Assessment of abnormal personality in childhood: a Delphi survey of questionnaire data

Tamara Elliott, BSc, MB[1]

Peter Tyrer, MD, FMedSci[1]

John Horwood, MSc[2]

David Fergusson, PhD[2]

[1]Centre for Mental Health, Imperial College, London, England

[2]Christchurch Health and Development Study, University of Otago, Christchurch School of Medicine and Health Sciences, Christchurch, New Zealand

Address correspondence to Peter Tyrer, Centre for Mental Health, Imperial College, St. Dunstan's Road, London W6 8RP, UK; E-mail: p.tyrer@imperial.ac.uk

ABSTRACT

As an investigation into the feasibility of recording personality status from questionnaire data in younger people we used a three phase Delphi survey to assess items from the Christchurch Health and Development Study, administered at ages 12 to 16 years. Twelve experts took part in Phase I, and 22 in Phases II and III, 16 of whom were experts in adult personality disorder (PD), and 6 were experts who work with children. In total, 189 questions (55% of the total (238) in the questionnaires) were identified as possibly being related to personality abnormality in one or more clusters with high consensus. Experts who work with children were less likely to label features as related to personality than experts in PD (p<0.001), and the four personality factors (equivalent to Mulder and Joyce's antisocial, asocial, asthenic and anankastic) chosen for assessment showed variable agreement. Confirmatory factor analysis showed the best fitting model of the data was a 3 factor solution involving asocial/asthenic, antisocial and anankastic factors. This represents the first attempt to use existing recorded data to code personality status and the results of this Delphi survey give some grounds for optimism that this approach has potential in the early identification of personality features.

Personality disorder is formally identified only in late adolescence or early adulthood and is not currently admitted as an official diagnosis before the age of 18 years. However, there is accumulating evidence to suggest that personality abnormality, possibly amounting to disorder, is a significant form of psychopathology in adolescence (Johnson et al. 1999, Kasen et al. 1999, Levy et al. 1999, Westen et al., 2005), and could be an important factor when considering treatment for co-morbid Axis I disorders. Adolescents with PD are at a greater risk of Axis I disorders and suicidal behaviour, whether or not these are present in adolescence (Johnson et al., 1999, Levy et al., 1999), and also for subsequent personality disorders (Skodol et al, 2007).

Nevertheless, the concept of PD in childhood and early adolescence remains understandably controversial (Paris, 2003). As well as the stigma attached to a diagnosis of PD, there is concern about making a diagnosis while personality is still in development and constantly fluctuating; any diagnosis in adolescents may reflect a point-in-time disturbance rather than a chronic disorder (Bernstein et al., 1993). Major adverse experiences in childhood may lead to personality disorders both in childhood and adult life (Coid et al, 2006; Vizard et al, 2007). Presence of more than one risk factor in those predisposed to PD (e.g., poor family environment, genetic loading) increases the chance of progression, but this can be offset by protective factors, such as having a positive social orientation, or good parent-child relationships (Coid, 2003).

Because of the uncertainty over the dating of onset of personality pathology there are few assessment procedures designed for adolescent personality pathology, and it is uncertain to what extent the diagnosis represented on Axis II is appropriate (Westen & Arkowitz-Westen, 1998). Currently, PD is assessed by various structured interview schedules tailored to DSM-IV and ICD-10 classifications. However, these often contradict the diagnosis of clinicians (Tyrer & Simonsen, 2003). Since personality is such a complex variable, at present there appears to be no satisfactory quick assessment instrument (Tyrer, 2004). The Axis II diagnostic criteria of DSM-IV classify symptoms and PD as present or absent, when most of these variables are continuous with no dividing line

between normal and abnormal personality, so a dimensional diagnosis may be more appropriate (Tyrer & Johnson, 1996). Axis II also ignores personality patterns not severe enough to warrant a diagnosis, but these appear to be as least as prevalent as diagnosable PDs (Westen & Arkowitz-Westen, 1998). Despite this, Durrett and Westen (2005) proposed that the structure of personality pathology as assessed by Axis II criteria in adolescents is similar to that outlined in Axis II for adults, suggesting that PDs can be assessed similarly in both adolescents and adults.

In the present study we aimed to investigate whether questions from the Christchurch Health and Development Study (CHDS), administered at ages 12 to 16 years, could be consistently identified as being associated with personality abnormality using the Delphi technique. Questions thus selected were then subjected to a confirmatory factor analysis. This combination of methodologies was hoped to provide a separate data base that could serve as a proxy for personality status in the longitudinal CHDS data set.

METHODS

We used the Delphi technique to determine which questions asked in the CHDS between the ages of 12 and 16 might relate to PD. The CHDS is a birth cohort comprising 1265 individuals born in the Christchurch region in New Zealand between 15th April and 5th August 1977. During this period, there were 1310 live births, of whom 1265 (97%) subsequently participated. Participants were assessed at birth, 4 months, 1 year, annual intervals to age 16, 18 and 21 years. Data were obtained in the following ways: parental interviews (birth to 16 years), teacher questionnaires (six to 13 years), child and young person interviews (eight to 21 years) and police record data (14 to 21 years). Information includes family change and stability, exposure to child abuse and family dysfunction, educational achievement, behavioural adjustment at school, peer affiliations and relationships, mental health and psychosocial adjustment in adolescence and young adulthood, and participation in tertiary education and the workforce. The CHDS is not specifically designed to measure PD, but because of the focus on mental health, many relevant data have been collected.

The Delphi method aims to gain consensus from a group of experts. The judgements of the individuals are aggregated in order to improve the quality of responses. The rationale is that when there are no established data or rules on a topic, no one person can provide an answer, and no one expert is more reliable than another, so a combination of opinions should give an answer close to the true answer (Sharkey & Sharples, 2001; Dalkey & Helmer, 1963). The Delphi survey involves a series of questionnaires; each stage builds on the results of the previous one. Participants are allowed to change their opinions in subsequent rounds based on feedback provided by the researchers.

The present Delphi survey consisted of three stages. Phase I was designed to eliminate questions clearly unrelated to personality abnormality. Phase II asked participants to rate the degree to which questions retained from Phase I related to personality abnormality, and Phase III gave participants the opportunity to change their answers based on feedback (Figure 1).

There is no obvious one group of experts who could be consulted when investigating PD in children; there are experts with a specialist interest in PD, and others with expertise in normal and abnormal child behaviour. Initially questionnaires were sent to 157 people, including members of the British and Irish Group for the Study of Personality Disorders (BIGSPD)(http ://www.bigspd.net) who have a specialist interest in PD, and experts working with children (child and adolescent psychiatrists and educational psychologists). Twelve (8%) agreed to take part in Phase I. A further 20 individuals were invited to take part in Phases II and III (32 people in total), of whom 22 (67%) agreed to do so (10 of whom had also completed Phase I). Of the 12 who took part in Phase I, nine were experts in PD, and 3 worked with children; of the 22 people who took part in Phases II and III, 16 were experts in PD and six worked with children.

The Phase I questionnaire consisted of all 345 questions in the CHDS administered between the ages of 12 and 16 considered by the authors to be possibly related to personality abnormality, including questions from the parent, teacher and young person questionnaires. The questions were not consistent in the 12, 13, 14, 15 and 16 year questionnaires, and similar questions were asked to parents, teachers and young people. For this reason similar questions were merged; for example, the 12 year parent questionnaire includes "Child is inattentive or easily distracted", while the 15 and 16 year parent questionnaires include "Child is inattentive to what others say" and "Child is distractible, easily diverted from the task at hand". These were merged as "Child is inattentive or easily distracted from the task at hand" in our questionnaire.

The participants were asked to rate each question on a scale from 0-2, where 0 = no suggestion of personality abnormality; 1 = suggestion of personality abnormality if the symptom is present on at least two occasions between 12 and 16 (i.e. shows evidence of persistence), and 2 = likely personality abnormality if the symptom is present on any occasion. Questions with a mean score < 0.50 (i.e., over half the participants gave a score of 0) were thought not to bear a relationship to PD and were eliminated at the end of this

round. A low threshold was used as the purpose at this stage was to exclude questions clearly unrelated to personality abnormality.

The Phase II questionnaire comprised 238 questions that remained after Phase I. This time participants were asked to rate questions by cluster, based on the DSM-IV classifications. DSM-IV classifies PD into ten categories, which are grouped into three clusters (American Psychiatric Association, 1994). However, there is a strong case for a four cluster classification system fitting in with the concept of four clusters (Tyrer & Alexander, 1979; Tyrer et al, 2007), and summarised by Mulder & Joyce (1997) as the four 'A's' – antisocial, asocial, asthenic and anankastic. In the present study, we therefore split Cluster C into C1 (anxious/fearful – dependent and avoidant) and C2 (obsessive/rigid – obsessive-compulsive). For each of these four clusters, participants were asked to rate questions from 1-5, where 1 = disagree strongly (no relation to personality abnormality), 2 = disagree, 3 = neutral, 4 = agree, and 5 = agree strongly (highly related to personality abnormality).

In Phase III participants were invited to re-rate their answers to the previous phase. They were provided with their original answer, the group median response, and the level of consensus (high, medium or low), calculated using the inter-quartile ranges (IQRs). Questions that reached high consensus had IQR ≤ 1.5, medium consensus IQR $\leq$ 2.5, and low consensus IQR > 2.5. Participants were told that the aim was to reach high consensus on every question. Questions that achieved high consensus in all four clusters after Phase II were not further tested in this phase (21 questions).

Means, medians, quartiles, IQRs, and percentages for the different levels of agreement were calculated.

Confirmatory factor analysis

To test the validity of the four cluster solution identified by the DELPHI process a confirmatory factor analysis was conducted using items from the CHDS database

assessed when study participants were aged 15 years. The analysis was limited to a subset of 50 items (9 items Cluster A, 22 items Cluster B, 14 items Cluster C1, 5 items Cluster C2) assessed using parental report. These items were selected on the basis that: (a) there was very high inter-rater agreement on the placement of the item in a specified cluster; and (b) the item had sufficient base rate for analysis. Model fitting was conducted using *Mplus* (Muthen & Muthen, 1998-2007). Model goodness of fit was assessed on the basis of the Satorra-Bentler chi square; the Comparative Fit Index (CFI); the Root Mean Squared Error of Approximation (RMSEA); and the Standardised Root Mean Squared Residual Correlation (SRMR). Values of CFI greater than .90, RMSEA less than .05, and SRMR close to zero are generally thought to be indicative of a well-fitting model.

RESULTS

In Phase I, median score for 122 (35%) questions was zero; for 21 (6%) questions, median score was 0.5; for 195 (57%) questions median score was 1; for two questions, median score was 1.5, and for five questions median score was 2. At the end of Phase I, 107 questions were eliminated. In Phase II, each of the remaining 238 questions were split into four parts (one for each cluster) totalling 952 parts. At this stage 515 (54%) parts reached high consensus (Table 1). High consensus was reached for 21 questions in Phase II, and therefore these were not reassessed in Phase III. Including these 21 questions, in Phase III, 691 (73%) parts reached high consensus, giving a difference of 176 (18.5%) parts compared with Phase II (Table 1).

In Phase III, 189 (84%) questions achieved a median score of > 3 at high consensus in one or more cluster (i.e. more than half the participants rated the statement as having some relationship to PD), of which eight were thought to be related to more than one cluster. Figure 2 shows the distribution of these questions by cluster (22 in A, 115 in B, 54 in C1 and 6 in C2), as well as the number of questions that achieved a median of 3.5 (only slightly related to personality), 4, 4.5 and 5 (highly related to personality) in each cluster. A further 12 questions achieved a median score of > 3, but

not at high consensus. The participants rated 37 (16%) questions as unrelated to PD in any cluster (not always at high consensus).

Levels of consensus were calculated for each cluster – both for questions that were and were not thought to be related to PD. In Cluster A, 54% of questions achieved high consensus, 84% in Cluster B, 84% in Cluster C1, and 69% in Cluster C2. Questions achieving low consensus (1%) were only found in Cluster C1. To examine whether there were differences in responses of participants by expertise, responses were split into two: those from experts in PD (16) and those from experts working with children (6). For the experts in PD, the number of parts reaching high consensus was 775 (81%), 176 (19%) parts reached medium consensus and one (0.1%) part achieved low consensus. For the experts working with children, numbers of parts reaching high consensus were 829 (87%), 118 (12%) reached medium consensus, and 5 (0.5%) parts reached low consensus. The distribution of responses between high, and medium or low, consensus differed between the two groups of experts (p < 0.001) (Table 2).

For the experts in PD, 205 questions were thought to be related to PD in one or more cluster at high consensus, with a further 3 at medium consensus. For the experts working with children, 139 questions were thought to be related to PD in one or more cluster at high consensus, and a further 48 at medium or low consensus. The difference in numbers of questions thought to be related to PD at high consensus differed significantly between the two groups of experts (p<0.0001). There was also a significant difference (p<0.001) in consensus levels among the two groups of experts for questions thought to be related to PD (Table 2).

Results of confirmatory factor analysis

To test the validity of the identified four cluster solution a confirmatory factor analysis was applied to a selected series of 50 items from the CHDS database (see Methods). This analysis showed that a four factor solution (Clusters A, B, C1, C2) did not fit the data well (Satorra-Bentler chi square = 183.7, df=82, p <.0001; CFI = .77; RMSEA = .035; SRMR =

.071), even when some elements of the error covariance matrix were permitted to be free. In addition, the fitted model produced some anomalous results, including an estimated correlation between the Cluster A and Cluster C1 factors equal to 1, suggesting that the underlying dimensionality of the model could be reduced.

Given these observations, further analyses were conducted fitting both a three factor solution (in which Clusters A and C1 were combined) and a two factor model (in which all internalising Clusters A, C1, C2 were combined). Comparison of the fit indices for the three models showed that while none of the models fitted the observed data well, the best fitting model appeared to be the 3 factor solution (Clusters A C1, Cluster B, Cluster C2). Jointly, these findings provide comparatively weak empirical support for the 4 cluster Delphi classification.

DISCUSSION

In this study, experts in PD and those working with children were asked to rate questions relating to children aged between 12 and 16 as potentially indicating abnormal personality traits. There were high levels of agreement concerning questions thought to be related to personality abnormality; this suggests that personality disorder in the under 18s may exist in some form, and identifiable. The experts in PD rated significantly more questions as related to PD, with significantly higher levels of consensus, than did the experts working with children.

Whether or not PD should be diagnosed in childhood and adolescence is a matter of controversy. Grilo et al. (1998) suggests that PDs seen in adolescence are valid forms of most PDs. Westen et al. (2003) reported good concurrent and predictive validity of a diagnosis of PD under the age of 18 (i.e., they are valid indicators of distress and dysfunction), and adequate internal consistency. However, Becker et al. (1999) found that internal consistency of PD criteria was generally lower in adolescents than adults, mainly limited to Cluster B and borderline disorders – suggesting that these criteria may not be applicable to this age group.

It has been reported that stability of PD in the under 18s is similar to that of adult PD (Chanen et al. 2004). Johnson et al. (2000) report that PD traits that are present in adolesence may disappear during the transition from adolescence to adulthood; and Skodol et al. (2007) found that people with PD who experience remission of symptoms during adolescence to adulthood often experience relatively little residual impairment by middle adulthood. Further research is needed using a longitudinal resource such as the CHDS to investigate to what extent symptoms of PD present in childhood or adolescence truly reflect PD that persists into adult life.

Most of the questions thought to be related to PD were classified as Cluster B, mainly reflecting the type of questions asked in the CHDS. The lowest levels of consensus were for Cluster A, with highest levels in Clusters B and C1. This may be because features within Cluster A are negative – these may be more difficult to identify than positive features (as in Cluster B).

When results were examined by expert group, those who work with children were more reluctant to label features as related to PD, and were less likely to agree, than the experts in PD. It is reasonable to conclude that those working with children may be more likely to view personality before the age of 18 as fluctuating and unstable and thereby likely to resolve with time. They may also be more inclined to select Axis I disorders such as attention deficit hyperactivity disorder (ADHD) and depressive disorder as formal diagnoses instead of personality disorder compared with those working with adults. It is also possible that the higher number of questions thought to be related to personality abnormality at high consensus among experts in PD is a consequence of their relating these features to adult rather than child behaviour.

STRENGTHS AND LIMITATIONS

Use of the Delphi technique allowed us to identify questions thought to be related to PD in the CHDS based on expert opinion. By including both experts in PD and experts working with children we increased the range and wealth of experience available. We only

took forward for consideration at each stage questions which were identified at high levels of consensus. These questions can now be investigated in the CHDS cohort to examine clusters of symptoms that may define PD in childhood and adolescence.

There are a number of limitations to this study. Firstly, there are intrinsic drawbacks to the Delphi technique. Although allowing participants to change their answers is generally seen as positive since any mistakes can be altered, and it encourages people to consider the question in more depth, participants may feel under pressure to change their answers to agree with the group response. There is also debate as to the reliability and validity of the Delphi technique. The extent that two different panels selected in the same way would produce the same results is not established (Williams & Webb, 1994). The number of participants in each stage of the present Delphi survey was limited, possibly affecting the validity and repeatability of our findings. Furthermore, response rate amongst experts invited into the study was relatively low, so findings may not have been truly representative of the target groups.

The present study was focussed on ages 12 to 16, and according to current thinking about PD, this is probably the lowest age limit that could be assessed. Further research will identify whether the questions selected in the Delphi study are able to characterise early onset PD at these ages. If personality abnormalities are to be identified in this age group, further studies could be undertaken at younger ages. Positive findings could also support accumulating evidence that the accepted age of diagnosis of PD should be lowered. If the age of diagnosis of PD is to be lowered, it is likely that clinicians would still be reluctant to make diagnoses before the age of 18 because of the stigma for the child attached. It is worth considering using a term for personality abnormalities that is less damaging than disorder. An alternative notion is that these conditions constitute a diathesis, i.e., there is a tendency to develop psychiatric disorders or abnormal behaviours, which are otherwise quiescent. Personality develops fairly early in life, and although it may not fully manifest itself until adulthood the notion of longstanding

vulnerability in the form of diathesis could allow the characteristics to be identified much earlier (Tyrer, 2007).

IMPLICATIONS OF FACTOR ANALYSIS

Application of confirmatory factor analysis using a subset of symptom data from the CHDS did not provide strong validation for the four cluster Delphi model. While none of the models fitted showed good fit to the observed data, the best solution appeared to be a three factor model comprising: Clusters A and C1 combined; Cluster B; and Cluster C2. These findings provide only partial validation for the classification approach developed by the investigators in this DELPHI process, and suggest the need for considerable reanalysis and item refinement to locate well fitting models.

CONCLUSION

Personality abnormalities in the under 18s are currently ignored, dismissed or misdiagnosed, and so consequently under-treated. Although PD in childhood is poorly understood, the Delphi technique was successful in obtaining high levels of consensus for specific questions thought to be related to personality abnormality. The high consensus reached amongst experts for a majority of questions suggests that personality abnormality may be identifiable in childhood and adolescence. The questions identified now need to be tested using prospective data from the CHDS, as this cohort offers the opportunity of testing whether the personality dimensions analysed are predictive of adult behaviors. If any of these personality characteristics under the age of 18 are supported, this will have important implications for both the treatment and prevention of PD.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Becker, D.F., Grilo, C.M., Morey, L.C., Walker, M.L., Edell, W.S., & McGlashan, T.H. (1999). Applicability of personality disorder criteria to hospitalised adolescents: evaluation of internal consistency and criterion overlap. *J Am Acad Child Adolesc Psychiatry, 38,* 200-205.

Bernstein, D.P., Cohen, P., Velez, C.N., Schwab-Stone, M., Siever, L.J., & Shinsato, L. (1993). Prevalence and stability of DSM-III-R personality disorders in a community-based survey of adolescents. *Am J Psychiatry, 150,* 1237-1243.

Chanen, A.M., Jackson, H.J., McGorry, P.D., Allot, K.A., Clarkson, V., & Yuen, H.P. (2004). Two-year stability of personality disorder in older adolescent outpatients. *J Personality Disorder*, *18,* 526-541.

Coid, J. (2003). Epidemiology, public health and the problem of personality disorder. *Brit J Psychiatry*, *182*, s3-s10.

Coid J, Yang M, Tyrer P, Roberts A, Ullrich S. (2006). Prevalence and correlates of personality disorder in Great Britain. *Br J Psychiatry* 2006; **188**: 423 -431.

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science, 2,* 458-467.

Durrett, C., & Westen, D. (2005). The structure of Axis II disorders in adolescents: a Cluster- and factor- analytic investigation of DSM-IV categories and criteria. *J Personality Disorders, 19,* 440-461.

Grilo, E.M., McGlashan, T.M., Quinlan, D.M., Walker, M.L., Greenfeld, D., & Edell, W.S. (1998). Frequency of personality disorders in two age cohorts of psychiatric inpatients. *Am J Psychiatry, 155,* 140-142.

Johnson, J.G., Cohen, P., Kasen, S., Skodol, A.E., Hamagami, F., & Brook, J.S. (2000). Age-related change in personality disorder trait levels between early adolescence and adulthood: a community-based longitudinal investigation. *Acta Psychiatr Scand, 102,* 265-275.

Johnson, J.G., Cohen, P., Skodol, A.E., Oldham, J.M., Kasen, S., & Brook, J.S. (1999). Personality disorders in adolescence and risk of major mental disorders and suicidality during adulthood. *Arch Gen Psychiatry, 56,* 805-811.

Kasen, S., Cohen, P., Skodol, A.E., Johnson, J.G., & Brook, J.S. (1999). Influence of child and adolescent psychiatric disorders on young adult personality disorder. *Am J Psychiatry, 156,* 1529-1535.

Levy, K.N., Becker, D.F., Grilo, C.M., Mattanah, J.J.F., Garnet, K.E., Quinlan, D.M., Edell, W.S., & McGlashan, T.H. (1999). Concurrent and predictive validity of the personality disorder diagnsosis in adolescent inpatients. *Am J Psychiatry, 156,* 1522-1528.

Mulder, R.T. & Joyce, P.R. (1997) Temperament and the structure of personality disorder symptoms. *Psychological Medicine* **27**, 99-106.

Muthen, L. K. & Muthen, B.O. (1998-2007). *Mplus user's guide* (5th ed.). Los Angeles, CA:

Paris, J. (2003). Personality disorders over time: Precursors, course and outcome. *J Personality Disorders, 17,* 479-488.

Sharkey, S.B., & Sharples, A.Y. (2001). An approach to consensus building using the Delphi technique: developing a learning resource in mental health. *Nurse Education Today 21,* 398-408.

Skodol, A.W., Johnson, J.G., Cohen, P., Sneed, J.R., & Crawford, T.N. (2007). Personality disorder and impaired functioning from adolescence to adulthood; *Brit J Psychiatry, 190*, 415-420.

Tyrer, P. (2004). New approaches to the diagnosis of psychopathy and personality

 disorder. *J Royal Soc Med, 94,* 371-374.

Tyrer, P. (2007). Personality diatheses: a superior explanation than disorder. *Psychol

 Med, 37*, 1521-1525.

Tyrer, P., Coombs, N., Ibrahimi, F., Mathilakath, A., Bajaj, P., Ranger, M., Rao, B., & Din,

 R. (2007). Critical developments in the assessment of personality disorder. *Brit J

 Psychiatry, 190*, s51-s59.

Tyrer, P. & Alexander, J. (1979) Classification of personality disorder. *British Journal of

 Psychiatry, 135*, 163-167.

Tyrer, P., & Johnson, T. (1996). Establishing the severity of personality disorder. *Am J

 Psychiatry, 153,* 1593-1597.

Tyrer, P., & Simonsen, E. (2003). Personality disorder in psychiatric practice. *World

 Psychiatry, 2*, 41-44.

Vizard E, Hickey N, McCrory E. (2007). Developmental trajectories associated with

 juvenile sexually abusive behaviour and emerging severe personality disorder in

 childhood: 3–year study. *Br J Psychiatry* , 190 (suppl 49): s27–32.

Westen, D., & Arkowitz-Westen, L. (1998). Limitations of Axis II in diagnosing personality

 pathology in clinical practice. *Am J Psychiatry, 155,* 1767-1771.

Westen, D., Dutra, L., & Shedler, J. (2005). Assessing adolescent personality pathology.

 Brit J Psychiatry, 186, 227-238.

Westen, D., Shedler, J., Durrett, C., Glass, S., & Martens, A. (2003). Personality

 diagnoses in adolescence: DSM-IV Axis II diagnoses and an empirically derived

 alternative. *Am J Psychiatry, 160,* 952-965.

Williams, P.L., & Webb, C. (1994). The Delphi technique: a methodological discussion.

 Journal of Advanced Nursing, 19, 180-186.

Table 1 – Numbers of parts of questions* reaching different levels of consensus in Phases II and III, and difference (Phase III – Phase II)

Level of consensus	Phase II	Phase III	Difference
High	515 (54%)	691 (73%)	+176
Medium	382 (40%)	258 (27%)	-124
Low	55 (6%)	3 (<0.5%)	-52
Total	952 (100%)	952 (100%)	

*Each question was divided into four parts corresponding with each of four clusters.

Table 2 – Distribution of consensus levels and p-values (Fisher's Exact test), by expertise of participants for all question parts and questions thought to be related to Personality Disorder (PD), Phase III*

Consensus level	All question parts (4 Clusters per question)		Questions thought to be related to PD	
	Experts in PD	Experts working with children	Experts in PD	Experts working with children
High	775 (81.4%)	829 (87.1%)	205 (98.6%)	139 (74.3%)
Medium/Low	177 (18.6%)	123 (12.9%)	3 (0.01%)	48 (25.7%)
Total	952	952	208	187
	p < 0.001		p < 0.001	

*Includes 21 questions from Phase II reaching high consensus

Figure Legends

Figure 1 – Summary of methods and Delphi process

CHDS Christchurch Health and Development Study
BIGSPD British and Irish Group for the Study of Personality Disorders
PD Personality Disorder

Figure 2 – Numbers of questions relating to each personality cluster by median score, Phase III*

*Includes 21 questions from Phase II reaching high consensus. Eight questions were attributed to more than one cluster.
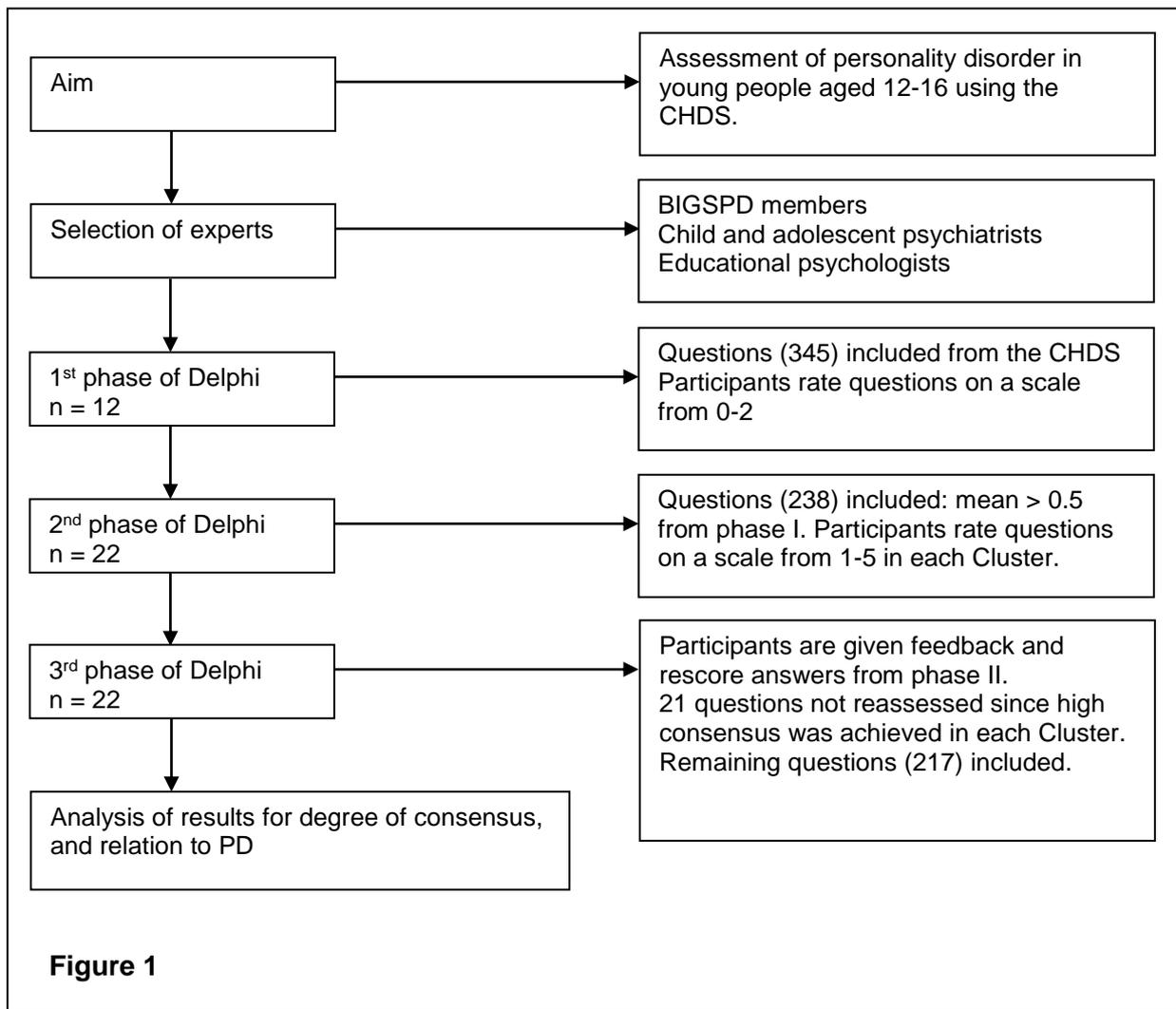
| Aim | → | Assessment of personality disorder in young people aged 12-16 using the CHDS. |

| Selection of experts | → | BIGSPD members
Child and adolescent psychiatrists
Educational psychologists |

| 1st phase of Delphi
n = 12 | → | Questions (345) included from the CHDS
Participants rate questions on a scale from 0-2 |

| 2nd phase of Delphi
n = 22 | → | Questions (238) included: mean > 0.5 from phase I. Participants rate questions on a scale from 1-5 in each Cluster. |

| 3rd phase of Delphi
n = 22 | → | Participants are given feedback and rescore answers from phase II.
21 questions not reassessed since high consensus was achieved in each Cluster.
Remaining questions (217) included. |

| Analysis of results for degree of consensus, and relation to PD |

Figure 1

Figure 2