to the Chapman and Chapman model. For convenience, the scale-scale correlations are presented in Table 6 as part of the matrix of intercorrelations with the dimensions of the Eysenck Personality Questionnaire. Presented separately for each sex, they indicate no significant correlations between the scales, except for a modest correlation between PER-MAG and HYP-IMP which is consistent in both sexes. These scales are derived directly from the data and have not been orthogonally transformed.

### Genetic Analysis of the Four Psychosis Proneness Scales

The genetic analysis for each scale is based upon the six polychoric correlations given in the upper part of Table 4, namely MZFF, MZMF, DZFF, DZMM, DZFM and DZMF (where F = female, M = male, and FM and MF refer to opposite-sex pairs in which the first born twin is female and male respectively). Few sex differences have been identified and the adequacy elsewhere in these analyses of the pooled data argues for genetic model-fitting just to the single MZ and single DZ correlations calculated from all the twin pairs together and shown in the lower part of Table 4. However, the simultaneous fit to all six correlations allows an immediate test both of the suggestion that there may be sex differences in genetic control (Claridge and Hewitt, 1987) and of the general assumption that data on these measures can be pooled over sexes. If there were sex-specific effects, then fitting the same models simultaneously to the six statistics would not work unless the estimates for the same parameter were allowed to vary between the sexes.

The results of the model-fitting are shown in Table 5. While path coefficients are estimated, for clarity these have been squared in Table 5 to give components of variance which sum to 1.0. The hypothesis of sex-specific effects is clearly rejected in that excellent fits were found for all four scales. There was uniformly no significant effect of C, so that fitting a reduced AE model gave essentially the same goodness of fit as the full ACE model. The significant failure of the chi-square tests of goodness of fit for the CE and E models indicates that additive genetic effects are an essential part of any model to explain the data. In the case of SAN, the fit of the AE model is poor because the DZ correlation is considerably less than expected under an additive model, given the size of the MZ values. The possibility of genetic dominance effects was therefore considered, and a model incorporating this greatly improved the fit to the data.

One particular strength of contemporary genetic modelling is the ability to dissect the multivariate genetic and environmental structure of multiple measures. However this was not attempted here, since the independence of the four scales has already been established. Given the small intra-individual correlations between the scales, the co-twin correlations were so small (even in MZ twins) that analysis was impractical.

### Relationship to the EPQ

Muntaner et al. (1988) has made explicit predictions of how the dimensions of the EPQ should relate to the psychosis-proneness scales. Table 6 shows such predictions are confirmed with these scales. PER-MAG correlates with N (neuroticism) and to a much lesser extent with P (Psychoticism) and HYP-IMP with both P and E (Extraversion). The only relationship with the Anhedonia scales is P, which correlates reasonably with SAN and to a lesser extent with PAN. The direction of this difference would be predicted from the antisocial component of P (Chapman, Chapman & Kwapil, 1994). Given the sex differences in means on the psychosis-proneness scales, it

### Table 3

Factor Loadings for the 12 Items and Pearson Product Moment Correlations between the Four Factors and the Four Scales. (Note that for ease of comparison, items are given in the same order as Table 1. This is not the randomised order used in the questionnaire).

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>PER1</td>
<td>0.44</td>
<td>0.29</td>
<td>0.00</td>
<td>0.27</td>
</tr>
<tr>
<td>PER2</td>
<td>0.68</td>
<td>0.04</td>
<td>0.00</td>
<td>0.16</td>
</tr>
<tr>
<td>MAG1</td>
<td>0.60</td>
<td>0.03</td>
<td>0.06</td>
<td>-0.24</td>
</tr>
<tr>
<td>MAG2</td>
<td>0.65</td>
<td>0.07</td>
<td>0.02</td>
<td>0.08</td>
</tr>
<tr>
<td>HYP1</td>
<td>0.09</td>
<td>0.47</td>
<td>-0.12</td>
<td>-0.25</td>
</tr>
<tr>
<td>HYP2</td>
<td>0.21</td>
<td>0.64</td>
<td>-0.06</td>
<td>-0.06</td>
</tr>
<tr>
<td>IMP1</td>
<td>0.20</td>
<td>-0.50</td>
<td>-0.26</td>
<td>-0.07</td>
</tr>
<tr>
<td>IMP2</td>
<td>0.09</td>
<td>0.67</td>
<td>-0.08</td>
<td>0.02</td>
</tr>
<tr>
<td>SAN1</td>
<td>-0.04</td>
<td>-0.06</td>
<td>-0.81</td>
<td>0.00</td>
</tr>
<tr>
<td>SAN2</td>
<td>-0.15</td>
<td>0.04</td>
<td>0.76</td>
<td>0.03</td>
</tr>
<tr>
<td>PAN1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.08</td>
<td>-0.69</td>
</tr>
<tr>
<td>PAN2</td>
<td>0.00</td>
<td>0.10</td>
<td>-0.04</td>
<td>0.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variance accounted for</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>PER-MAG</td>
<td>0.16</td>
<td>0.14</td>
<td>0.14</td>
<td>0.11</td>
</tr>
</tbody>
</table>

### Table 4

Polychoric Correlations (with standard errors) for Psychosis-Proneness Scores in Twins

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Scale 1 (PER-MAG)</th>
<th>Scale 2 (HYP-IMP)</th>
<th>Scale 3 (SAN)</th>
<th>Scale 4 (PAN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MZFF</td>
<td>0.34 (0.06)</td>
<td>0.23 (0.05)</td>
<td>0.512 (0.06)</td>
<td>0.37 (0.06)</td>
</tr>
<tr>
<td>MZMM</td>
<td>0.33 (0.08)</td>
<td>0.35 (0.07)</td>
<td>0.479 (0.09)</td>
<td>0.41 (0.06)</td>
</tr>
<tr>
<td>DZFF</td>
<td>0.21 (0.06)</td>
<td>0.12 (0.06)</td>
<td>0.070 (0.09)</td>
<td>0.10 (0.07)</td>
</tr>
<tr>
<td>DZMM</td>
<td>0.13 (0.08)</td>
<td>0.14 (0.09)</td>
<td>0.137 (0.12)</td>
<td>0.21 (0.10)</td>
</tr>
<tr>
<td>DZFM</td>
<td>0.07 (0.09)</td>
<td>0.23 (0.08)</td>
<td>0.105 (0.12)</td>
<td>0.14 (0.10)</td>
</tr>
<tr>
<td>DZMF</td>
<td>0.12 (0.10)</td>
<td>0.07 (0.10)</td>
<td>0.042 (0.12)</td>
<td>-0.05 (0.11)</td>
</tr>
<tr>
<td>MZ pooled</td>
<td>0.34 (0.05)</td>
<td>0.28 (0.04)</td>
<td>0.51 (0.05)</td>
<td>0.39 (0.04)</td>
</tr>
<tr>
<td>DZ pooled</td>
<td>0.14 (0.04)</td>
<td>0.15 (0.04)</td>
<td>0.09 (0.06)</td>
<td>0.10 (0.05)</td>
</tr>
</tbody>
</table>
should be noted that the patterns of correlations were identical in both sexes.

Discussion

While it may seem ambitious and psychometrically unwise to reduce the 264 items of Chapman and Chapman's six scales to 12 items on four scales, the results in Table 3 show the efficacy of this procedure. Despite the imprecision which would be predicted with such a small range of items, these 12 fall into exactly the four factors which have been hypothesised from studies of the full-scale instruments and with minimal overlap between factors. There are three reasons why this procedure may have worked here. Firstly the original scales are robust and homogeneous with high coefficient alphas, which does imply some degree of item redundancy. Secondly our two stage procedure in identifying items has selected those items with a particularly high relationship to specific scales and little loading on other scales. Thirdly, items were rejected if they had a particularly low or high frequency of response and discriminability is maximised on yes-no responses such as these when the probability of response in either direction is 50%.
It is possible the construction of this short form has a significant advantage over the full version. The long form is “transparently concerned with psychopathology” (Chapman et al., 1994) and individuals may be defensive about their replies. One criterion for the items chosen for the short form was that they be less “transparent” (while still loading highly on the underlying scale). Taking such items and embedding these in a more general personality inventory may have resulted in more accurate and honest answers.

These results clearly counter the obvious objection that this highly selective culling of items has resulted in a scale with no relationship to the original psychosis-proneness measures. If that were so, then the four scales derived explicitly from the previous full-scale studies would not have been identified so clearly with the four factors. In addition, there is close correspondence between the results of this genetic analysis and of the much smaller, previous ones. It was necessary here to include a component for genetic dominance in the analysis of Social Anhedonia because the DZ twin correlation was much less than half that for MZ twins. Kendler and Hewitt (1992) reported MZ and DZ correlations on the full Chapman and Chapman scale of 0.70 and 0.23 respectively. Few of their other measures showed such an MZ-DZ discrepancy. For comparison, their Physical Anhedonia correlations were 0.57 and 0.34. While their numbers of twin pairs were so small that it was not necessary to fit a dominance component, it is reassuring for the validity of the short scales that there was this correspondence between the two Social Anhedonia measures. This result is corroborated even further by the extreme difference of MZ and DZ correlations on the Trait Anhedonia scale of the Kendler et al. (1991) study discussed earlier. Admittedly their scale loaded also on Physical Anhedonia, albeit to a lesser extent than on Social Anhedonia. But it should be noted that there was some suggestion here of dominance also for the Physical Anhedonia.

The twin analysis of genetic factors is subject to the usual assumptions underlying this methodology (Kendler, 1993). However it would be very difficult with these data to posit any pattern of environmental influences which would mimic the role of genetic factors in the results of Tables 4 and 5. Such environmental factors would have to be totally uniform across the sexes (even in opposite sex pairs), would have to be specific to each scale (given the lack of any correlation between the scales or of any significant twin-twin cross correlations) and would have to act just on each twin individually with no influences common to both members of the pair (in the absence of any common environmental effects on any of the scales).

In the case of the SAN measure, various other explanations exist for why the DZ correlation is so much less than predicted, given the MZ value. At least that of non-additive genetic effects does make an explicit prediction of the MZ/DZ relative magnitudes which turns out to be consistent with the data. An even more convincing rationale concerns the relationship of SAN to Eysenck’s extraversion, discussed by Muntaner et al. (1988) and confirmed here by the high loadings in Table 6. Extraversion is the one personality dimension where genetic dominance has previously been identified (Eaves et al., 1989) although the magnitude of the effect was not as large or unambiguous as that observed here.

The pattern of MZ/DZ differences on the Social Anhedonia scale is compatible with a single dominant gene or with the cumulative effects of dominance at many loci. Certainly there are other possibilities such as epistatic interactions between two or more genes and competition/co-operation between twins (Neale & Cardon, 1992). Dominance has implications for the role of natural selection and it is possible to speculate what may be the selective issues specifically concerned with this dimension of psychosis proneness. Kendler et al. (1991) expressed some surprise at finding trait anhedonia had such a broad relationship to all facets of negative symptoms and concluded that it “may reflect, to a moderate extent, the underlying deficit that was detected as negative schizotypal symptoms at personal interview”. If Anhedonia and a general indifference to others, does have a significant relationship to social and specifically sexual interactions as the putative negative symptoms relationship would imply (Harvey & Walker, 1987), then its survival value and hence its pattern of inheritance are consistent (Hay, 1985). As Chapman et al. (1994) found in their ten year follow-up study, Social Anhedonia may also be a significant variable in confounding the long-term impact of the positive symptom measures. High Magical Ideation plus high Social Anhedonia identifies those highest in psychotic-like experiences at the initial assessment and on schizotypal symptoms ten years later. The distinction here between the Anhedonia and the more cognitive-perceptual components of psychosis-proneness may support the argument of Clementz et al. (1991) that Anhedonia is the most useful familial indicator of liability to psychosis.

There are two directions in which this scale and these extensive twin data can be extended. First is multivariate genetic analysis of the relationship of these scales to the Eysenck and Cloninger Personality dimensions. As has repeatedly been demonstrated in behavior genetics, modest phenotypic correlations between measures may reflect large genetic and environmental commonalities but in different directions, for example in the relationship of the suspiciousness and the tough-minded items on EPQ Psychoticism (Heath & Martin, 1990). A fuller analysis is planned to explore and hopefully resolve some of the current discrepancies between the previous, much smaller studies of these interrelationships. The genetic and environmental covariances of the EPQ and Cloninger’s TPQ have already been explored in another sample of 2680 older twin pairs in the Australian Twin Registry, leading to the conclusion that a five factor model (the TPQ Harm Avoidance, Novelty Seeking and Reward Dependence plus the EPQ Extroversion and Lie scales) provide the best description of the genetic structure of personality (Heath et al., 1994). To date there has been little theoretical, far less empirical evidence presented for the relationship between the TPQ and psychosis proneness.

Second is the possibility of developmental and longitudinal studies, which are increasingly the focus of general use of the psychosis-proneness scales. In addition to the longitudinal studies that focussed on the schizophrenia spectrum (Kwapil et al., 1997), Hypomania and Impulsivity-Nonconformity have recently been associated with bipolar disorder and major depressive episodes (Kwapil et al., 2000).
and the association of with substance use disorders confirmed (Kwapil, 1996). The report (Meyer & Hautzinger, 1999) that Magical Ideation is much more stable over two years than Perceptual Aberration warrants consideration of a developmental genetic perspective on the relationship between these two measures that are so often combined.

Psychosis-proneness is now being integrated into the neurodevelopmental hypothesis of schizophrenia. Changes in psychosis-proneness in early adulthood have been interpreted (Verdoux et al., 1998) as “a physiological developmental stage favouring the expression of psychosis proneness”, an argument strengthened by the recent associations with timing of puberty (Kaiser & Gruzelier, 1999) and with fluctuating asymmetry (Rosa et al., 2000).

The Australian Twin Registry and the increasing number of registries that have regular contact with twins are ideal for the study of psychosis-proneness, in addition to the more conventional measures of personality and temperament. The Australian twins were generally enrolled as children in the Twin Registry by their parents which may have reduced some of the usual biases (Eaves et al., 1989) seen when adult twins are recruited de novo. Such biases may be especially important in studies of behaviours associated with psychoses (Gottesman, 1991) and at longitudinal studies will provide the ultimate test of the validity of this short form of the psychosis-proneness scales.

Acknowledgments

This study was supported by a program grant to the NHMRC Schizophrenia Research Unit and by grants from the U.S. National Institute of Alcoholism and Alcohol Abuse (AA07535, AA07728) and Australian National Health and Medical Research Council (941177). The authors thank Ann Bruce, Joanne Caldwell, Ann Eldridge, Sue Mason, Theresa Pangan, John Pearson, Theresa Theobald, Luke Xantidis and Olivia Zheng for assistance and the twins for their cooperation.

References


