Psychological Medicine, 1999, **29**, 539–553. Printed in the United Kingdom
© 1999 Cambridge University Press

Fears and phobias: reliability and heritability

K. S. KENDLER,[1] L. M. KARKOWSKI AND C. A. PRESCOTT

From the Virginia Institute for Psychiatric and Behavioral Genetics and the Departments of Psychiatry and Human Genetics, Medical College of Virginia of Virginia Commonwealth University, VA, USA

ABSTRACT

Background. Familial factors, which are partly genetic, influence risk for phobias. Prior family and twin studies, however, were based on a single lifetime assessment, which may be only moderately reliable.

Methods. We obtained, 8 years apart, two assessments of lifetime history of five unreasonable fears and phobias (agoraphobia and social, situational, animal and blood-injury phobia) from face-to-face and telephone interviews from 1708 individual female twins from a population-based registry. We also obtained, 1 month apart, test–retest reliability on 192 twins. We fitted, using the program Mx, a measurement model that estimates the role of genetic and environmental risk factors correcting for measurement error.

Results. Short-term reliability of the five phobias was modest (mean $\kappa = 0{\cdot}46$), but higher than long-term stability (mean $\kappa = 0{\cdot}30$). Unreliability occurred both for subject recall of unreasonable fears and for interviewer assessment of which fears constituted phobias. Examining fears and phobias together, in a multiple threshold model, results suggested that twin resemblance was due solely to genetic factors, with estimated total heritabilities, corrected for unreliability, of: any 43%, agoraphobia 67%, animal 47%, blood/injury 59%, situational 46% and social 51%. With the exception of animal phobia, similar results were obtained analysing phobias alone.

Conclusions. Lifetime histories of unreasonable fears and phobias assessed at personal interview have substantial unreliability. Correcting for unreliability, the liability to fears and their associated phobias is moderately heritable. Individual-specific environmental experiences play an important role in the development of phobias, while familial–environmental factors appear to be of little aetiological significance.

INTRODUCTION

We have previously reported, from a population-based sample of female twins, that the liability to agoraphobia, social phobia and animal phobia was modestly influenced by genetic factors with heritabilities estimated between 30 and 40% (Kendler *et al.* 1992*a*). For situational and blood-injury phobia, by contrast, twin resemblance appeared to be due to familial–environmental risk factors, which were responsible for $\sim 30\%$ of the variance in liability (Kendler *et al.* 1992*a*; Neale *et al.* 1994*b*).

As with most recent epidemiological and genetic studies in psychiatry, we assessed the lifetime history of phobias by a single interview. In accord with studies of the accuracy of recall in the survey research literature (Wentland & Smith, 1993), previous investigations have found that, especially in non-clinical populations, the assessment of lifetime histories of psychiatric illness is not highly reliable (Bromet *et al.* 1986; Aneshensel *et al.* 1987; Prusoff *et al.* 1988; Fendrich *et al.* 1990; Kendler *et al.* 1993).

Unreliability of assessment has important implications for psychiatric genetic studies (Rice *et al.* 1992; Kendler, 1993). In family, twin or adoption studies utilizing a one-time assessment of lifetime history, the degree of familial resemblance or heritability will be attenuated, perhaps substantially, by error of measurement. Because reliability of recall is influenced by the 'memorability' or 'salience' of disorder-related experi-

* Address for correspondence: Professor K. S. Kendler, Virginia Institute for Psychiatric and Behavioral Genetics, PO Box 980126, Richmond, VA 23298-0126, USA.

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

ences (Tanur, 1992), it may be particularly low for disorders like phobias which, for most of those affected, are clinically mild.

We have now re-assessed the lifetime history of phobias in these female twins. In this report, we examine the short-term test–retest reliability and long-term stability of the diagnosis of fears and phobias and then explore, using a model that incorporates both assessments, the role of genetic and environmental risk factors in the aetiology of these five phobias correcting for the role of measurement error. We seek to address the following three questions.

1. How reliable is the lifetime history of phobias as assessed at psychiatric interview?

2. Correcting for measurement error, how heritable are the phobias? Using the increased statistical power from two times of measurement, do we find evidence for real differences in the causes of twin resemblance for agoraphobia, social phobia and animal phobia *v.* situational and blood-injury phobias as suggested by our earlier work?

3. How much of the previous evidence for a strong role of individual-specific environment in the aetiology of phobias is due to measurement error *v.* 'true' environmental experiences?

METHOD

Sample and assessment procedures

Our data were derived from a study of genetic and environmental risk factors for common psychiatric and substance use disorders in Caucasian female same-sex twin pairs from the Virginia Twin Registry (Kendler *et al.* 1992*b*) – a population-based register formed from a systematic review of all birth certificates in the Commonwealth of Virginia. Twins were initially ascertained through mailed surveys to female twin pairs in the Registry, the response to which was $\sim 64\%$ (1176 pairs). The true cooperation rate was undoubtedly higher as such large scale mailings always contain a significant percentage of questionnaires which do not reach the intended recipient. Twins were then interviewed face-to-face, at which time our refusal rate was $\sim 8\%$ (Kendler *et al.* 1992*b*) (2163 individual twins, mean age of 29·6). This initial interview included an assessment of agoraphobia and social, situational and animal phobia. We reported, in 1992, the results of model-fitting in these data (Kendler *et al.* 1992*a*). Attempts were then made to recontact all originally interviewed twins a minimum of 1 year after their initial interview. We succeeded in completing time 2 personal interviews in 92·5% of them (2002 individual twins, mean age 31·1). This second interview, completed largely by phone, contained a short module assessing blood-injury phobia. We reported on these data in 1994 (Neale *et al.* 1994*b*).

In the current phase of the project, 2293 members of female–female pairs from the Virginia Twin Registry were eligible to participate in a structured telephone interview. These twins were unselected except that they had participated in previous face-to-face interviews in this project. Of these 2293 twins, 1942 were successfully interviewed in 1995–7 (mean age 37·7), 3 had died, 33 were lost to follow-up, 4 were too medically ill to be interviewed, 58 neither refused participation nor completed an interview by the end of the study and 253 refused. However, two of these individuals terminated the interview prior to the section on phobias. Thus, we succeeded in interviewing 84·7% of the entire and 86·1% of the eligible sample. As outlined previously (Kendler *et al.* 1992*b*), zygosity was initially determined by a blind review of two experienced twin researchers that utilized standard questions (Eaves *et al.* 1989) and photographs. Blood samples were obtained from both members of 119 pairs of uncertain zygosity and analysed using eight RFLP markers (Spence *et al.* 1988). More recently, we performed polymerase chain reaction zygosity tests on an additional 269 twin pairs over-sampling those where our prior zygosity assignment was questionable. On the basis of these tests (where the mean number of markers tested per pair was 17·5 (S.D. = 8·4)), zygosity was changed in 12 pairs (4·5% of those tested).

All interviews were conducted blind to information about the co-twin. This project was approved by our local Institutional Review Board. Written informed consent was obtained prior to face-to-face interviews and verbal assent prior to phone interviews. Of the 1940 twins interviewed in this phase of the project with complete data on phobias, 1708 (88%) also participated in the first personal interview (containing our first phobia assessment) and

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

1648 participated in our second interview (that contained our assessment of blood-injury phobia).

The interviewers, each of whom had at least a Master's degree in Social Work or another mental health-related discipline or a Bachelor's degree in a similar area and 2 years of clinical experience, were initially trained for 40 hours and received regularly scheduled review sessions over the course of the study. All interviews with twins were conducted blind to knowledge about the psychiatric history of the co-twin.

As outlined previously (Kendler *et al.* 1992*a*), we assessed a lifetime history of phobias by an adaptation of the Phobic Disorders section of the Diagnostic Interview Schedule (DIS) Version III-A (Robins & Helzer, 1985), which in turn was based on the DSM-III criteria (American Psychiatric Association, 1980). While we added additional unreasonable fears in our most recent interview, for these analyses, we only analysed those fears assessed previously, which were: agoraphobia – unreasonable fear of going out of the house alone, being in crowds and being in open spaces; animal phobia – unreasonable fear of bugs, spiders, mice, snakes, bats or other animals; blood-injury phobia – unreasonable fear of sight of blood, needles or injections, dentists or hospitals or certain diseases, such as cancer or AIDS; situational phobia – unreasonable fear of tunnels, other closed places, bridges, airplanes and other high places; social phobia – unreasonable fear of meeting new people, giving a speech, using public bathrooms and eating in public. In addition to these 22 specific unreasonable fears, we also asked respondents whether there was 'Anything else you've been unreasonably terrified to do or be near?'. If any phobia described in response to this question best belonged with one of the five specific subtypes, it was so treated. Other phobias mentioned in response to this question (e.g. fear of water or darkness) were included in our analysis of 'any' phobia, but were not counted as belonging to one of the specific subtypes.

The major change in our assessment of phobias from the DIS was the criteria whereby an unreasonable fear was considered a phobia. For the DIS, one of three criteria had to be met: seeing a doctor about the fear; taking medications for the fear; or reporting that the fear or avoiding the fear 'interfered with life or activities a lot'. Given the low rates of treatment seeking for phobias (Kessler *et al.* 1998) and the range of extraneous factors influencing treatment seeking, we adopted a modification of the third criterion: objective behavioural impact of fear on respondent behaviour. Furthermore, in contrast to the DIS (developed for non-clinician interviewers), in which the respondent makes the judgement about fear-associated interference, in our interview, designed for clinicians, the interviewer made this assessment. Our interviewers were trained to rate whether the unreasonable fear and/or the avoidance of the feared object or situation had a significant objective behavioural impact on the respondent's life. If the respondent stated that they suffered substantial distress from the irrational fear but never changed their behaviour as a result of it, this was not scored as a phobia. Averaged across the two waves, the proportion of 'unreasonable' fears judged to meet criteria for phobia were: agoraphobic, 61%; animal, 29%; blood/injury, 31%; situational, 43% and social, 36%.

Phobia was defined in this report, then, as the presence of a fear which the respondent recognizes as unreasonable and that, in the judgement of the interviewer, objectively interferes with the respondent's life. It is noteworthy that this definition is different in several ways from that currently proposed in DSM-IV (American Psychiatric Association, 1994). For all analyses reported in this paper, a twin is considered affected for all subtypes of phobias for which she meets criteria.

We previously reported, in our wave 1 data, high levels of inter-rater reliability for our assessment both of the presence or absence of an unreasonable fear (agreement rate = 98·5%; kappa = $+0{\cdot}95 \pm 0{\cdot}02$ (Cohen, 1960)) and, when an unreasonable fear was present, the rating of the presence/absence of significant interference (agreement rate = 90·3%; kappa = $+0{\cdot}81 \pm 0{\cdot}07$) (Kendler *et al.* 1992*a*).

A methodological concern in twin studies is whether the equal environment assumption – that the exposure of monozygotic (MZ) and dizygotic (DZ) twins to environmental risk factors for unreasonable fears and phobias was equally correlated is correct. We have previously tested this assumption by attempting to predict twin concordance for fears or phobias, con-

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

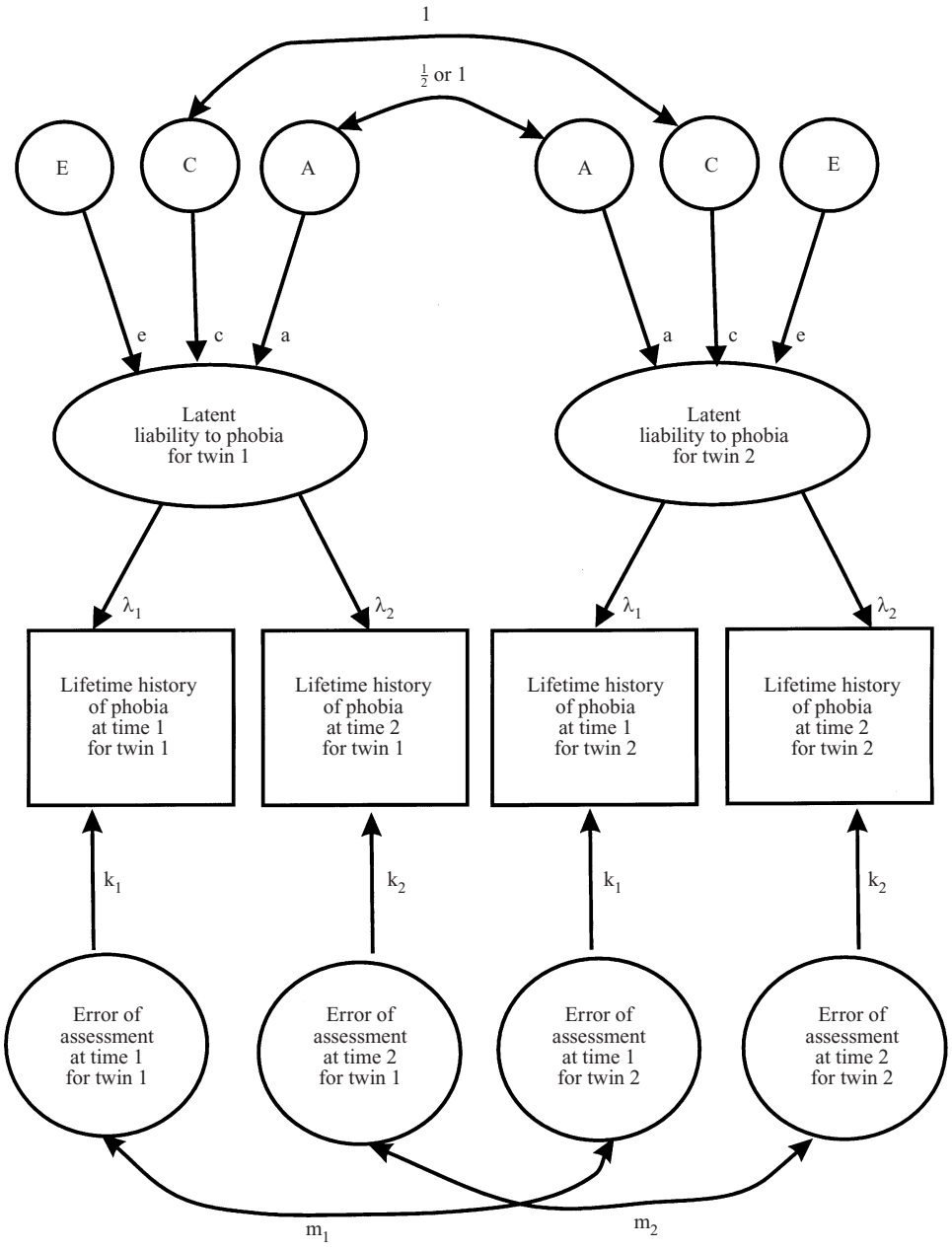


FIG. 1. A twin model for the liability to phobia including error of measurement. Observed variables are depicted in boxes and latent variables in circles and ellipses. This model assumes that individual twins have a true latent liability to phobia, which is imperfectly indexed by two assessments, at time 1 and time 2. The paths λ_1 and λ_2 represent the degree to which these two assessments, respectively, reflect the true level liability to phobia. The other path to the observed history of phobia (k_1 and k_2) represent occasion-specific effects (i.e. 'error') on the measurement obtained at time 1 and time 2, respectively. The model is constrained so that for each twin assessment, $\lambda^2 + k^2 = 1{\cdot}0$.

In addition, the model contains paths m_1 and m_2 which reflect the magnitude of the correlation of the occasion-specific effects on the two members of the twin pair at times 1 and 2, respectively.

The origin of the latent liability to phobia is modelled as in a standard biometrical twin design (Kendler, 1993), with the sources of variance divided between additive genetic (A), familial environmental (C) and individual-specific environmental factors (E). The correlation between familial environment is assumed to be 1 in both MZ and DZ twins. The correlation between additive genetic factors is assumed to be 1 in MZ and $\frac{1}{2}$ in DZ twins.

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

trolling for zygosity, by measures of the similarity of environmental experiences of the twins. Of the 10 such tests (childhood and adult environmental similarity for five phobias) with our wave 1 data (Kendler *et al.* 1992*a*; Neale *et al.* 1994*b*), only one was significant – a finding consistent with chance effects (Feild & Armenakis, 1974). We repeated these analyses separately for unreasonable fears and phobias for our five subtypes and for any fear/phobia using our wave 2 data and new assessments of the similarity of environmental experiences of twins that we have called 'childhood treatment', 'co-socialization' and 'similitude' (see Kendler & Gardner, Jr., 1998, for details of these measures). Of the 36 separate analyses, three were significant at the 5% level (a result not different from chance expectations (Feild & Armenakis, 1974)), only one of which was in the expected direction: increased co-socialization in teenage years predicting increased concordance for unreasonable social fears.

Statistical analysis

The twin measurement model utilized in this report has been described previously (Kendler *et al.* 1993; Kendler, 1997). Briefly, we assume that the variation in liability to phobias can be ascribed to four potential sets of factors: (*i*) additive genetic (A), which contribute twice as much to the correlation in MZ twins as DZ twins (because MZ twins share all their genes identical by descent, while DZ twins share on average half their genes); (*ii*) dominance genetic effects (D), which contribute four times as much to the correlation in MZ *v.* DZ twins (because MZ twins will always share both copies of a gene identical by descent while DZ twins will share both copies only one time in four, on average); (*iii*) family or 'common' environment (those familial factors such as parental attitudes which are shared by members of a twin pair) (C), which contributes equally to the correlation in MZ and DZ twins; and (*iv*) 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 reported history of phobias. We will call the impact of additive and dominance genetic effects together as total heritability, equivalent to the broad heritability of quantitative genetics (Falconer, 1989).

As outlined in Fig. 1, the twin measurement model utilizes the lifetime diagnoses of phobias of twins at both time 1 and time 2. The model assumes that each twin has a 'true' but unobserved (or 'latent') liability to phobias. We have measured the lifetime history of phobias twice in these twins, but each measurement is fallible, in that it partially reflects their true liability to phobias and partly reflects error. The paths λ_1 and λ_2 represent the degree to which the assessments of the lifetime history of phobias obtained at time 1 and time 2, respectively, reflect this true liability. The other path to assessments of phobias at the two time points (k_1 and k_2, respectively) represent error in the individual assessments. By definition, $\lambda^2 + k^2 = 1{\cdot}0$. To reiterate, error is defined in this model as any factors which influence the results of interview-based assessments other than the true latent liability to illness. The 'true' or latent liability to phobias is then modelled in a standard twin design, as outlined above, with the sources of variance in liability divided between additive genetic, dominance genetic, common environmental and individuals specific environmental factors.

While the standard measurement model assumes that any errors of measurement are uncorrelated within time, this assumption can be formally tested in our model by the addition of 'correlated error' or 'm' paths. Since it is possible that different degrees of correlated error were present in MZ and DZ twins and at time 1 and time 2, a full model would contain four m paths (where, for example, m_{1MZ}, represents the correlation in the errors of assessment for MZ twins at time 1): m_{1MZ}, m_{2MZ}, m_{1DZ}, and m_{2DZ}. For example, if new genetic risk factors for phobias were activated between the time 1 and time 2 assessments, this should be seen in positive values for the m path that were greater in MZ than in DZ twins.

The measurement model also assumes that the two interviews are assessing the same risk period. Given the early average age at onset of phobias in this sample (Kendler *et al.* 1992*a*), we assumed, as a first approximation, that all phobias reported at the time 2 interview began prior to the time 1 interview. We explore the validity of this assumption below. Three differences between a measurement model and the standard twin model, based on a single time

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

of assessment, are noteworthy. First, this model provides separate estimates for error of measurement (k) and true individual specific environment (e) that are confounded in the standard twin model. Secondly, it provides a direct estimate of the reliability of the phenotypic assessment support (λ^2). Thirdly, while a standard twin model estimates the heritability of the observed phenotype (including error), the measurement model estimates the heritability of the latent phenotype, correcting for the effects of error. The latter will always be greater than the former.

Because unreasonable fears are more common than phobias, an analysis that included fears would have considerably increased statistical power (Neale *et al*. 1994*a*). Such an analysis would only be appropriate if unreasonable fears and phobias were both manifestations of a single continuum of liability. This hypothesis is testable using a multiple-threshold model (Reich *et al*. 1972).

Model fitting is performed with the program Mx (Neale, 1994) using weighted least squares to fit to matrices of tetrachoric or polychoric correlations and the full weight matrix computed by PRELIS 2.14 (Joreskog & Sorbom, 1996). The best-fit model is selected using Akaike's information criterion (Akaike, 1987), where lower or more negative values indicates a model with a better balance of parsimony and explanatory power. The fit of nested models can also be compared by the χ^2 difference text.

RESULTS

Lifetime prevalences at times 1 and 2

Table 1 depicts the lifetime prevalences for the five subtypes of unreasonable fears and phobias. From the time 1 to the time 2 interview, the proportion of twins reporting unreasonable fears increased for two subtypes (animal and situational) and decreased for three subtypes (agoraphobia, blood/injury and social). However, for phobias, the lifetime prevalence rates were lower for all subtypes at the time 2 *v*. time 1 interview. The decline in lifetime prevalence was modest for animal, blood/injury and social phobia, moderate for situational phobia and substantial for agoraphobia.

Reliability and stability

The short-term test–retest reliabilities ($N = 192$, mean interval $\pm$ S.D. of $4{\cdot}3 \pm 1{\cdot}5$ weeks) for unreasonable fears are seen in Table 2. The test–retest reliabilities for the five classes of fears, measured by the κ coefficient and tetrachoric correlation (r), averaged $+0{\cdot}53$ and $+0{\cdot}77$, respectively.

Table 3 depicts the percent of subjects (total $N = 1708$) reporting no unreasonable fear, an unreasonable fear but no phobia or a phobia for each of the five subtypes at our time 1 and time 2 interview (mean interval of $92{\cdot}5 \pm 6{\cdot}9$ months). As seen in Table 2, the long-term stability of unreasonable fears was consistently lower than the test–retest reliability, with an average κ coefficient and tetrachoric correlation of $+0{\cdot}37$ and $+0{\cdot}59$, respectively. The reliability and stability of the individual classes of unreasonable fears were relatively similar to one another.

The short-term test–retest reliabilities for phobias are also seen in Table 2. For the five classes of phobias, the test–retest reliabilities, as measured by the κ coefficient and tetrachoric correlation, averaged $+0{\cdot}46$ and $+0{\cdot}77$, respectively. As with fears, the long-term stability of phobias was consistently lower, with an average κ coefficient and tetrachoric correlation of $+0{\cdot}30$ and $+0{\cdot}57$, respectively. The assessment of blood-injury phobia was less reliable than that of the other phobia subtypes.

To obtain a diagnosis of phobia, the respondent must report both a history of an unreasonable fear and describe objective interference due to that fear that are of sufficient significance to be so rated by the interviewer. To explore the reliability of that second stage, we examined – for individuals who admitted to an unreasonable fear on two occasions – the reliability of the assessment that the fear had sufficient practical impact on their lives to be judged a phobia. As can be seen from Table 2, both the short-term reliability (mean $\kappa = +0{\cdot}41$ and tetrachoric $= +0{\cdot}60$) and the long-term stability (mean $\kappa = +0{\cdot}23$ and tetrachoric $= +0{\cdot}35$) of this assessment was lower than that seen for either the assessment of fears or the diagnosis of phobias.

Fitting the measurement model to phobias

The results of fitting the model outlined in Fig. 1 to any phobia and to the five subtypes of phobias, assessed on two occasions, are outlined in Tables 4 and 5. We will describe in detail the

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

Table 1. *Lifetime prevalences for unreasonable fears and phobias at the two times of assessment*

	Time 1		Time 2	
Subtype	Unreasonable fear (%)	Phobia (%)	Unreasonable fear (%)	Phobia (%)
Agoraphobia	12·7	9·1	9·2	4·3
Animal	34·5	10·5	36·5	10·1
Blood/Injury	20·2	6·6	18·6	5·6
Situational	30·4	11·3	33·1	8·4
Social	27·5	14·1	26·6	13·9

Table 2. *Test–retest reliability and stability of reporting unreasonable fears, phobias, and phobias given the presence of fears*

	Short-term reliability*		Long-term stability†	
	κ	r	κ	r
Fears				
Agoraphobia	0·450	0·751	0·366	0·642
Animal	0·586	0·801	0·378	0·570
Blood/Injury	0·596	0·844	0·331	0·551
Situational	0·553	0·765	0·431	0·641
Social	0·470	0·697	0·368	0·536
Phobias				
Agoraphobia	0·515	0·843	0·302	0·611
Animal	0·493	0·783	0·289	0·554
Blood/Injury	0·283	0·644	0·243	0·526
Situational	0·523	0·798	0·281	0·514
Social	0·472	0·786	0·381	0·669
Phobias given fear‡				
Agoraphobia	NE (7)	NE	0·235 (81)	0·356
Animal	0·497 (52)	0·718	0·216 (363)	0·343
Blood/Injury	0·296 (19)	0·453	0·138 (145)	0·219
Situational	0·338 (42)	0·498	0·183 (332)	0·285
Social	0·520 (28)	0·729	0·380 (249)	0·558

* $N = 192$; mean inter-interview interval of 4·3 weeks.
† $N = 1708$; mean inter-interview interval of 7·7 years.
‡ Values in parentheses indicate sample size; NE = not estimable due to small sample size.
κ, kappa; r, tetrachoric correlation.

Table 3. *A comparison of subjects reporting no fear or phobia, fear only or phobia at time* 1 *and time* 2 *assessment*

		Time 2		
Phobia	Time 1	No fear/phobia	Fear	Phobia
Agoraphobia	No fear/phobia	82·4	2·9	1·4
	Fear	2·9	0·8	0·5
	Phobia	5·6	1·2	2·3
Animal	No fear/phobia	52·0	9·7	2·3
	Fear	14·4	7·9	2·8
	Phobia	4·8	3·2	3·0
Blood/Injury	No fear/phobia	72·7	6·2	1·7
	Fear	9·3	2·6	1·4
	Phobia	3·4	1·3	1·3
Situational	No fear/phobia	54·5	8·8	4·5
	Fear	7·2	6·9	4·0
	Phobia	4·6	4·2	5·3
Social	No fear/phobia	60·5	9·6	2·4
	Fear	8·8	5·7	1·6
	Phobia	4·1	2·9	4·3

results for any phobia. Model I assumes that the latent liability to any phobia is due to additive genetic and common and unique environmental factors. Furthermore, the degree to which each occasions of measurement reflected the true liability (assessed by the λ paths) or occasion-specific factors (e.g. 'error') (assessed by the k paths) and the degree to which these occasion-specific factors were correlated within times of measurement in MZ and DZ twins (assessed by the m paths) was left to be estimated by the model. For any phobia, this model fits quite well ($\chi^2 = 1{\cdot}3$, df $= 6$, $P = 0{\cdot}97$). Model II differed from model I only in that we constrained the correlated error terms to be the same in MZ and DZ twins. The AIC value for model II for any phobia was better (i.e. more negative) than for model I. Model III, then, sets all the m paths correlating the occasion-specific factors to zero, still assuming that twin resemblance results for both additive genetic and common environmental effects. For any phobia, model III produced a better AIC value than model II, indicating no evidence for occasion-specific factors. Assuming that the m paths are set to zero, Models IV to VI test the other three major possible explanations for twin resemblance: model IV, additive and dominance genetic effects; Model V, additive genetic effects only; and Model VI, common environmental effects only. For any phobia, model V produced the best AIC value. In addition, it is noteworthy that model VI could be rejected against model III by the rigorous χ^2 difference test ($\chi^2 = 8{\cdot}4$, df $= 1$, $P = 0{\cdot}004$). Assuming the best-fitting model from models III through VI, model VII then tests whether the two times of measurement were similarly accurate indices of the liability to phobias by constraining the two λ paths to be equal. For any phobias, the AIC for model VII improved from that found for model V, meaning

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

Table 4. *A description of the measurement models applied to phobias*

Model	Parameters estimated											
	A	C	D	E	λ_1	λ_2	k_1	k_2	m_{1MZ}	m_{2MZ}	m_{1DZ}	m_{2DZ}
I	f	f	0	f	f	f	f	f	f	f	f	f
II	f	f	0	f	f	f	f	f	f	f	m_{1MZ}	m_{2MZ}
III	f	f	0	f	f	f	f	f	0	0	0	0
IV	f	0	f	0	f	f	f	f	0	0	0	0
V	f	0	0	f	f	f	f	f	0	0	0	0
VI	0	f	0	f	f	f	f	f	0	0	0	0
VII	Best-fit models III–VI				f	λ_1	f	f	0	0	0	0

f, Parameter was estimated in the model.
0, Parameter was set to zero.
'specific parameter', Parameter was set to equal the one indicated.

that we could assume that the two times of measurement equally reflected the underlying liability to any phobia. Model VII was, overall, the best-fit model.

Examining the results of model fitting to all the phobias, four results are noteworthy. First, for any phobia and all five subtypes of phobia, we found no evidence for occasions specific twin resemblance in our assessments. That is, the m paths for all models could be set to zero, with an improvement in the AIC value. Secondly, for any phobia and for all phobia subtypes, the best-fitting model assumed that the two interviews assessed the true latent liability with equal reliability. That is, both the λ_1 and λ_2 and the k_1 and k_2 paths could be equated. Thirdly, for any phobia and for blood/injury, situational and social phobia, the best-fitting model was a very simple one which assumed that twin resemblance was due solely to additive genetic factors. For agoraphobia, the best-fit model postulated that twin resemblance was due to both additive and dominance genetic effects. For animal phobia, the best-fit model assumed that twin resemblance was due to familial–environmental factors. Fourthly, the estimated heritability of the latent liability to any phobia and to the three phobia subtypes with the best-fit AE model was moderate and (along with 95% CIs) was estimated as follows: any phobia 0·55 (0·41–0·71); blood/injury phobia 0·60 (0·34–0·91); situational phobia 0·50 (0·32–0·72); and, social phobia 0·50 (0·35–0·66). For agoraphobia, additive and dominance genetic effects were estimated at 0 (0·00–0·58) and 0·61 (0·00–0·82). (The estimate for additive genetic effects is so small and the standard errors are so large in the ADE model because of the high negative correlation for the estimates of additive and dominance genetic variance (Neale & Cardon, 1992)). For animal phobia, by contrast, the best-fit model predicted that twin resemblance was due to familial–environmental factors accounting for 34% of the variance in liability (16–50%).

Fitting to a multiple-threshold model

Given the low reliability with which unreasonable fears and phobias could be distinguished, we examined whether fears and phobias could be assumed to represent differing levels of severity on the same continuum of liability. Since we had four tests for any phobia and each phobia subtype (MZ and DZ twins each at time 1 and time 2), we performed 24 individual tests. Of these, three were significant at the 5% level, a result that does not differ significantly from chance expectation (Feild & Armenakis, 1974). We therefore assumed that unreasonable fears and phobias reflect varying points of severity on a single liability continuum.

The results of fitting the measurement model to the data on fears/phobias are outlined in Table 6. Four results are noteworthy. First, for any fear/phobia and all five subtypes of fear/phobia, we found no evidence for occasions specific twin resemblance in our assessments. The m paths for all models could be set to zero. Secondly, for any, 'agora', blood/injury and social fears/phobias, the best-fitting model assumed that the two interviews assessed the true latent liability with equal reliability. However, for animal and situational phobia, the best model indicated that the second phone interview was a better index of the latent liability to fears/phobias than the first face to face interview. Thirdly, for any fear/phobia, and for

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

Table 5. *Model fitting results for phobias*

	Model	I	II	III	IV	V	VI	VII	Parameter estimates of best-fitting model							
Phobia	df	6	8	10	10	11	11		a^2	c^2	d^2	e^2	λ_1	λ_2	k_1	k_2
Any	χ^2	1·31	2·44	3·32	2·96	3·32	11·56	3·53‡	0·55	—	—	0·45	0·73	0·73	0·69	0·69
	AIC	−10·7	−13·6	−16·7	−17·0	−18·7	−10·4	−20·5*								
Agoraphobia	χ^2	15·66	12·82	16·11	12·92	16·11	25·10	14·3†	0·00	—	0·61	0·39	0·85	0·85	0·53	0·53
	AIC	−3·7	−3·2	−3·9	−7·1	−5·9	3·1	−7·7*								
Animal	χ^2	6·70	7·47	7·47	9·54	8·15	7·63	7·87‡	—	0·34	—	0·66	0·75	0·75	0·66	0·66
	AIC	−5·3	−8·5	−12·5	−10·5	−13·9	−14·4	−16·1*								
Blood/Injury	χ^2	2·72	3·05	3·08	11·76	3·08	7·51	3·70‡	0·60	—	—	0·40	0·76	0·76	0·65	0·65
	AIC	−9·3	−13·0	16·9	−10·2	−18·9	−14·5	−20·3*								
Situational	χ^2	12·91	15·33	15·44	16·58	15·44	20·66	16·9‡	0·50	—	—	0·50	0·74	0·74	0·67	0·67
	AIC	0·9	−0·7	−4·6	−5·4	−6·6	1·4	−7·1*								
Social	χ^2	1·87	2·48	2·96	4·21	3·01	5·83	4·20‡	0·50	—	—	0·50	0·86	0·86	0·52	0·52
	AIC	−10·1	−13·5	−17·0	−17·8	−19·0	−16·2	−19·8*								

AIC, Akaike's Information Criteria.
* Best-fitting model by AIC.
—, Not estimated.
† df = 11, ‡ df = 12.

Table 6. *Multiple threshold model fitting results for unreasonable fears and phobias*

	Model	I	II	III	IV	V	VI	VII	Parameter estimates of best-fitting model							
Phobia	df	6	8	10	10	11	11		a^2	c^2	d^2	e^2	λ_1	λ_2	k_1	k_2
Any	χ^2	2·37	4·36	4·99	4·69	4·99	15·73	5·81‡	0·43	—	—	0·57	0·73	0·73	0·68	0·68
	AIC	−9·6	−11·6	−15·0	−15·3	−17·0	−6·3	−18·2*								
Agoraphobia	χ^2	14·79	17·94	18·40	11·59	18·40	37·99	11·6†	0·00	—	0·67	0·33	0·84	0·84	0·54	0·54
	AIC	2·8	1·9	−1·6	−8·5	−3·6	16·0	−10·4*								
Animal	χ^2	5·16	6·15	6·29	6·32	6·32	11·94	9·21‡	0·47	—	—	0·53	0·66	0·87	0·76	0·48
	AIC	−6·8	−9·9	−13·7	−13·7	−15·7*	−8·1	−12·8								
Blood/Injury	χ^2	7·21	7·75	7·77	6·75	7·77	17·18	7·77‡	0·59	—	—	0·41	0·74	0·74	0·68	0·68
	AIC	−4·8	−8·3	−12·2	−13·3	−14·2	−4·8	−16·2								
Situational	χ^2	10·78	11·11	12·21	9·71	12·21	25·16	13·5†	0·00	—	0·46	0·53	0·63	0·90	0·77	0·41
	AIC	−1·2	−4·9	−7·8	−10·3*	−9·8	3·2	−8·5								
Social	χ^2	5·43	5·69	8·55	10·35	10·35	10·37	11·6‡	0·51	—	—	0·48	0·78	0·78	0·63	0·63
	AIC	−6·6	−10·3	−11·5	−9·7	−11·7	−11·7	−12·4*								

AIC, Akaike's Information Criteria.
* Best-fitting model by AIC.
—, Not estimated.
† df = 11, ‡ df = 12.

animal, blood/injury and social fear/phobia, the best-fitting model assumed that twin resemblance was due entirely to additive genetic factors. Of note, we could not replicate the result obtained with phobias alone that twin resemblance for animal phobias was due to family environment. As expected, we had greater statistical power because of higher prevalences and more measurement categories. For example, the level of statistical confidence in rejecting the model that assumed that twin resemblance was due to familial–environmental factors by the χ^2 difference test (comparing models III and VI) was substantially greater than that seen when analysing phobias alone with the exception of social phobia: any fear/phobia $\chi^2 = 10{\cdot}7$, df = 1, $P = 0{\cdot}001$; 'agora' fear/phobia $\chi^2 = 19{\cdot}6$, df = 1, $P < 0{\cdot}0001$; animal fear/phobia $\chi^2 = 5{\cdot}3$, df = 1, $P = 0{\cdot}02$; blood/injury fear/phobia $\chi^2 = 9{\cdot}4$, df = 1, $P = 0{\cdot}002$; situational fear/phobia $\chi^2 = 13{\cdot}0$, df = 1, $P = 0{\cdot}0003$. With social fears/phobia, by contrast, model VI could not be rejected against model III ($\chi^2 = 1{\cdot}9$, df = 1, $P = 0{\cdot}17$) and the fits of models V and VI were indistinguishable. We fit model VII (constraining the λ paths to be the same) separately to the AE and CE models, and the former provided modestly better fit ($\chi^2 = 11{\cdot}63$, AIC $= -12{\cdot}37$ *v.* $\chi^2 = 12{\cdot}77$, AIC $= -11{\cdot}63$) and so is presented as the best-fit model VII. Fourthly, the estimated heritability of the latent liability to any fear/phobia and to the subtypes of fears/phobias were generally similar to that found for phobias alone, although the 95% CIs were tighter. For any fear/phobia and three of the subtypes, the AE model was the best fit with estimated heritabilities as follows: any 0·43 (0·32–0·54); animal 0·47 (0·34–0·59); blood/injury 0·59 (0·43–0·78); and, social 0·51 (0·39–0·64). The ADE model was the best fit for two fear/phobia subtypes: 'agora', additive 0 (0·00–0·41) and dominance 0·67 (0·24–0·63), and situational, additive 0 (0·00–0·51) and dominance 0·46 (0·00–0·61), respectively.

DISCUSSION

We sought to address three major questions about the reliability and heritability of unreasonable fears and phobias in a population-based sample of female twins. We will examine these issues in turn.

The reliability and stability of the assessment of unreasonable fears and phobias

Substantial evidence now suggests that the lifetime history of most psychiatric disorders, especially when assessed in non-clinical populations, have only moderate reliability (Bromet *et al.* 1986; Aneshensel *et al.* 1987; Prusoff *et al.* 1988; Fendrich *et al.* 1990; Kendler *et al.* 1993). To examine this question for unreasonable fears and phobias, we assessed both short-term test–retest reliability and long-term stability. As expressed by the κ coefficient, we found that the short-term reliability of the five phobias was moderate (mean $\kappa = 0{\cdot}46$), but long-term stability was even lower (mean $\kappa = 0{\cdot}30$). However, these results are not inconsistent with the few similar previous studies. For example, Canino *et al.* in a mixed sample of out-patients and community respondents, found that the one-day test–retest reliability for a diagnosis of simple phobia obtained by the DIS was $\kappa = 0{\cdot}50$ and 0·66 for social phobia (Canino *et al.* 1987*a*). In a summary of short-term (< 3 day) test–retest evaluations with the CIDI, in a mixture of in-patient, out-patient and community samples, Wittchen reported reliability for simple and social phobia of $\kappa = 0{\cdot}59$ and 0·64, respectively (Wittchen, 1994). In a small general population sample, Wacker *et al.* found that the test–retest reliability (as measured by Yule's Y) over 20 months was 0·68 for social phobia but only 0·22 for specific phobias (Wacker *et al.* 1992). In what may be the most comparable study, Prusoff *et al.* interviewed a mixed group of probands and their relatives twice over a mean of 4·7 years, using the SADS-L interview. The test–retest reliability for phobic disorder was $\kappa = 0{\cdot}32$ (Prusoff *et al.* 1988).

Long-term stability of phobias was lower than the short-term reliability. Since the unreliability due to different interviewers should be similar for both measures, these results suggest that the reliability of respondent reporting for fears and phobias declines over time. Some insight into the process of unreliability is gained by our findings of particularly low long-term stability for the assignment of a phobia diagnosis given the report of an unreasonable fear. We have previously shown high inter-rater reliability in these assessments ($\kappa = +0{\cdot}81$ (Kendler *et al.* 1992*a*)), suggesting that variation in respondent

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

reporting is likely to be responsible for most of the unreliability. We would conclude that, over the long-term people may be better able to remember the existence of an unreasonable fear than to recall the associated level of impairment.

Using the same assessment, we found little overall change in the rates of reported unreasonable fears. However, rates of phobias declined across all the subtypes, particularly for agoraphobia. This counter-intuitive finding, termed the 're-test' artefact, has been found in other studies and may be due to increased forgetting and/or learning to deny questions to decrease the duration of interviews (Jorm *et al.* 1989). That respondents 'forget' phobias as they age is suggested by analyses of the Epidemiologic Catchment Area Study showing decreasing lifetime prevalence rates of phobias with increasing age (Simon & VonKorff, 1992). In our data, most of the decline is due to a lower proportion of individuals reporting unreasonable fears who made criteria for phobia. Although we strived for consistency, our requirement for phobia could have become more stringent across waves. Alternatively, the respondents may have 'learned' to deny interference from the unreasonable fear as a way of speeding up the interview. We previously found that prevalence rates of phobias did not differ when assessed by phone *v.* face-to-face (Kendler *et al.* 1992*a*), so mode of assessment is unlikely to be responsible for these prevalence changes.

The heritability of phobias correcting for measurement error

The low reliability of the diagnosis of phobias suggests that twin studies based on a single diagnostic assessment may seriously underestimate the heritability of phobias. To address this question, we applied a measurement model to our data on phobias from two waves of interviews. As expected, we found a substantial increase in the heritability of phobias when measurement error was taken into account. Compared to our previous results obtained from our time 1 interviews (Kendler *et al.* 1992*a*), for example, total heritability increased for any phobia from 32 to 55%, for agoraphobia from 39 to 61% and for social phobia from 30 to 50%. Our findings of significant heritable influences on the liability to phobias is consistent with most previous twin studies of self-report irrational fears (Torgersen, 1979; Rose *et al.* 1981; Rose & Ditto, 1983; Neale & Fulker, 1984; Phillips *et al.* 1987) and family (Noyes *et al.* 1986; Fyer *et al.* 1990, 1993, 1995) and twin studies (Carey & Gottesman, 1981; Torgersen, 1983) of clinically diagnosed phobias.

As noted above, twin model-fitting applied to wave 1 of our data suggested that the familial aggregation for agoraphobia, social phobia and animal phobia was due to genetic factors while for situational and blood-injury phobia, the aggregation appeared due to the family environment (Kendler *et al.* 1992*a*; Neale *et al.* 1994*b*). We noted, however, in these publications, that these conclusions were reached with a low level of statistical confidence (being based on differences in models of only 1·5 and 0·7 χ^2 units, respectively (Kendler *et al.* 1992*a*; Neale *et al.* 1994*b*). In the model-fitting using data from both waves, we did not replicate our prior finding that family environment played a major aetiological role in either situational or blood-injury phobias. Instead, we found for animal phobias only, model fitting to both waves of data suggested an important role for familial–environmental factors. Since this result was replicated neither in our prior analysis of our time 1 interviews (Kendler *et al.* 1992*a*), nor in our combined analysis of animal fears and phobias, we explored our data to try to understand the origin of this anomalous result. In tetrachoric correlation matrices for animal phobia, MZ correlations substantially exceeded DZ correlations at both time 1 (+0·31 *v.* +0·07) and at time 2 (+0·23 *v.* +0·06), as would be expected for a disorder where twin resemblance was due solely to genetic factors. However, an inexplicably high correlation of +0·39 was seen in DZ twins between the diagnosis of animal phobia in twin 1 at time 1 and in twin 2 at time 2. This 'cross-twin/cross-time' correlation was much higher than either those seen in MZ twins (+0·07 and +0·18) or the other DZ cross-twin cross-time correlation (i.e. twin 2 at time 1 and twin 1 and time 2 +0·09). It is likely that this single high cross-twin/cross-time DZ correlation was responsible for the apparent evidence for family environment. Since no known mechanism could explain a higher cross-twin/cross-time than cross-twin/within-time correlation, and the MZ and DZ cross-twin within-time correlations

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

at time 1 and time 2 clearly indicated genetic effects, we would regard our evidence for family environment in the aetiology animal phobia to be highly questionable.

Aside from these unusual results with animal phobias, we consistently found twin resemblance to be due to shared genetic rather than shared environmental factors. Given the low power of twin studies for the detection of dominance genetic effects (Neale *et al*. 1994*a*), the high negative correlation for the estimates of additive and dominance genetic variance (Neale & Cardon, 1992) and the unreliability of assigning genetic variance to dominance and additive components even with very large samples (Kendler *et al*. 1994), we do not attach much significance to our evidence for dominance genetic variance for agoraphobia and for situational fears/phobias.

The importance of environmental risk factors in the aetiology of phobias

Even when correcting for measurement error, we found that individual-specific environment appeared to play an important aetiological role in phobias. It is useful to compare the present findings with those that we previously reported for major depression. Correcting for measurement error using the same model applied here to phobias, we found that the heritability of liability to major depression was estimated at 71% with individual specific environment accounting for the remaining 29% of the variation in liability (Kendler *et al*. 1993).

By contrast, for the phobias examined here, estimates for the role of 'true' individual-specific environment ranged from 40 to 60%, substantially higher than that seen with major depression. These findings indicate that, even after correcting for the effects of measurement error, environmental experiences play a major aetiological role in the development of phobias. These 'phobia-genic' experiences are, apparently, rarely shared with the co-twin.

Implications

In psychiatric genetic studies utilizing lifetime diagnoses, the reliance on a single assessment is likely to result in a substantial underestimation of heritability and/or familial–environmental effects and an overestimation of the aetiological role of the individual-specific environment. When accounting for errors of measurement, the liability to any fear/phobia and to 'agora', animal, blood/injury, situational and social fears/phobias are approximately 40–60% due to genetic factors.

Our results are inconsistent with two major aetiological theories of phobia acquisition: social learning and classical conditioning. Social learning theory predicts that phobias are acquired from the observation of fear responses of others (Bandura, 1986). Since most phobias begin in childhood (Burke *et al*. 1990), other family members, particularly parents and co-twins, are likely models for 'vicarious' phobia acquisition. Classical condition, by contrast, predicts that phobias will arise as a result of the accidental pairing of benign and fear-inducing stimuli (Watson & Rayner, 1920; Rachman, 1977; Marks, 1987). Social learning theory predicts that familial–environmental factors should play a major aetiological role in phobias. Conditioning theory predicts that phobias should be entirely environmental in origin. Most of these environmental experiences are likely to be random in nature and hence emerge as individual-specific environment. However, exposure to fear-inducing stimuli might be correlated in family members hence predicting some familial–environmental component of liability to phobias.

By contrast, our results are supportive of an 'inherited phobia-proneness' model (Seligman, 1971; Gray, 1982), which suggests that, through natural selection, man has evolved an inherited predisposition to form phobic reactions to certain stimuli. Such phobia-proneness might have evolved via stabilizing selection (Hartl, 1980) in which a moderate level of phobia-proneness was maximally fit and fitness declined when individual levels of phobia-proneness were either too high or too low. Substantial additive genetic variation can emerge in populations under such evolutionary models (Mather, 1966; Lande, 1976; Hartl, 1980). Because environmental exposure to the specific stimuli would be needed to 'activate' the inherited phobia-proneness, this model predicts – consistent with our observations – that both genetic and environmental factors would be needed to produce clinical phobias.

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

Limitations

Our results should be interpreted in the context of three potentially significant limitations. First, these results apply only to Caucasian females. Although no studies have examined the sex-specific transmission of phobias, the substantial gender difference in prevalence (Eaton *et al.* 1991; Kessler *et al.* 1994) suggests that genetic and environmental risk factors may differ in their impact on phobias in males *v.* females.

Secondly, the lifetime prevalence rates for phobias in this sample are toward the upper end of the range of those previously reported in women (Canino *et al.* 1987*b*; Bland *et al.* 1988; Wells *et al.* 1989; Eaton *et al.* 1991). Whether this is due to differences in population or assessment is not clear. Our interviewer-based assessment of impairment, while more objective, may produce a lower 'threshold' for interference than the self-report measure used in the DIS. Our definition may also be broader than that proposed in DSM-III-R criteria (American Psychiatric Association, 1987), in that we require objective 'impact' on respondent behaviour rather than the 'significant interference with normal routine' in DSM-III-R (American Psychiatric Association, 1987). However, DSM-III-R would include an unreasonable fear that produced 'marked distress', while we did not. However, the National Comorbidity Survey, based on DSM-III-R criteria, produced lifetime prevalence rates in women of social phobia and agoraphobia (the two comparable phobias categories across studies) that were quite similar to those found in our population (Kessler *et al.* 1994).

Thirdly, of the twins who were interviewed at time 2 and met criteria for one of our five phobia subtypes, 10% reported that the age at onset of that phobia was after their time 1 interview. Since the measurement model assumes that the two interviews are assessing the same risk period, it can be argued that for model-fitting purposes, these 'new onset' cases should be considered unaffected. When this is done, the best-fit phobia model was the same as that found previously in all five subtypes and the parameter estimates were quite similar. Excluding all such cases, however, is almost certainly too conservative as 44% of them reported, at their time 1 interview, that they had the same unreasonable fear or phobia that they reported at time 2 but said had not begun till after the time 1 interview.

This work was supported by grants MH-40828, MH-54150, MH-01277 and AA-09095 from the United States National Institutes of Health. The Virginia Twin Registry, established and maintained by Drs W. Nance and L. Corey, is supported by the United States National Institutes of Health grants HD-26746 and NS-31564. Data were collected under the direction of Lisa Halberstadt and Barbara Brooke.

REFERENCES

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika* **52**, 317–332.

American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders*, *3rd edn*. American Psychiatric Association: Washington, DC.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders*, *3rd edn. Revised*. American Psychiatric Association: Washington, DC.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*, *4th edn*. American Psychiatric Association: Washington, DC.

Aneshensel, C. S., Estrada, A. L., Hansell, M. J. & Clark, V. A. (1987). Social psychological aspects of reporting behavior: lifetime depressive episode reports. *Journal of Health and Social Behavior* **28**, 232–246.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall: Englewood Cliffs, NJ.

Bland, R. C., Orn, H. & Newman, S. C. (1988). Lifetime prevalence of psychiatric disorders in Edmonton. *Acta Psychiatrica Scandinavica* **77** (suppl. 338), 24–32.

Bromet, E. J., Dunn, L. O., Connell, M. M., Dew, M. A. & Schulberg, H. C. (1986). Long-term reliability of diagnosing lifetime major depression in a community sample. *Archives of General Psychiatry* **43**, 435–440.

Burke, C. B., Burke, J. D., Regier, D. A. & Rae, D. S. (1990). Age at onset of selected mental disorders in five community populations. *Archives of General Psychiatry* **47**, 511–518.

Canino, G. J., Bird, H. R., Shrout, P. E., Rubio-Stipec, M., Bravo, M., Martinez, R., Sesman, M., Guzman, A., Guevara, L. M. & Costas, H. (1987*a*). The Spanish diagnostic interview schedule: reliability and concordance with clinical diagnoses in Puerto Rico. *Archives of General Psychiatry* **44**, 720–726.

Canino, G. J., Bird, H. R., Shrout, P. E., Rubio-Stipec, M., Bravo, M., Martinez, R., Sesman, M. & Guevara, L. M. (1987*b*). The prevalence of specific psychiatric disorders in Puerto Rico. *Archives of General Psychiatry* **44**, 727–735.

Carey, G. & Gottesman, I. I. (1981). Twin and family studies of anxiety, phobic and obsessive disorders. In *Anxiety: New Research and Changing Concepts* (ed. D. F. Klein and J. G. Rabkin), pp. 117–136. Raven Press: New York.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20**, 37–46.

Eaton, W. W., Dryman, A. & Weissman, M. M. (1991). Panic and phobia. In *Psychiatric Disorders in America: The Epidemiologic Catchment Area Study* (ed. L. N. Robins and D. A. Regier), pp. 155–179. Free Press: New York.

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

Eaves, L. J., Eysenck, H. J., Martin, N. G., Jardine, R., Heath, A. C., Feingold, L., Young, P. A. & Kendler, K. S. (1989). *Genes, Culture and Personality: An Empirical Approach*. Oxford University Press: London.

Falconer, D. S. (1989). *Introduction to Quantitative Genetics, 3rd edn*. Wiley: New York.

Feild, H. S. & Armenakis, A. A. (1974). On use of multiple tests of significance in psychological research. *Psychological Reports* **35**, 427–431.

Fendrich, M., Weissman, M. M., Warner, V. & Mufson, L. (1990). Two-year recall of lifetime diagnoses in offspring at high and low risk for major depression: the stability of offspring reports. *Archives of General Psychiatry* **47**, 1121–1127.

Fyer, A. J., Mannuzza, S., Gallops, M. S., Martin, L. Y., Aaronson, C., Gorman, J. M., Liebowitz, M. R. & Klein, D. F. (1990). Familial transmission of simple phobias and fears: A preliminary report. *Archives of General Psychiatry* **47**, 252–256.

Fyer, A. J., Mannuzza, S., Chapman, T. F., Liebowitz, M. R. & Klein, D. F. (1993). A direct interview family study of social phobia. *Archives of General Psychiatry* **50**, 286–293.

Fyer, A. J., Mannuzza, S., Chapman, T. F., Martin, L. Y. & Klein, D. F. (1995). Specificity in familial aggregation of phobic disorders. *Archives of General Psychiatry* **52**, 564–573.

Gray, J. A. (1982). *The Neuropsychology of Anxiety*. Oxford University Press: New York.

Hartl, D. L. (1980). *Principles of Population Genetics*. Sinauer: Sunderland, MA.

Joreskog, K. G. & Sorbom, D. (1996). *PRELIS 2: User's Reference Guide*. Scientific Software International: Chicago, IL.

Jorm, A. F., Duncan-Jones, P. & Scott, R. (1989). An analysis of the re-test artefact in longitudinal studies of psychiatric symptoms and personality. *Psychological Medicine* **19**, 487–493.

Kendler, K. S. (1993). Twin studies of psychiatric illness: current status and future directions. *Archives of General Psychiatry* **50**, 905–915.

Kendler, K. S. (1997). Social support: a genetic-epidemiologic analysis. *American Journal of Psychiatry* **154**, 1398–1404.

Kendler, K. S. & Gardner, C. O., Jr. (1998). Twin studies of adult psychiatric and substance dependence disorders: are they biased by differences in the environmental experiences of monozygotic and dizygotic twins in childhood and adolescence? *Psychological Medicine* **28**, 625–633.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1992*a*). The genetic epidemiology of phobias in women: the inter-relationship of agoraphobia, social phobia, situational phobia and simple phobia. *Archives of General Psychiatry* **49**, 273–281.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1992*b*). A population-based twin study of major depression in women: the impact of varying definitions of illness. *Archives of General Psychiatry* **49**, 257–266.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1993). The lifetime history of major depression in women: reliability of diagnosis and heritability. *Archives of General Psychiatry* **50**, 863–870.

Kendler, K. S., Walters, E. E., Truett, K. R., Heath, A. C., Neale, M. C., Martin, N. G. & Eaves, L. J. (1994). The sources of individual differences in depressive symptoms: an analysis of two twin-family samples. *American Journal of Psychiatry* **151**, 1605–1614.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H.-U. & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Archives of General Psychiatry* **51**, 8–10.

Kessler, R. C., Olfson, M. & Berglund, P. A. (1998). Patterns and predictors of treatment contact after first onset of psychiatric disorders. *American Journal of Psychiatry* **155**, 62–69.

Lande, R. (1976). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetic Research at Cambridge* **26**, 221–235.

Marks, I. M. (1987). *Fears, Phobias, and Rituals*. Oxford University Press: New York.

Mather, K. (1966). Variability and selection. *Proceedings of the Royal Society, Series B* **164**, 328–340.

Neale, M. C. (1994). *Mx: Statistical Modeling, 2nd edn*. Box 980126, Department of Psychiatry: Richmond, VA 23298-0126.

Neale, M. C. & Cardon, L. R. (1992). *Methodology for Genetic Studies of Twins and Families*. Kluwer Academic Publishers B.V.: Dordrecht, The Netherlands.

Neale, M. C. & Fulker, D. W. (1984). A bivariate path analysis of fear data on twins and their parents. *Acta Genetica Medica et Gemellologia* **33**, 273–286.

Neale, M. C., Eaves, L. J. & Kendler, K. S. (1994*a*). The power of the classical twin study to resolve variation in threshold traits. *Behavior Genetics* **24**, 239–258.

Neale, M. C., Walters, E. E., Eaves, L. J., Kessler, R. C., Heath, A. C. & Kendler, K. S. (1994*b*). The genetics of blood-injury fears and phobias: a population-based twin study. *American Journal of Medical Genetics* (*Neuropsychiatric Genetics*) **54**, 326–334.

Noyes, R., Crowe, R. R., Harris, E. L., Hamra, B. J., McChesney, C. M. & Chaudhry, D. R. (1986). Relationship between panic disorder and agoraphobia. *Archives of General Psychiatry* **43**, 227–232.

Phillips, K., Fulker, D. W. & Rose, R. J. (1987). Path analysis of seven fear factors in adult twin and sibling pairs and their parents. *Genetic Epidemiology* **4**, 345–355.

Prusoff, B. A., Merikangas, K. R. & Weissman, M. M. (1988). Lifetime prevalence and age of onset of psychiatric disorders: recall 4 years later. *Journal of Psychiatric Research* **22**, 107–117.

Rachman, S. (1977). The conditioning theory of fear-acquisition: a critical examination. *Behaviour Research and Therapy* **15**, 375–387.

Reich, T., James, J. W. & Morris, C. A. (1972). The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. *Annals of Human Genetics* **36**, 163–184.

Rice, J. P., Rochberg, M., Endicott, J., Lavori, P. W. & Miller, C. (1992). Stability of psychiatric diagnoses: an application to the affective disorders. *Archives of General Psychiatry* **49**, 824–830.

Robins, L. N. & Helzer, J. E. (1985). *Diagnostic Interview Schedule* (*DIS*)*: Version III-A*. Washington University School of Medicine: St. Louis, MO.

Rose, R. J. & Ditto, W. B. (1983). A developmental-genetic analysis of common fears from early adolescence to early adulthood. *Child Development* **54**, 361–368.

Rose, R. J., Miller, J. Z., Pogue-Geile, M. F. & Cardwell, G. F. (1981). Twin-family studies of common fears and phobias . In *Twin Research 3: Intelligence, Personality, and Development* (ed. L. Gedda, P. Parisi and W. E. Nance), pp. 169–174. Alan R. Liss, Inc.: New York.

Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy* **2**, 307–320.

Simon, G. E. & VonKorff, M. (1992). Reevaluation of secular trends in depression rates. *American Journal of Epidemiology* **135**, 1411–1422.

Spence, J. E., Corey, L. A., Nance, W. E., Marazita, M. L., Kendler, K. S. & Schieken, R. M. (1988). Molecular analysis of twin zygosity using VNTR DNA probes. *American Journal of Human Genetics* **43**, A159 (Abstract).

Tanur, J. M. E. (1992). *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*. Russell Sage Foundation: New York.

Torgersen, S. (1979). The nature and origin of common phobic fears. *British Journal of Psychiatry* **134**, 343–351.

Torgersen, S. (1983). Genetic factors in anxiety disorders. *Archives of General Psychiatry* **40**, 1085–1089.

Wacker, H. R., Müllejans, R., Klein, K. H. & Battegay, R. (1992). Identification of cases of anxiety disorders and affective disorders in the community according to ICD-10 and DSM-III-R by using the Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research* **2**, 91–100.

Watson, J. B. & Rayner, B. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology* **3**, 1–14.

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press

Wells, J. S., Bushnell, J. A., Hornblow, A. R., Joyce, P. R. & Oakley-Browne, M. A. (1989). Christchurch psychiatric epidemiology study, part I: methodology and lifetime prevalence for specific psychiatric disorders. *Australian and New Zealand Journal of Psychiatry* **23**, 315–326.

Wentland, E. J. & Smith, K. W. (1993). *Survey Responses: An Evaluation of Their Validity*. Academic Press, Inc.: San Diego.

Wittchen, H.-U. (1994). Reliability and validity studies of the WHO Composite International Diagnostic Interview (CIDI): a critical review. *Journal of Psychiatric Research* **28**, 57–84.

https://doi.org/10.1017/S0033291799008429 Published online by Cambridge University Press