ORIGINAL PAPER

Helen Sweeting · Robert Young · Patrick West

GHQ increases among Scottish 15 year olds 1987–2006

■ **Abstract** *Background* Increases in a number of psychosocial disorders have been identified among Western youth in the second half of the Twentieth century. However findings are not consistent, trends are complex, and comparisons over time are hampered by methodological problems. *Methods* Data were drawn from three samples identical in respect of age (15 years), school year (final year of statutory schooling) and geographical location (the West of Scotland). Each sample was administered the 12-item General Health Questionnaire, a measure of self-report psychological distress, in 1987 ($N = 505$), 1999 ($N = 2,196$) and 2006 ($N = 3,194$). Analyses were conducted to examine changes in: GHQ 'caseness'; individual items; and factors, derived via confirmatory factor analysis representing (a) 'negative' and 'positive' items, and (b) 'anxiety and depression', 'loss of confidence or self-esteem' and 'anhedonia and social dysfunction'. *Results* Based on the standard (2/3) cut-off, 'caseness' rates in 1987, 1999 and 2006 were 12.7, 15.1 and 21.5% (males) and 18.8, 32.5 and 44.1% (females). Similar increases were observed with more stringent 'caseness' cut-offs. Examination of individual items showed some to have increased much more markedly over time than others. There were larger increases among females for all except two items and some evidence, among both genders, of steeper increases among 'negative' items compared with 'positive' ones. However, the differences in slope were very small compared with the overall increases in both types. *Conclusions* Data from three samples identical in respect of age, school year and geographical location, show marked increases in GHQ-12 'caseness' among females between 1987 and 1999 and among both males and females between 1999 and 2006. Although slightly steeper increases in 'negative' items raise the possibility that endorsing such symptoms may have become more acceptable, these were small in comparison with increases in all dimensions of psychological distress. The next step is to identify causal explanations for the increases reported here.

■ **Key words** GHQ-12 – time-trends – young people – gender differences – confirmatory factor analysis (CFA)

H. Sweeting (✉) · R. Young · P. West
MRC Social and Public Health Sciences Unit
4, Lilybank Gardens
Glasgow (SCO) G12 8RZ, Scotland, UK
E-Mail: helen@sphsu.mrc.ac.uk

Introduction

Reviews conducted in the 1990s identified substantial increases among young people in a number of psychosocial disorders including depression, suicide, alcohol and drug use, and crime and conduct disorder in most Western countries since the Second World War [14, 30]. Within the UK, comparison of data from the parents of 15 to 16-year-olds in 1974, 1986 and 1999 found increases in conduct problems from 1974 and emotional problems from 1986. Further analyses in respect of conduct problems suggested these increases could not be attributed to greater willingness to report, since associations with adult outcomes such as occupational, relationship or parenting problems remained stable over time, which would not be expected if reporting thresholds had reduced [8]. Similarly, while studies of Dutch 6–16 year olds based on parent and teacher-report emotional and behavioural problems found few differences between 1983 and 1993 [39], small increases in parent-reported problems were evident when the period of comparison was extended, most of which had occurred between 1993 and 2003. This study found no clear evidence for gender- or age-specific trends [36].

However, not all studies have found increases in behavioural and emotional disorders among young people, and trends are complex [25]. For example, a meta-analysis performed on 310 samples of North American 8–16 year olds who had responded to the self-report Children's Depression Inventory found that 'contrary to expectation', between 1980 and 1998 among males, there was a slight decrease in scores, and no change in those of females. There was no age effect but a small birth cohort effect for males, and a small age effect but no birth cohort effect for females [38]. Problems identified by the parents of American 7–16 year olds via the Child Behaviour Checklist increased from 1976 to 1989, but reduced from 1989 to 1999 among both males and females and regardless of age [1]. A meta-analysis of international studies using structured diagnostic interviews to make formal diagnoses of depression on representative samples of children born between 1965 and 1996 provided no evidence of increasing rates in later-born cohorts, this result remaining unchanged when analysis was restricted to studies of under 13s, 13 to 18-year-old males or 13 to 18-year-old females [9]. A comparison of information relating to 8 to 9-year-old Finnish children, obtained from parents and teachers using Rutter's questionnaires to identify emotional and behavioural difficulties in 1989 and 1999 found a decrease in psychiatric symptoms among males and no change among females. [32]. Within the UK, parent and teacher ratings of 5–15 year olds via the Strengths and Difficulties Questionnaire showed no change or small declines in problem subscales and increases in levels of prosocial behaviours between 1999 and 2004. Parent-rated total difficulties declined among both younger (5–10) and older (11–15) groups, and regardless of gender [26].

As observed both by others and ourselves, comparisons of mental health problems over time are hampered by methodological difficulties. The potential evidence includes official statistics, which cover only a minority of problems and are subject to changes in recording; data on service use and prescribing which are influenced by changes in knowledge, diagnostic criteria, availability and attitudes; lifetime prevalence as reported by different age cohorts, which can be influenced by recall bias; and the comparison of data from repeated community based surveys [8, 26, 36, 41]. The latter include studies of secular changes based on comparisons of parent or teacher reports as outlined above. However, despite the importance of self-reports when investigating adolescent mental health [37], we are aware of only a very few repeat cross-sectional surveys using the same self-administered instrument on socially and geographically comparable groups of young people.

Even among the few studies which meet these criteria, results are mixed. Children in the Finnish study (above) completed the Child Depression Inventory, reporting more depressive symptoms in 1999 than 1989 [32]. A comparison of Dutch 11–18 year olds who completed the Youth Self-Report, assessing emotional and behavioural problems in 1993 and 2003, found small decreases in externalising and social problems among males, but increases in thought problems, suicidal ideation and self-harm among females. Some differences were found when the sample was divided by age (11–14 and 15–18 years) and gender; for example, somatic complaint scores decreased for older but not younger males, while total problem scores increased for younger but not older females [37]. A study of Greek 'adolescents' (ranging from 10 to 26 year olds), using the self-report General Health Questionnaire (GHQ), found significant and 'alarming' increases among both males and females between 1980 and 1998. The version of the GHQ used in this study provides sub-scores for 'somatic complaints', 'anxiety', 'social dysfunction' and 'depression', greater increases in 'depression' being found over time in females. These analyses included age as a covariate, but did not test whether trends varied in different age groups [13].

Analyses of our own data, derived from surveys which administered the shortest, 12-item GHQ (GHQ-12) to Scottish 15 year olds resident in the same geographical area in 1987 and 1999, found increased levels of 'psychological distress' among females but not males [41]. Although consistent with a number of other studies [8, 14, 30], this result contrasts with another Scottish study which compared data obtained from 11 to 15 year olds school children in 1994 and 2006, finding increases in the proportion who were 'very happy' (and, among females, 'always confident'), and decreases in the proportion reporting two or more common psychological/psychosomatic symptoms [21]. Our results also differ from those obtained from 11 to 15 year olds who completed the Strengths and Difficulties Questionnaire within the UK national studies of child and adolescent mental health; between 1999 and 2004, levels of emotional difficulties and conduct problems reduced while the other problem sub-scales and prosocial behaviours remained stable [26].

Following on from our earlier comparison, a further time-point is provided by a survey which administered the GHQ-12 to a third sample of 15 year olds in the same geographical area in 2006. These data allow us to examine more recent trends, and are the subject of the present paper. The key question is whether the increases in rates of 'caseness' previously seen among females between 1987 and 1999 [41] continued between 1999 and 2006, or whether, as some have suggested, there is evidence of stabilisation or improvements in mental well-being among young people in the UK [21, 26]. In addition to changes in overall levels of self-report psychological distress, it is possible that time-trends may vary for different items or groups of items within this measure, and/or for males compared with females [13]. If so, this would

allow us to say more about trends in particular dimensions of distress.

To do this, we have examined the evidence for time trends in respect of: (1) GHQ 'caseness'; (2) individual items; and (3) factors representing firstly, 'negative' and 'positive' items and secondly, 'anxiety and depression', 'loss of confidence or self-esteem' and 'anhedonia and social dysfunction' (see below for further details of GHQ scoring and factor structure).

Methods

■ Samples

All three cohorts included 15 year olds in Secondary 4 (S4), their final year of (Scottish) statutory education, and resident in the Central Clydeside Conurbation, a predominantly urban area centred around Glasgow. They were drawn from the 'West of Scotland Twenty-07 Study: Health in the Community' ('Twenty-07' [3]), the 'West of Scotland 11 to 16 Study: Teenage Health' ('11 to 16' [41]) and the most recent study in the series, 'Peers and Levels of Stress' ('PaLS' [20]). Although precise methods of sampling differed, each was a representative sample of 15 year olds from the Central Clydeside Conurbation.

Twenty-07 included a youth cohort (one of three age cohorts) sampled using a two stage stratified random procedure based on postcodes and targeted individuals within postcodes, selected with a probability proportional to the total population [11], and first surveyed in 1987 at age 15. At that time, a response rate of 65% (excluding those who had moved house prior to first contact) of the eligible sample was obtained, resulting in a total of 1,009 respondents. An examination of bias due to non-response revealed no significant gender or social class differences [10]. Two home interviews were conducted, the second by a nurse. This was completed by 96% of the sample and included a number of self-completion instruments, one of which was the GHQ-12. At the time of the nurse interview, just over half the sample were in their S4 school year, 30% in a higher school year (S5) and 15% had left. To maintain comparability with the two later studies, respondents in S5 or who had left school, together with a small number in special schools, were excluded from the analyses. The total number of respondents included in the comparison was therefore 505 (48% males) S4 pupils, mean age 15 years 8 months (SD 3.5 months), who completed the GHQ-12 between February and June 1987.

11 to 16 was a school-based study of a cohort in mainstream education, first surveyed in their final year of primary school, and followed up on two occasions in their secondary schools, including a contact at age 15 (S4) in 1999. The sampling scheme involved a number of steps to ensure a representative sample at both the primary and secondary school stages [12]. At the second follow-up, 2,196 respondents (51% males), mean age 15 years 5 months (SD 3.6 months), in 43 secondary schools, took part between January and March 1999. This group represented 85% of the baseline and 79% of the original eligible samples. During classroom sessions, respondents completed questionnaires, including the GHQ-12. Nurses helped with questionnaire completion if necessary. Comparison with census data showed the baseline sample to be representative in respect of gender and social class; thereafter, differential attrition (e.g. persistent school truants) made it less so [33].

PaLS was also mainstream school-based. Details of the sampling scheme are available [34]. This aimed to obtain a representative sample, and within selected schools, all pupils in the S4 year group were invited to participate between January and March 2006. The total sample comprised 3,194 (49%) males, mean age 15 years 5 months (SD 3.8 months), representing 81% of the eligible sample. As with '11 to 16', respondents completed questionnaires, including the GHQ-12, with help available if required. Participating schools did not differ significantly from the remainder in the area in respect of a number of socio-demographic dimensions, nor for pupil achievement by the end of statutory schooling. However within selected schools, pupils completing a questionnaire differed from non-responders in respect of gender and deprivation [34].

Thus, while all respondents filled in the GHQ-12 as part of a self-complete questionnaire administered within the context of a health and lifestyle survey, the earliest study differed from the other two in respect of setting (home in the presence of a nurse-interviewer versus school in the presence of peers, researchers and survey assistants).

■ Measure: the GHQ-12

The GHQ was designed as a measure of state, focusing on inability to carry out normal functions and the emergence of distressing symptoms. Each item includes four answer options which can be scored in several ways. Traditional, binary scoring indicates deviations from normal (0-0-1-1), and GHQ 'caseness' is generally defined via thresholds based on this method. Items representing inability to carry out normal functions (e.g. 'been able to enjoy your normal day-to-day activities'), which employ response options from 'better/more than usual' to 'much less', have been termed 'positive'. In contrast, new distressing symptoms (e.g. 'felt constantly under strain'), with options from 'not at all' to 'much more than usual' have been termed 'negative' items [16].

The version which we employed, the GHQ-12, has been validated for use with both older (age 17 [2]) and younger (ages 11–15 [35]) adolescents. Although originally conceived as measuring a single dimension, there has been debate over whether it is actually multidimensional. The literature on this, which relates almost entirely to studies of adults, is unclear, and studies have identified one, two and three-factor structures. Differences may result, at least partly, from different scoring and analytic methods [7, 28].

Most two-factor models identify dimensions representing 'social functioning' and 'anxiety and depression' [6, 31, 40]. 'Social functioning' generally includes the 'positive' items and 'anxiety–depression' 'negative' ones, raising the possibility that the factors may have methodological (response set) origins [19]. There may be both cultural and gender differences in the way these factors pattern. Thus, a comparison showed that Latin American adults scored higher than those in the UK, but only on the 'negative' items; that is those scoring when the response is 'more than usual'. The authors suggested this result may have occurred because of cultural differences, symptoms of psychiatric disorder being more socially acceptable in Latin America than the UK [22]. More recently, analyses of an adult UK sample found that females scored higher on the 'negative', but not the 'positive' factor [19], suggesting it may be more acceptable for females than males to respond 'more than usual' to 'negative' symptoms.

More recent studies, using confirmatory rather than exploratory factor analysis, have generally favoured a three-factor structure to the GHQ-12, particularly a model proposed by Graetz [18], comprising factors including elements of 'anxiety and depression' (four 'negative' items), 'anhedonia and social dysfunction' (six 'positive' items) and 'loss of confidence or self-esteem' (two 'negative' items). This model, which was based on data from a large representative Australian sample of 16–25 year olds, has been supported in confirmatory analyses of samples of adults [7, 24, 28, 31] and adolescents [15].

■ Analyses

Analyses of change over time in GHQ-12 'caseness' and individual items were based on categorical data, thus binary scoring was used. The cut-off for GHQ 'caseness' was defined as 2/3 [2, 16]. In addition, the robustness of the findings in respect of 'caseness' was checked with more stringent cut-offs of 3/4 and 4/5.

In order to examine whether patterns of change differed for particular dimensions, the first step was the use of binary CFA to evaluate the fit of the unitary, two ('negative' and 'positive' items)

and three ('anxiety and depression', 'loss of confidence or self-esteem' and 'anhedonia and social dysfunction') factor models at each time-point and overall. The Mplus statistical package [27] was used, with the weighted least squares mean and variance adjusted estimator. Model fit was assessed by several fit indices, among them the CFI (comparative fit index) and RMSEA (root mean square error of approximation, a statistic which is relatively unaffected by sample size). The results indicated a good fit to the data for all three models (for example, CFI fits ranging from 0.93 to 0.98), with the three factor model producing a marginally better fit than the others (further details available upon request from RY). Reflecting this, correlations between the factors were extremely high (anxiety–depression with anhedonia–social dysfunction $r = 0.989$; anxiety–depression with loss of confidence $r = 0.957$; anhedonia–social dysfunction with loss of confidence $r = 0.964$). Factor scores from the two and three factor solutions were analysed in respect of change over time. Note that saved scores from binary factor analysis are not distributed normally [27] and do not have a mean of 0 or SD of 1.

Analyses of both GHQ 'caseness' and individual items (i.e. categorical variables) comprised: (a) cross-tabulations according to date, for males and females separately; (b) logistic regressions, resulting in mutually adjusted odds ratios for females compared with males and for 1999 and 2006 compared with 1987; and (c) further logistic regressions including the gender by date interactions to examine whether the pattern of change between 1987 and 1999, and between 1999 and 2006 differed for males compared with females. Analyses of the factor scores (continuous variables) comprised: (a) means according to date, for males and females separately; (b) one-way ANOVAs, resulting in F values for differences in means between 1987 and 1999, and between 1999 and 2006, separately for males and females; (c) further ANOVAs including the gender by date interactions to examine whether the changes in factor scores over each of these two time periods differed for males compared with females; and (d) repeated measures ANOVA, including gender and year, with factor type as a within subjects factor in order to test for differences in trends.

Probabilistic weights have been constructed to compensate for differential attrition in '11 to 16' [33] and for socio-demographic differences between responders and non-responders in 'PaLS' [34]. However, since the results of analyses of total GHQ-12 'caseness' and individual items using weighted and unweighted data were very similar, those based on unweighted data are presented here.

Results

Figure 1 shows GHQ-12 'caseness' rates based on the standard (2/3), and more stringent (3/4 and 4/5) cut-offs, for males and females at each date. Based on the standard cut-off, 'caseness' rose steadily over time for females, with significant increases between 1987 and 1999 and then again between 1999 and 2006. The small increase for males from 1987 to 1999 was not significant (as previously described [41]), but that from 1999 to 2006 was. Analyses (not shown) found the ORs in 1999 and 2006 compared with 1987 to be 2.09 and 3.41 for females but 1.23 and 1.89 for males. The ORs for female excess 'caseness' therefore increased markedly over time, from 1.59 in 1987 to 2.88 in 2006. Patterns of change based on the other cut-offs were much the same. Thus between 1987 and 1999, rates based on the most stringent cut-off increased from 3.4 to 5.5% among males, while among females they almost tripled, from 6.6 to 18.4%. By 2006 the rates were 10.2% for males (three times that in 1987) and 26.7% for females (four times that in 1987).

Table 1 shows GHQ-12 'caseness' (2/3 cut-off) together with individual items according to gender and date. The items are listed according to the three Graetz factors, 'anxiety and depression' (GHQ-12 items 5, 9, 2 and 6—all negative items), 'loss of confidence or self-esteem' (items 10 and 11—negative items) and 'anhedonia and social dysfunction' (items 12, 3, 4, 8, 7 and 1—all positive items). Some items ('constantly under strain', 'unhappy or depressed') were reported far more frequently than others ('capable about making decisions'). After adjustment for date, there was a significant female excess for all items, although this ranged from ORs greater than 2.7 ('thinking of yourself as a worthless person', 'feeling unhappy or depressed') to less than 1.5 ('able to enjoy normal day-to-day activities').

Examination of the individual items showed some to have increased much more markedly over time than others. Thus, after adjustment for gender, endorsement of 'thinking of yourself as a worthless person' rather or much more than usual was around three and four times more likely in 1999 and 2006 respectively than in 1987, while being less or much less 'able to concentrate on whatever you're doing' was around three and five times higher. This contrasts with ORs of 1.08 and 1.65 in 1999 and 2006 compared with 1987 for 'playing a useful part' and 1.05 and 1.23 for 'able to enjoy normal day-to-day activities'. The latter was the only item which did not show a significant increase from 1987 to 2006.

As Table 1 shows, the significance of the gender by date interaction for 'caseness' over the earlier time period was 0.053, reflecting the much greater increase in rates for females than males between 1987 and 1999. Significant gender by date interactions over this period were found for 'felt constantly under strain', 'lost much sleep over worry', 'been able to enjoy your normal day-to-day activities' and 'been able to con-
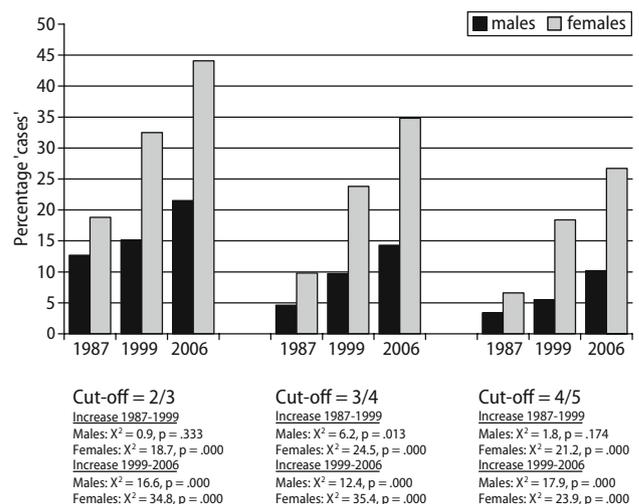

Figure 1 GHQ-12 'caseness': percentages of males and females at each date according to increasingly stringent cut-offs.

Table 1 GHQ-12 'caseness' and endorsement of each item[a] according to date, males and females

| | GHQ 'case' and 'case' scoring on each item—percentages (and 95% CIs) of males and females at each date | | | | | | Odds ratios (and 95% CIs) of GHQ 'case' and 'case' scoring on each item according to gender and date (mutually adjusted)[b] | | | | | Significance of gender by date interactions | |
| | Males | | | Females | | | Gender | | Date | | | | |
	1987	1999	2006	1987	1999	2006	Male	Female	1987	1999	2006	1987–1999	1999–2006
GHQ-12 'case'	12.7 (9.0–17.5)	15.1 (13.1–17.4)	21.5 (19.5–23.6)	18.8 (14.4–24.0)	32.5 (29.8–35.4)	44.1 (41.6–46.5)	1.00	2.73 (2.41–3.08)	1.00	1.73 (1.32–2.26)	2.75 (2.13–3.56)	(0.053)	(0.647)
Anxiety and depression													
5. Felt constantly under strain?[b]	17.6 (13.3–22.9)	19.0 (16.8–21.4)	21.9 (19.9–24.0)	20.9 (16.4–26.3)	32.7 (30.0–35.6)	36.5 (34.2–38.9)	1.00	1.99 (1.77–2.24)	1.00	1.49 (1.16–1.90)	1.77 (1.39–2.24)	(0.040)	(0.942)
9. Been feeling unhappy or depressed?	9.6 (6.5–14.0)	14.7 (12.8–16.9)	18.5 (16.7–20.5)	20.6 (16.1–26.0)	34.0 (31.2–36.9)	37.2 (34.9–39.6)	1.00	2.73 (2.41–3.09)	1.00	1.85 (1.41–2.42)	2.24 (1.73–2.91)	(0.487)	(0.326)
2. Lost much sleep over worry?	10.5 (7.2–15.0)	9.2 (7.7–11.1)	16.4 (14.7–18.3)	11.6 (8.3–16.1)	19.4 (17.1–21.8)	29.5 (27.3–31.8)	1.00	2.12 (1.85–2.43)	1.00	1.36 (1.00–1.85)	2.47 (1.84–3.32)	(0.019)	(0.511)
6. Felt you could not overcome your difficulties?	5.4 (3.2–9.1)	9.9 (8.3–11.8)	14.9 (13.2–16.7)	10.1 (7.0–14.4)	18.2 (16.1–20.7)	26.2 (24.1–28.4)	1.00	2.03 (1.76–2.34)	1.00	1.96 (1.38–2.78)	3.13 (2.22–4.39)	(0.908)	(0.993)
Loss of confidence or self-esteem													
10. Been losing confidence in yourself?	3.8 (2.0–7.0)	10.1 (8.4–12.0)	12.8 (11.2–14.6)	13.2 (9.6–17.9)	22.7 (20.3–25.3)	26.4 (24.3–28.6)	1.00	2.55 (2.21–2.94)	1.00	2.13 (1.52–2.98)	2.67 (1.92–3.71)	(0.341)	(0.631)
11. Been thinking of yourself as a worthless person?	1.3 (0.4–3.6)	4.3 (3.3–5.7)	6.8 (5.7–8.2)	5.0 (3.0–8.4)	14.1 (12.1–16.3)	16.1 (14.4–18.0)	1.00	2.96 (2.44–3.59)	1.00	3.14 (1.87–5.29)	4.05 (2.42–6.75)	(0.838)	(0.115)
Anhedonia and social dysfunction													
12. Been feeling reasonably happy, all things considered?	5.9 (3.5–9.6)	8.2 (6.7–9.9)	9.7 (8.3–11.2)	9.3 (6.3–13.5)	18.6 (16.4–21.0)	21.7 (19.8–23.8)	1.00	2.52 (2.15–2.95)	1.00	1.92 (1.34–2.74)	2.32 (1.64–3.29)	(0.238)	(0.952)
3. Felt you were playing a useful part in things?	5.9 (3.5–9.6)	6.4 (5.1–7.9)	8.1 (6.8–9.6)	8.9 (6.0–13.0)	9.5 (7.9–11.5)	15.1 (13.4–16.9)	1.00	1.83 (1.53–2.18)	1.00	1.08 (0.75–1.57)	1.65 (1.16–2.35)	(0.986)	(0.193)
4. Felt capable about making decisions about things?	3.3 (1.7–6.5)	2.6 (1.8–3.7)	5.2 (4.2–6.4)	4.7 (2.7–8.0)	6.3 (5.0–7.9)	13.3 (11.8–15.1)	1.00	2.62 (2.10–3.27)	1.00	1.13 (0.69–1.86)	2.51 (1.58–4.00)	(0.263)	(0.687)
8. Been able to face up to your problems?	3.8 (2.0–7.0)	5.9 (4.7–7.5)	6.9 (5.7–8.3)	5.8 (3.6–9.4)	13.3 (11.4–15.4)	16.6 (14.8–18.5)	1.00	2.53 (2.11–3.03)	1.00	2.15 (1.39–3.32)	2.71 (1.77–4.14)	(0.349)	(0.608)
7. Been able to enjoy your normal day-to-day activities?	12.1 (8.6–16.9)	9.4 (7.9–11.3)	11.1 (9.7–12.8)	10.1 (7.0–14.4)	13.6 (11.7–15.8)	15.2 (13.5–17.1)	1.00	1.40 (1.19–1.64)	1.00	1.05 (0.77–1.44)	1.23 (0.91–1.66)	(0.049)	(0.750)
1. Been able to concentrate on whatever you are doing?	6.3 (3.8–10.1)	10.6 (8.9–12.5)	15.4 (13.7–17.2)	5.4 (3.3–8.9)	19.1 (16.9–21.6)	32.0 (29.7–34.3)	1.00	2.30 (2.00–2.64)	1.00	2.89 (1.95–4.30)	5.22 (3.55–7.68)	(0.037)	(0.088)
N[c]	237	1,085	1,505	256	1,057	1,539							

[a]Based on binary scoring (0-0-1-1) to indicate deviations from normal

[b]Numbers show order of items as presented in GHQ-12

[c]N total number included in 'caseness' analyses. N's for individual items vary slightly, but since all items must be answered for determination of 'caseness', are never less than the N shown

centrate on whatever you are doing'. In each case increases were larger for females. None of the gender by date interactions over the later time period reached significance. Only two items showed somewhat larger increases for males over the entire 19-year-period, 'losing confidence in yourself' (OR in 2006 compared with 1987 of 3.74 for males and 2.36 for females) and 'thinking of yourself as a worthless person (OR in 2006 of 5.77 for males and 3.61 for females). However, the gender by date interactions were not significant, and even in 2006, the female excess in these two items remained striking.

These two items, 'losing confidence in yourself' and 'thinking of yourself as a worthless person', comprise Graetz's [18] 'loss of confidence or self-esteem' factor. However, Table 1 suggests no patterning in respect of increases in the other items. Table 2 investigates this further by examining changes in the factor scores representing two-factor ('negative' and 'positive' items) and three-factor ('anxiety and depression', 'loss of confidence or self-esteem' and 'anhedonia and social dysfunction') solutions. The table focuses on changes between 1987 and 1999, and 1999 and 2006 separately. Over the earlier time period, there were significant gender by date interactions for all factors; increases in all factor scores were much smaller for males than females, with the factor 'loss of confidence or self-esteem' being the only one to show a significant increase among males. In contrast, between 1999 and 2006 all factors increased for both males and females. Repeated measures ANOVA, with factor type as a within subjects factor, indicated small, but significant differences between the factors in their trends over time. The rate of increase was (1) steeper for the negative than the positive factor, and (2) steepest for 'loss of confidence or self-esteem' and least steep for 'anhedonia and social dysfunction' (graphs available from HS). Three-way interactions of factor type by gender by year were non-significant in respect of both the two factor ($P = 0.910$) and three factor ($P = 0.796$) solutions.

Discussion

Using data from representative samples of Scottish 15 year olds resident in the same geographical area, we have previously shown that GHQ-12 'caseness' rates increased among females, but not males, between 1987 and 1999 [41]. The present analyses extend these findings, demonstrating increases among both males and females between 1999 and 2006. The robustness of this result is underlined by finding the same pattern of marked increases in rates with more stringent definitions of 'caseness'; between 1987 and 2006, rates based on a cut-off of 4/5 increased three-fold for males and fourfold for females.

Table 2 GHQ-12 two and three factor solutions: factor scores at each date, males and females

	Mean (SD) scores of males and females at each date						1987–1999: F value (sig) according to date for males and females (and sig of gender by date interaction)			1999–2006: F value (sig) according to date for males and females (and sig of gender by date interaction)		
	Males			Females			Males	Females	Gender × date (sig)	Males	Females	Gender × date (sig)
	1987	1999	2006	1987	1999	2006	F (sig)	F (sig)		F (sig)	F (sig)	
Two factors												
Negative items	−0.139 (0.362)	−0.082 (0.429)	−0.003 (0.484)	−0.033 (0.460)	0.172 (0.568)	0.286 (0.584)	3.6 (0.059)	28.6 (0.000)	(0.002)	18.4 (0.000)	24.3 (0.000)	(0.238)
Positive items	−0.114 (0.334)	−0.070 (0.386)	−0.002 (0.430)	−0.032 (0.406)	0.149 (0.504)	0.257 (0.523)	2.6 (0.105)	28.4 (0.000)	(0.002)	17.1 (0.000)	27.7 (0.000)	(0.123)
Three factors												
Anxiety and depression	−0.121 (0.372)	−0.069 (0.427)	0.010 (0.482)	−0.020 (0.460)	0.180 (0.561)	0.296 (0.577)	2.9 (0.086)	28.0 (0.000)	(0.002)	18.7 (0.000)	25.8 (0.000)	(0.210)
Loss of confidence or self-esteem	−0.136 (0.386)	−0.060 (0.490)	0.024 (0.549)	−0.003 (0.524)	0.229 (0.657)	0.344 (0.667)	5.1 (0.024)	27.7 (0.000)	(0.005)	16.1 (0.000)	18.7 (0.000)	(0.366)
Anhedonia and social dysfunction	−0.100 (0.332)	−0.057 (0.382)	0.010 (0.425)	−0.019 (0.402)	0.158 (0.497)	0.266 (0.516)	2.6 (0.109)	28.3 (0.000)	(0.002)	17.1 (0.000)	27.9 (0.000)	(0.122)
N	237	1,085	1,505	256	1,057	1,539						

The main strength of the study is that it overcomes many of the problems encountered by other similar comparisons, in that the samples (15 year olds in the same school year), geographical location, measure and survey context ('health and lifestyle') were identical at each time-point. One limitation is that the response rate in the earliest study was lower than that in the later two, but while this might have impacted on differences between 1987 and 1999, it has no bearing on those between 1999 and 2006. It is also possible that the different setting in which the GHQ was administered in the earliest study (during a home interview) compared with the other two (school-based survey sessions) might have had some impact. Situational effects including the presence of others and perceived degree of confidentiality, anonymity and privacy have been shown to bias the completion of self-report questionnaires [4]. If, as is possible, gender differences exist in respect of these effects, this might have contributed to the greater increases seen among females in the earlier period. However, since the methodology of the 1999 and 2006 studies was identical, alternative explanations must be found for the increases seen among both males and females in the later period. Another potential methodological explanation, time of year, can probably also be ruled out; fieldwork for the earliest study ran from February to June, that for both later studies was conducted between January and March. This makes it less likely that stressors associated with proximity to examinations contributed to the increases seen.

Findings in respect of trends in the mental health of children and young people are mixed. Ours are broadly consistent with some [8, 13, 14, 30] but contrast with others [1, 9, 21, 26, 32, 38]. Scottish data in respect of somewhat older groups also provides a mixed picture. Between 1995 and 2003, the proportion of male 16 to 24-year-old Scottish Health Survey respondents scoring four or more on the GHQ-12 increased only very slightly from 9% to 10% while the proportion of females remained constant at 16% [5]. However, between 1980–82 and 2000–2002, the death rate per 100,000 population attributable to intentional self-harm and events of undetermined intent increased from 17 to 37 among Scottish 15 to 29-year-old males and 5–9 among females [23].

While our key question related to overall 'caseness' rates, we were also interested in whether patterns of change differed for particular items or groups of items and/or for males compared with females. Some individual items increased much more markedly over time than others, increases being greater for females than males for all except two items. Among both males and females, there was evidence of somewhat steeper increases among the 'negative' items; that is those that score when the response is 'more than usual', raising the possibility that endorsing such symptoms may have become more acceptable. However, it is important not to overstate this finding. The differences in slope were very small compared with the overall increases in both the 'positive' and 'negative' factors. In addition, although the results of CFA showed the fit of the three-factor model to be marginally better than the unitary or two-factor models, the correlations between the factors were extremely high. Our findings strongly suggest that rather than over-interpreting the factors, the GHQ-12 is best used as a unitary measure [7, 15, 31].

As suggested by others [13], the increases we have demonstrated over time might simply reflect generally greater willingness to express psychological distress or social acceptability of symptoms. We have suggested that this hypothesis might be supported by increases in 'negative' items, the assumption being that an increase in 'positive' items indicates 'real' change. However, this is not necessarily so, since it remains possible that both types have become more socially acceptable.

In evaluating the evidence, cultural differences in the predictive value of the GHQ have been demonstrated in comparisons against standardised diagnostic interviews [17, 22]. We are not aware of studies which have used this 'gold standard' method to examine time trends in GHQ scores, but it is possible that greater willingness to express psychological distress might have increased the threshold for 'caseness'. However, analyses conducted in respect of increases in conduct problems among UK adolescents between 1974 and 1999 found associations with adult outcomes did not change, suggesting the trends did not result from reduced reporting thresholds [8].

On the assumption that our results are consistent with the body of evidence indicating increases in psychosocial disorders among young people, they suggest the need for greater attention at the level of primary care or, given evidence that this age group find it difficult to consult their GP with mental health concerns [29], within school or alternative counselling services.

Conclusion

Using data from three samples identical in respect of age, school year and geographical location, we have built on previous analyses showing increases in GHQ-12 'caseness' among females but not males between 1987 and 1999. Results show increases among both males and females between 1999 and 2006. The next step is to identify causal explanations for the increasing levels of self-report psychological distress identified here

586

also due to the young people, nurse interviewers, schools, and all those from the MRC Social and Public Health Sciences Unit involved in the three studies described here.

References

1. Achenbach TM, Dumenci L, Rescorla LA (2003) Are American children's problems still getting worse? A 23-year comparison. J Abnorm Child Psychol 31:1–11
2. Banks MH (1983) Validation of the General Health Questionnaire in a young community sample. Psychol Med 13:349–354
3. Benzeval M, Der G, Ellaway A, Hunt K, Sweeting H, West P, Macintyre S (2008) Cohort profile: West of Scotland Twenty-07 Study: Health in the Community. Int J Epidemiol doi:10.1093/ije/dyn213
4. Brener N, Billy J, Grady W (2003) Assessment of factors affecting the validity of self-reported health-risk behaviour among adolescents: evidence from the scientific literature. J Adolesc Health 33:436–457
5. Bromley C, Sproston K, Shelton N (2005) The Scottish Health Survey-2003 results (vol 2—adults). Scottish Executive Health Department, Edinburgh
6. Campbell A, Walker J, Farrell G (2003) Confirmatory factor analysis of the GHQ-12: can I see that again? Aust N Z J Psychiatry 37:475–483
7. Campbell A, Knowles S (2007) A confirmatory factor analysis of the GHQ12 using a large Australian sample. Eur J Psychol Assess 23:2–8
8. Collishaw S, Maughan B, Goodman R, Pickles A (2004) Time trends in adolescent mental health. J Child Psychol Psychiatry 45:1350–1362
9. Costello EJ, Erkanli A, Angold A (2006) Is there an epidemic of child or adolescent depression? J Child Psychol Psychiatry 47:1263–1271
10. Der G. A comparison of the West of Scotland Twenty-07 Study sample with the 1991 Census SARs. Working Paper No. 60: Glasgow, MRC Medical Sociology Unit 1998
11. Ecob R. West of Scotland Twenty-07 Study: the sampling frame and procedures for the cohort studies. Working Paper No. 6: Glasgow: MRC Medical Sociology Unit 1987
12. Ecob R, Sweeting H, West P, Mitchell R. The West of Scotland 11 to 16 Study: schools, sample design and implementation issues. Working Paper No. 61: Glasgow: MRC Medical Sociology Unit 1996
13. Fichter MM, Xepapadakos F, Quadflieg N, Georgopoulou E, Fthenekis WE (2004) A comparative study of psychopathology in Greek adolescents in Germany and Greece in 1980 and 1998–18 years apart. Eur Arch Psychiatry Clin Neurosci 254:27–35
14. Fombonne E (1998) Increased rates of psychosocial disorders in youth. Eur Arch Psychiatry Clin Neurosci 248:14–21
15. French D, Tait R (2004) Measurement invariance in the General Health Questionnaire-12 in young Australian adolescents. Eur Child Adolesc Psychiatry 13:1–7
16. Goldberg D, Williams P. A User's Guide to the General Health Questionnaire. NFER-Nelson, Windsor 1988
17. Goldberg D, Oldehinkel T, Ormel J (1998) Why GHQ threshold varies from one place to another. Psychol Med 28:915–921
18. Graetz B (1991) Multidimensional properties of the General Health Questionnaire. Soc Psychiatry Psychiatr Epidemiol 26:132–138
19. Hu Y, Stewart-Brown S, Twigg L, Weich S (2007) Can the 12-item General Health Questionnaire be used to measure positive mental health? Psychol Med 37:1005–1013
20. Kelly SJ, Young R, Sweeting H, Fischer JE, West P (2008) Levels and confounders of morning cortisol collected from adolescents in a naturalistic (school) setting. Psychoneuroendocrinology 33:1257–1268
21. Levin K, Todd J, Currie D, Currie C Mental well-being of young people in Scotland: 1994–2006. HBSC Briefing Paper 14, Child and Adolescent Health Research Unit, University of Edinburgh, Edinburgh, UK 2007. http://www.education.ed.ac.uk/cahru/publications/BriefingPaper_14.pdf
22. Lewis G, Araya RI (1995) Is the General Health Questionnaire (12 item) a culturally biased measure of psychiatric disorder? Soc Psychiatry Psychiatr Epidemiol 30:20–25
23. Leyland AH, Dundas R, McLoone P, Boddy FA. Inequalities in Mortality in Scotland 1981–2001. Occasional Paper No. 16: Glasgow: MRC Social & Public Health Sciences Unit 2007
24. Makikangas A, Feldt T, Kinnunen U, Tolvanen A, Kinnunen ML, Pulkkinen L (2006) The factor structure and factorial invariance of the 12-item General Health Questionnaire (GHQ-12) across time: evidence from two community-based samples. Psychol Assess 18:444–451
25. Maughan B, Iervolino AC, Collishaw S (2005) Time trends in child and adolescent mental disorders. Curr Opin Psychiatry 18:381–385
26. Maughan B, Collishaw S, Meltzer H, Goodman R (2008) Recent trends in UK child and adolescent mental health. Soc Psychiatry Psychiatr Epidemiol 43:305–310
27. Muthén LK, Muthén B. Mplus User's Guide, 5th edn. Los Angeles: Muthén & Muthén, 1998–2007
28. Penninkilampi-Kerola V, Miettunen J, Ebeling H (2006) A comparative assessment of the factor structure and psychometric properties of the GHQ-12 and the GHQ-20 based on data from a Finnish population-based sample. Scand J Psychol 47:431–440
29. Potts Y, Gillies M, Wood S (2001) Lack of mental well-being in 15-year olds: an undisclosed iceberg? Family Practice 18:95–100
30. Rutter M, Smith DJ (eds) Psychosocial disorders in young people: time trends and their causes. Chichester: Wiley, 1995
31. Shevlin M, Adamson G (2005) Alternative factor models and factorial invariance of the GHQ-12: a large sample analysis using confirmatory factor analysis. Psychol Assess 17:231–236
32. Sourander A, Santalahti P, Haavisto A, Piha J, Ikaheimo K, Helenius H (2004) Have there been changes in children's psychiatric symptoms and mental health service use? A 10-year comparison from Finland. J Am Acad Child Adolesc Psychiatry 43:1134–1145
33. Sweeting H, Der G, West P. Bias, attrition and weighting in respect of the West of Scotland 11 to 16 Study's baseline, S2 and S4 surveys. Working Paper No. 9: Glasgow, MRC Social & Public Health Sciences Unit, 2001
34. Sweeting H, Young R, West P. The Peers and Levels of Stress ('PaLS') Study: basic frequencies and documentation. Working Paper No. 17: Glasgow, MRC Social & Public Health Sciences Unit, 2008
35. Tait R, French D, Hulse G (2003) Validity and psychometric properties of the General Health Questionnaire-12 in young Australian adolescents. Aust N Z J Psychiatry 37:374–381
36. Tick NT, van der Ende J, Verhulst FC (2007) Twenty-year trends in emotional and behavioural problems in Dutch children in a changing society. Acta Psychiatrica Scandinavica 116:473–482
37. Tick NT, van der Ende J, Verhulst FC (2008) Ten-year trends in self-reported emotional and behavioural problems of Dutch adolescents. Soc Psychiatry Psychiatr Epidemiol 43:349–355
38. Twenge JM, Nolen-Hoeksema S (2002) Age, gender, race, socioeconomic status, and birth cohort differences on the Children's Depression Inventory: A meta-analysis. J Abnorm Psychol 111:578–588
39. Verhulst FC, van der Ende J, Rietbergen J (1997) Ten-year time trends of psychopatholgy in Dutch children and adolescents: no evidence for strong trends. Acta Psychiatrica Scandinavica 96:7–13
40. Wernecke U, Goldberg DP, Yalcin I, Ustun BT (2000) The stability of the factor structure of the General Health Questionnaire. Psychol Med 2000:823–829
41. West P, Sweeting H (2003) Fifteen, female and stressed: changing patterns of psychological distress over time. J Child Psychol Psychiatry 44:399–411