Linköping University Post Print

Predicting Nonresponse Bias from Teacher Ratings of Mental Health Problems in Primary School Children

Kjell Morten Stormark, Einar Heiervang, Mikael Heimann,
Astri Lundervold and Christopher Gillberg

N.B.: When citing this work, cite the original article.

The original publication is available at www.springerlink.com:

Kjell Morten Stormark, Einar Heiervang, Mikael Heimann, Astri Lundervold and Christopher Gillberg, Predicting Nonresponse Bias from Teacher Ratings of Mental Health Problems in Primary School Children, 2008, Journal of Abnormal Child Psychology, (36), 3, 411-419.

<http://dx.doi.org/10.1007/s10802-007-9187-3>

Copyright: Springer Science Business Media

<http://www.springerlink.com/>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-16788>

Predicting nonresponse bias from teacher ratings of mental health problems
in primary school children

Kjell Morten Stormark 1)

Einar Heiervang 2)

Mikael Heimann 1), 3)

Astri Lundervold, 1) 4)

Christopher Gillberg, 5)

- 1) Centre for Child and Adolescent Mental Health, University of Bergen, Norway
- 2) Department of Child and Adolescent Mental Health, Haukeland University Hospital, Bergen, Norway
- 3) Department of Behavioural Sciences and Learning, Linköping University, Sweden
- 4) Department of Biological and Medical Psychology, University of Bergen, Norway
- 5) Child Neuropsychiatry Clinic, Queen Silvia's Children's Hospital, Sahlgren University Hospital, Gothenburg, Sweden

Address all correspondence to Kjell Morten Stormark, Centre for Child and Adolescent Mental Health, University of Bergen, PO Box 7800, 5020 BERGEN. E-mail: kjell.stormark@rbup.uib.no, Fax: +47 555 88 379.

Abstract

The impact of nonresponse on estimates of mental health problems was examined in a prospective teacher screen in a community survey of 9155 seven-nine year olds. For 6611 of the children parents consented to participation in the actual study (Responders) and for 2544 children parental consent was not obtained (Nonresponders). The teacher screen involved assessment of a broad set of problems and symptoms (emotional, conduct/oppositional defiance, hyperactivity, inattention, peer relationships, social functioning, obsessive compulsive disorder (OCD), tics and also an assessment of possible functional impairment (impact). Calculations of non-response coefficients, a function of effect sizes and non-response proportion, revealed only ignorable non-response bias for both mean scores and correlations between subscale scores. However, the results from binary logistic regressions revealed that moderate symptom scores on the teacher screen were associated with increased risk of children not obtaining parental consent to participate. This was most frequent among children with no impact, but it was a consistent finding that children with high scores of attention problems were more likely not to obtain parental consent to participate. These findings suggest that effect size estimates of non-response bias should be accompanied by a logistic regression approach to estimate if characteristic of the child is related to study participation.

Introduction

Nonresponse is a challenge in all epidemiologic research. It introduces a possible bias when there are substantial and systematic differences between the research sample and the target population. This is a threat to the generalizability of the research findings. Such a bias may arise for a number of reasons; sample frame exclusion, differential mortality or greater reluctance to participate. Surveys of mental health problems in children are often thought to be particularly susceptible to nonresponse bias since the children's caregivers, who consent to participation, may be reluctant to report such problems (e.g Owens et al, 2002). This is also based on the observation that many children with mental health problems never seek treatment (e.g. McMiller & Weisz, 1996; Pavuluri, Luk, & McGee, 1996). As a consequence, estimates of mental health problems from population-based surveys may be inaccurate, with possible serious consequences for the tailoring of appropriate health services. Differences in the magnitude of non-response bias may also be one of the reasons for the considerable differences in prevalence estimates for child mental health problems (see Roberts et al. 1998).

Response proportion is often used as an indicator of the likelihood of nonresponse bias, but a high response proportion is neither a necessary nor a sufficient condition for unbiased estimates (e.g. Stang, 2003). In fact, substantial differences between Responders and Nonresponders has been

found in surveys of mental health problems in children with low nonresponse proportion (e.g. Noll et al. (1997), whereas high nonresponse proportion might be associated with little bias, as in French and Waas (1985). An explanation for a possible inverse relation could be that a large nonresponse proportion comprise multiple reasons for nonresponse and that these level each other out, while when nonresponse proportion is low it may be more directly linked to the subject of the study, and thereby creating more bias. Noll et al. (1997) found that the small group of children who did not participate in a study of social skills were judged to be less socially accepted by their peers. However, to substantiate that amount of nonresponse bias is linked to the reason for nonresponse, one would need information from the Nonresponders about their reasons to not participate, which most surveys find difficult to provide.

Surveys on mental health problems in children instead typically evaluate the impact of nonresponse by comparing various child and family background variables (e.g. Weinberger et al, 1990). However, even if there are differences between the two groups, they do not necessarily imply that this causes bias. To draw such a conclusion, one would need to substantiate that characteristics of nonresponders are associated with the outcome measures; something which is rarely done.

Gerrits et al. (2001) argued that the magnitude of nonresponse bias could be estimated by the use of standardized mean scores (or effect sizes)

between Responders and nonresponders adjusted for nonresponse proportion. Gerrits et al. (2001) used the data from an original survey of social development to demonstrate that even for data sets with substantial nonresponse proportions, there were only negligible non-response bias (ranging from .00 to $\pm .12$), due to small differences between the Responders and Nonresponders..

However, effect size estimates of overall group comparisons may be of limited interest, since most mental health problems are relatively low frequent in children. What is interesting is thus not only if a response and nonresponse group differ in estimates of mental health problems at a group level, but if the number of children with high scores are over- or under-represented among the nonresponders. A way to accommodate this would be to examine the frequency distribution of scores in Responders and Nonresponders and calculate their risk for being high scorers on the screening questionnaire.

To understand nonresponse bias in surveys of mental health problems in children, both Responders and Nonresponders should be assessed prospectively within the same study to ensure that all children are screened on the same items and under the same circumstances. This was achieved in the first wave of the Bergen Child Study, a total population community survey of primary school children, where teachers filled in the questionnaire for all their pupils independent of parental consent. Teacher

screen for children with parental consent were subsequently made identifiable (Responders), while the teacher screen for the children without parental consent remained anonymous (Nonresponders).

Nonresponse bias was evaluated against a broad range of symptoms/problem areas: emotional symptoms, conduct problems, hyperactivity/inattention, problems with peer relations, problems with social functioning, tics, obsessive compulsive disorders (OCD) and Oppositional Defiance Disorder (ODD). Teachers also rated the impact of the symptoms/problems, to yield an estimate of the child's functional impairment. Two different statistical approaches were used to answer the question of nonresponse bias. The first approach comprised comparisons of mean scores and correlations using bias estimates and effect size calculations according to Gerrits et al.'s (2001) suggestions. The second approach comprised comparisons of the frequency distributions of subscale scores between the two groups, and logistic regression to evaluate the extent to which the teachers' report of mental health problems in the children predicted the likelihood of the children to take part in the study.

Method

Sample

The target population comprised all 9430 children who, during October-November 2002, attended 2nd-4th grade (i.e. 7 – 9 year olds) at any primary

school (public or private) in the municipality of Bergen, Norway. We received informed consent and teacher and/or parent screen for 7007 children (74,3% of target population). A teacher screen, independent of parental consent was obtained for 9155 children (97,2 % of target population), which comprised 6611 children whose parents consented to participation (response group) and 2544 children whose parents did not consent (nonresponse group).

Procedure

The Bergen Child study is a prospective study of children's mental health from primary school age to adolescence. The first wave of the study was organized in three-steps, where step 1 comprise the screening phase (where the present study draws its data), step 2 the diagnostic phase and step 3 the clinical assessment phase (see also Heiervang et al., 2007).

Members of the research group visited all 71 primary (private or public) schools in the municipality of Bergen to give all form teachers direct information about the background and the procedure of the study. The teachers received written and oral information explaining how teachers should distribute envelopes containing the material for the study via the pupil's knapsack to the parents from a given date. The parents were allowed a four week period to return consent and parental questionnaire forms to the school in two separate envelopes. The form teachers were asked

to fill in the teacher questionnaire for every pupil during the same four-week period. The teachers used a temporary identification code on a sticker attached to the questionnaire until the end of the data collection period, which they replaced with the unique identification number found on the parental consent form for the pupils when the teachers received these from the parents (Responders). For the pupils whose parents did not consent, the teachers simply removed the sticker with the temporary identification code, so that the teacher questionnaire for these pupils (Nonresponders) became anonymous. In order to ensure the confidentiality for the children in the Nonresponse group, the teacher questionnaire did not involve demographic information which could lead to involuntary disclosure for children in the non-response group.

The study was approved by the Regional Committee for Medical Research Ethics in Western Norway (REKIII 184.01) and the Data Inspectorate in Norway.

Questionnaire

The BCS comprises both a parent- and teacher-screen. Since this study draws its data solely from the teacher-screen, only the teacher questionnaire is described here. However, the parent questionnaire was almost identical to the teacher questionnaire. The teacher questionnaire was made up of three main instruments: the Strengths and Difficulties

Questionnaire (SDQ; Goodman et al , 1999), the Autism Spectrum Screening Questionnaire (ASSQ; Ehlers, Gillberg & Wing, 1999), the Swanson, Nolan and Pelham (SNAP-IV) rating scale (Swanson, 1992), which also coincides with the DSM-IV diagnostic criteria for Oppositional Defiant Disorder (ODD) and Attention-Deficit/Hyper-activity Disorder (ADHD).

The SDQ comprise four symptom/problem subscales; emotional symptoms, conduct problems, hyperactivity/inattention and peer relationship problems. In addition the questionnaire comprised five items addressing Obsessive Compulsive Disorder (OCD; taken from Thomsen, 1998), five items addressing tics and Tourette syndrome (from Apter et al ., 1993). The symptom questions from the SDQ and the rest of the above questions comprised the first three pages of the questionnaire. For all these questions there were three boxed response categories; Not True, Somewhat True and Certainly True. Finally, the questionnaire comprised questions on impairment, Impact section from the SDQ, with five questions: If the teacher thinks the child has a problem, and subsequent questions of chronicity, distress, social impairment (with peer relationships and classroom teaching) and burden to teacher or class. These items had four boxed categories; No, Yes – minor difficulties, Yes- definite-difficulties, Yes – severe difficulties.

Data analyses

The effect of a possible nonresponse bias was examined with respect to the following twelve subscale scores on the teacher questionnaire: SDQ Emotion problem subscale, SDQ Conduct problem subscale, SDQ Hyperactivity problem subscale, SDQ Peer problem subscale, SDQ Total difficulties (which consists of the first four subscales combined), SDQ-impact score, the ASSQ totalscore, the OCD total score, the Tics total score, and the SNAP-ODD, SNAP-inattention and SNAP- Hyperactivity subscales.

The statistical approach for this research question involved two sets of analyses. The first set concerned the computation of nonresponse bias coefficients for each of the symptom and the impact subscale scores. Differences between the response and the nonresponse group were tested using t-tests.. Then the standardized mean score (effect sizes) for each subscale were calculated as the difference between the Nonresponders and the Responders' mean score, divided by the pooled standard deviation for the total sample (Rosenthal & Rosenow, 1991). A nonresponse bias coefficient was then calculated as a function of the effect size and non-response proportion, according to Gerrits et al. (2001). In addition, non-response bias in correlations was calculated as the difference between coefficients for the response group and the total sample, in accordance with Gerrits et al. (2001).

The second set of statistical analyses involved comparisons of frequency distributions between the response and the non-response group first by using the Kolmogorov-Smirnov test. Then the Responders-Nonresponders ratio was calculated for increasing cut-off scores (for the 75th, 90th, 95th, 98th, and the 99.5th percentile). The relative risk of the response-non response ratio was calculated to examine if there were overall differences in the ratio between the two groups and if this ratio increased or decreased with increasing percentile scores. The next step involved conducting binary logistic regression analyses to examine if teacher-screen scores on each of the subscales were associated with increased odds ratios (ORs) for non-response.

For the regression analyses, children were divided into three groups; Zero- (zero score), Moderate (score above zero but below the 95th percentile score) and High (equal to or higher than within the 95th percentile) symptom scorers for each symptom subscale score. In addition, the children were divided into two group on the basis of whether the teacher found that the child's problems had definite or severe impact on one or more of the five questions concerning the child's everyday life (Impact group) or that the child were found to have no or minor problems (No impact group).

The 95th percentile for the symptom score was judged to be appropriate cut-off criteria for High symptom group in light of the estimated prevalence rates for the different diagnostic categories reported for this

sample elsewhere (Heiervang et al., 2007). The frequency distribution on all subscales was highly skewed, where around 50% of the children had zero problem/symptom scores on the subscales. Thus, the children in the Moderate symptom group comprised individuals with scores from around the 50th to the 95th percentile on all subscales.

Two sets of regressions were conducted. The first sets involved analyses where Impact was entered at the first step and symptom score was entered at the second step with nonresponse as outcome measure. Thus ORs for Non-response were calculated on the basis of Impact (with No-impact as reference) and Symptom (Moderate and High symptom group with Zero symptom group as reference). The second sets were conducted to examine if there were different relations between symptom-scores and nonresponse for children with and without impact, yielding ORs for non-response for Moderate and High scorers with Zero symptom group as reference, stratified for the impact status for the children.

The Zero symptom group served as a reference for the estimation of ORs since this was the largest group on most subscales, but we calculated separate logistic regressions using the Moderate symptom group as reference in order to estimate the ORs for the High symptom group which could not be extracted from the logistic regressions described above.

Since this approach to the study of nonresponse bias involved multiple comparisons of the same set of data, a more conservative alpha

level of $p=.01$ was used for all tests, to compensate for the risk of inflated p -values.

Results

Non-response bias estimates

The teacher-screen yielded significantly higher scores for non-responders than Responders on all subscales, except for the tics subscale, in the questionnaire (all other comparisons $p<.001$). Effect sizes for most subscales were small with moderate effect size (according to Cohen, 1992), for only one subscale; the SNAP-Inattention subscale ($d=.303$). Correcting for non-response proportion the bias estimates for all mean scores were low (.016 to .084; see Figure 1).

Insert Figure 1 about here

All subscale scores were significantly correlated in both groups (see Table 1; all correlations $p. <.001$). There were no substantial differences between the correlations for the response- and the non-response group, neither in number of large size effect correlations (according to Rosenthal

and Rosenow, 1991) nor in the estimates of bias in correlations introduced by the non-response group (.001 to .045).

Insert Table 1 about here

Frequency distribution

Kolmogorov-Smirnov Z-test yielded statistically significant differences in the distribution across Responders and Nonresponders for all subscales, except for the tics and the OCD subscales (all other comparisons $p < .001$).

The relative risk for the nonresponse/response group (NR/R) ratio was statistically significant for the 75th, the 90th and the 95th percentile for all subscale scores, and for the 98th percentile for most subscale scores, except for the SDQ-emotion, SDQ-hyper and Tics subscales. For the 99.5th percentile the relative risk for the NR/R ratio score was not statistically significant, except for the ASSQ, OCD, SNAP-Inattention and SNAP-Hyperactivity subscale scores. For these four subscales, the NR/R ratio actually increased compared to the other percentiles see Figure 2).

Insert Figure 2 about here

The binary logistic regression where Impact was entered at the first and Symptom at the second step yielded a main effect of impact for all subscales; $p < .001$; see Table 2), demonstrating that children whose mental health problems were described to have an impact on the children's everyday life, were significantly more likely to not participate in the study than children with no impact. There were also an effect of symptom-score on all ($p < .01$) except the OCD and Tics subscales and an interaction effect between impact and symptom score for the SDQ-emotional and SDQ-conduct (both comparisons $p < .01$).

For the children with no impact, the likelihood of participation in the study was significantly related to teacher reports of problems/symptoms on all subscale scores (except the tics and OCD tics subscales). Specifically, children rated by their teacher to have moderate symptoms on most subscales and high symptom scores on emotional, hyperactivity and attentional problems were more likely than children with zero symptoms to be Nonresponders. The children with high symptom scores on SNAP-Inattention ($OR = 1.98$ (1.09-3.61; $p < .01$)) was also more likely than children with moderate symptom scores to be Nonresponders.

For the children with impact, only teacher reports of peer problems and inattention in high symptom relative to zero symptom were associated with increased likelihood of nonresponse. For all other subscales, level of symptom scores was not associated with likelihood of participation.

Insert Table 2 about here

Discussion

The main findings were that teachers, prospectively, rated Nonresponders significantly higher on all symptom (except for the tics) subscales and with more impact than Responders. The effect size scores indicated, however, that the differences between Responders and Nonresponders were overall small in magnitude, except for the inattention subscale score. Second, the frequency distribution demonstrated that there were more Responders than Nonresponders with zero symptoms or problems, while there were more Nonresponders than Responders in the 70th-95th percentile. Finally, binary logistic regressions revealed that both impact and symptom scores from the teacher screen was related to the likelihood of nonresponse. For the children with impact, only inattention

and problems with peer relationships were associated with increased likelihood of non-participation.

Differences between Responders and Nonresponders in mean scores and/or correlations are considered necessary conditions for nonresponse bias to occur (e.g. Smith, 1983). In the present study there were no indications that the association between symptom and functional impairment scores were different in Responders and Nonresponders. Nonresponse was, however, associated with overall higher problem scores. However, the calculation of effect sizes on the basis of the mean score differences for the low nonresponse proportion, demonstrated that the bias introduced by nonresponse was ignorable, which accords with Gerrits et al's (2001) general conclusion that mean score (or correlation) differences does not necessarily indicate non-response bias.

In the present study we used an alternative approach to the estimate of nonresponse bias involving assessment of differences in frequency distribution between Responders and Nonresponders, thereby calculating ORs for nonresponse on the basis of a prospective teacher screen of almost all children in a total population of primary school children. We found that teachers rated relatively more of the Responders than the Nonresponders to be without symptoms on all subscales, and that the relation between number of children with zero and moderate symptoms were different between Responders and Nonresponders, except for the tics and the OCD subscales.

For all other problem areas children rated by their teacher to have moderate symptoms were more likely than children with zero symptoms to be Nonresponders in the No impact group. In addition children rated by their teachers to have symptoms of emotional problems, inattention and hyperactivity were also less likely to obtain parental consent for participation.

For children with impact, the relation between symptom scores and nonresponse was only evident for the subscales of peer problems and inattention, reflecting that high symptom scores on these two subscales were more likely than children with no symptoms to be Nonresponders. The fact that this association was only evident for two subscales probably reflects the magnitude of impact in itself as a predictor for nonresponse. However, it also suggests that for peer relationship problems and symptoms of inattention, impact and symptom scores had an additive effect in predicting nonresponse.

The results from the logistic regression approach to nonresponse bias demonstrated more differences between Responders and Nonresponders than did the comparisons using standardized mean scores. The reason nonresponse more consistently could be predicted from symptomscores in the No-impact than the Impact group is probably due to the fact that the impact score in itself was a predictor for nonresponse, while the relation

between nonresponse and symptomscore would be more evident in the larger and more heterogeneous No impact group.

Nonresponse was also related to problem areas. Interestingly, both symptoms associated with the diagnostic groups with the highest (attentional deficit/hyperactivity disorder; ADHD) and the lowest (emotional disorders) level of service use among the Responders in this study were related to nonresponse. If nonresponse in community surveys of mental health problems in children simply reflected parental reluctance to disclose problems in their children, one would not expect that the largest nonresponse bias should arise to symptoms related to ADHD, which was the diagnostic group with the highest level of service use (Heiervang et al., 2007). One could therefore speculate if nonresponse could be due to other aspects of the children's symptoms. As an example, ADHD symptoms may make it more difficult for families to organize everyday life, thereby making it more difficult for parents to return the consent form.

The relation between emotional problems and nonresponse could, on the other hand more readily be attributed to parents' reluctance to disclose such problems, since we (Heiervang et al , 2007) found parental reports of low health service use for children with emotional disorders. This is probably related to how these services are organized, but could also indicate the extent to which parents fail to recognize emotional problems in children or that they are not perceived to be disruptive for the child. In light of this it

was interesting to note that teachers described more symptoms of emotional problems in the children in the nonresponse group, since this could suggest that parental reluctance to participate in fact is related to worries about emotional problems in children. Finally, there was also an association between nonresponse and peer relationship problems, which accords with Noll et al (1997), suggesting that parents who perceive their children to have poor social skills or little social acceptance by their peers, are more reluctant to participate.

Both the effect size estimates and the distributional approach suggested little non-response bias from teacher ratings of OCD and tics. One should, however, note that these conditions may be difficult to screen for in young children. The problem areas of conduct /oppositional defiance and autism spectrum symptoms non-response yielded only small bias effects, in that Moderate (but not High) symptom scorers were more likely than Zero symptom scorers to be Nonresponders.

A limitation of this study is that it did not include demographic information about the children. Thus it is not possible to evaluate if the association between indices of mental health problems and nonresponse bias could be accounted for by differences in the characteristics of Responders and Nonresponders. Another limitation is that we did not probe Nonresponders for their reasons to not participate. Therefore it was not possible to determine if magnitude of bias differed between subjects whose

motivations for not participating was related to the topic of the study compared to those who refused to participate for other reasons.

The main implications of this work is that it demonstrates how an approach based on frequency distributions and logistic regressions will yield a different picture than the usual approach based on calculations of mean differences and effect sizes. A logistic regression approach also represents the possibility to compensate for nonresponse bias in terms of calculating probably weights for the children in terms of how representative they are for the target population, thereby adjusting for the nonresponse bias in estimates of e.g. prevalence rates.

The main conclusion from this study was that nonresponse in a community survey of children's mental health problems could be predicted from a prospective teacher screen. Children whose mental health problems were judged to have an impact on everyday life were more likely to not obtain parental consent for participation. In addition, children with emotional problems, ADHD and peer relationship problems were more likely to not participate, suggesting that nonresponse was related to problem area. In contrast, effect sizes estimates of mean score differences and correlations suggested only minimal non-response bias in this study. This argues in favour of a multi-analytic approach in assessing the impact of nonresponse in population surveys of mental health problems in children.

References

- Apter, A., Pauls D.L., Bleich A., Zohar A.H., Kron S., Ratzoni G., Dycian A., Kotler, M., Weizman A., Gadot N. (1993) An epidemiologic study of Gilles de la Tourette's syndrome in Israel. *Archives of General Psychiatry*, 50, 734-8.
- Cohen, J. (1992). Quantative methods in psychology. A power primer. *Psychological Bulletin*, 112, 155-159.
- Ehlers S., Gillberg C., Wing L. (1999) A screening questionnaire for Asperger syndrome and other high- functioning autism spectrum disorders in school age children. *Journal of Autism and Developmental Disorders*, 29, 129-141.
- Frame, C.L. Strauss, C.C (1987). Parental informed consent and sample bias in grade-school children. *Journal of Social and Clinical Psychology*, 5, 227-236.
- French, D.D. & Waas, G.A. (1985). Teachers' ability to identify peer-rejected children: A comparison of sociometrics and teacher ratings. *Journal of School Psychology*, 23, 347-353.
- Gerrits, M.H., van den OOrd, E.J.C.G, & Voogt, R. (2001). An evaluation of nonresponse bias in peer, self and teacher ratings of children's psychosocial adjustment. *Journal of Child Psychology and Psychiatry*, 42, 593-602.

- Goodman, R. (1999). The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child and Adolescent Psychiatry, 40*, 791-801.
- Heiervang, E., Stormark, K.M., Lundervold, A.J., Heimann, M., Goodman, R., Posserud, M-B., Ullebø, A.K., Plessen, K.J., Bjelland, I, Lie, S.A., & Gillberg, C (2007). Psychiatric disorders in Norwegian 8- to 10-year olds: An epidemiological survey of prevalence, risk factors and service use. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 438-447.
- McMiller, W.P., & Weisz, J.R. (1996). Help-Seeking Preceding Mental Health Clinic Intake among African-American, Latino, and Caucasian Youths. *Journal of the American Academy of Child and Adolescent Psychiatry. 35*, 1086-1094.
- Noll, R.B., Zeller, M.H., Vanatta, K., Bukowski, WM & Davies, WH (1997). Potential bias in classroom research: Comparison of children with permission and those who do not receive permission to participate. *Journal of Clinical Child Psychology, 26*, 36-42.
- Pavuluri, M.N., Luk, S-L., McGee, R. (1996). Help-Seeking for Behavior Problems by Parents of Preschool Children: A Community Study. *Journal of the American Academy of Child & Adolescent Psychiatry. 35*(2):215-222.

- Roberts, R.E, Atkinsson, C.C., Rosenblatt, A. (1998). Prevalence of Psychopathology among Children and Adolescents
The American Journal of Psychiatry, 155, 715-727.
- Rosenthal, R. & Rosenow, R.W (1991). *The essentials of behaviour research: Methods and data analysis*. McGraw-Hill; New York.
- Smith, T.W. (1983). The hidden 25 percent: An analysis of nonresponse on the 1980 general social survey. *Public Opinion Quarterly*, 47, 386-404.
- Stang, A. (2003). Non-response research. An underdeveloped field in epidemiology. *European Journal of Epidemiology*, 18, 929-932.
- Swanson, J.M. (1992). *School-based assessments and interventions for ADD students*. K.C. Publishing, Irvine, CA.
- Thomsen, P.H. (1998). Obsessive compulsive disorder in children and adolescents. Clinical guidelines. *European Child and Adolescent Psychiatry*, 7, 1-11.
- Weinberger, DA., Tublin, S.K, Ford, M.E, Feldman, S.S (1990). Preadolescents social-emotional adjustment and selective attrition in family research. *Child Development*, 61, 1374-1386.

Figure captions

Figure 1 Mean score differences between Responders and Nonresponders for each of the subscales. Number in brackets refers to the estimated nonresponse bias= $p_n(y_r - y_n)$, calculated as a function of proportion of nonrespondents (p_n) and effect size (mean score difference between respondents (y_r) and nonrespondents (y_n , see Gerrits et al., 2001).

Subscales:

SDQ Total difficulties score (20 items, which again is based on SDQ-Emotional symptoms, SDQ-Conduct problems, SDQ-Hyperactivity/inattention, SDQ-Peer relationship problems (all five items each). Autism Spectrum Screening Questionnaire (ASSQ; 27 items), Obsessive Compulsive (OCD) problems (5 items), Tics (5 items), Swanson, Noland and Pelham (SNAP) checklist for symptoms of Oppositional Defiance Disorder (ODD; 8 items), SNAP-Inattention (9 items) and the SNAP- hyperactivity (9 items), Impact score (5 items).

Figure 2. The ratio between Nonresponders and responders for the 75th, 90th, 95th, 98th and the 99.5th percentiles for each of the subscale scores. Small bars= 99th confidence interval.

Acknowledgement.

The authors want to thank Stein Atle Lie for valuable inputs to the statistical analyses. We thank the whole Bergen Child Study group for their involvement in carrying out the study. Finally we thank all teachers who filled out the questionnaire. Without the high participation rate this study would not have been possible.

The study was financially supported by the Norwegian Research Council, The Norwegian Directorate for Health and Social Affairs, Western Norwegian Regional Health Authority, the L. Meltzer legacy for the University of Bergen, and the City of Bergen.

Table 1

Intercorrelations between Subscales for Responders (below diagonal) and Nonresponders (above diagonal)

	SDQ- Total	SDQ- Emotion	SDQ- Conduct	SDQ- Hyper	SDQ- Peer	ASSQ	OCD	Tics	SNAP- ODD	SNAP- Inattention	SNAP- Hyper	Impairment
SDQ-Total												
Correlation		.61	.73	.83	.75	.74	.39	.37	.69	.75	.70	.76
Bias		.005	.001	.005	.011	.005	.029	.0003	.008	.007	.007	.004
SDQ-Emotion												
Correlation	.63		.26	.25	.45	.44	.38	.24	.31	.35	.23	.40
Bias	.005		.02	.002	.006	.009	.002	.004	.021	.002	.0004	.014
SDQ-Conduct												
Correlation	.77	.34		.55	.41	.49	.20	.22	.78	.51	.60	.58
Bias	.001	.02		.003	.011	.018	.014	.012	.005	.005	.007	.017
SDQ-Hyper												
Correlation	.82	.26	.55		.41	.53	.25	.30	.56	.77	.72	.64
Bias	.005	.002	.003		.022	.017	.032	.005	.006	.007	.005	.004
SDQ-Peer												
Correlation	.72	.44	.45	.36		.71	.34	.34	.40	.48	.38	.60
Bias	.011	.006	.011	.022		.015	.034	.019	.013	.025	.016	.012
ASSQ												
Correlation	.73	.47	.55	.50	.68		.47	.60	.54	.62	.55	.70
Bias	.005	.009	.018	.002	.015		.039	.02	.012	.008	.0003	.001
OCD												
Correlation	.32	.39	.17	.17	.26	.37		.37	.25	.34	.28	.37
Bias	.029	.002	.014	.032	.034	.039		.045	.015	.041	.025	.032
Tics												
Correlation	.37	.23	.26	.32	.28	.54	.24		.23	.34	.30	.37
Bias	.0003	.004	.012	.005	.019	.02	.045		.01	.005	.027	.001
SNAP-ODD												
Correlation	.72	.38	.78	.55	.45	.61	.21	.28		.52	.67	.62
Bias	.008	.021	.005	.006	.013	.02	.018	.01		.007	.009	.019
SNAP-Inattention												
Correlation	.74	.36	.53	.75	.43	.61	.25	.35	.55		.62	.71
Bias	.007	.002	.004	.007	.025	.008	.04	.005	.007		.005	.001
SNAP-Hyper												
Correlation	.68	.24	.59	.71	.35	.56	.22	.38	.65	.64		.60
Bias	.007	.0004	.007	.005	.016	.0003	.025	.027	.009	.005		.006
Impairment												
Correlation	.78	.45	.63	.63	.57	.71	.29	.37	.68	.72	.63	
Bias	.004	.014	.017	.004	.012	.001	.018	.001	.009	.001	.006	

Note Bias refers to Nonresponse bias using effect size estimate(the difference in correlation between the respondents and the target population ; $r_t - r$; see Gerrits et al., 2001)

Table 2. Odds ratios for nonresponse for each symptom score

Table 2: Odds ratios for nonresponse for each symptom score													
	<u>SDQ-Total</u>			<u>SDQ-Emotion</u>		<u>SDQ-Conduct</u>		<u>SDQ-Hyper</u>		<u>SDQ-Peer</u>		<u>ASSQ</u>	
	N	n	OR (CI 99%)	n	OR (CI 99%)	n	OR (CI 99%)	N	OR (CI 99%)	n	OR (CI 99%)	n	OR (CI (99%))
Impair.													
No	7799	2026	1 (ref)	2026	1 (ref)	2026	1 (ref)	2026	1(ref)	2026	1(ref)	2026	1(ref)
Yes	1356	494	1.71(1.40-2.07)	494	1.74(1.46-2.08)	494	1.71(1.42-2.07)	494	1.69(1.40-2.04)	494	1.50(1.24-1.81)	494	1.56(1.28-1.91)
Symptom													
<i>No impair.</i>													
Zero	7403	352	1 (ref)	1478	1 (ref)	1494	1 (ref)	582	1(ref)	1380	1(ref)	1115	1(ref)
Moderate	1645	1633	1.49(1.26-1.77)	472	1.21(1.07-1.37)	499	1.31(1.12-1.54)	1415	1.39(1.20-1.61)	627	1.45(1.25-1.68)	891	1.27(1.11-1.45)
High	229	11	1.44(0.57-3.61)	76	1.50(1.13-1.98)	33	1.74(0.98-3.06)	29	2.36(1.26-4.42)	19	1.50(0.73-3.09)	20	1.97(0.94-4.14)
<i>Impair.</i>													
Zero	481	1	1 (ref)	174	1 (ref)	123	1(ref)	11	1(ref)	84	1(ref)	32	1(ref)
Moderate	584	313	0.65(0.17-24.8)	214	0.90(0.70-1.17)	253	0.98(0.68-1.42)	333	1.55(0.60-4.01)	274	1.40(0.94-2.09)	287	1.10(0.60-2.01)
High	291	180	0.66(0.17-25.4)	106	0.82(0.60-1.19)	118	0.76(0.50-1.16)	150	1.40(0.53-3.67)	136	1.57(1.01-2.47)	175	1.25(0.67-2.33)

	<u>OCD</u>		<u>Tics</u>		<u>SNAP-ODD</u>		<u>SNAP-Inattention</u>		<u>SNAP-Hyper</u>	
	n	OR (CI 99%)	n	OR (CI 99%)	n	OR (CI 99%)	N	OR (CI 99%)	n	OR (CI (99%))
Impair.										
No	2026	1 (ref)	2023	1 (ref)	2023	1 (ref)	2023	1(ref)	2023	1(ref)
Yes	494	1.79(1.51-2.13)	494	1.86(1.56-2.21)	494	1.69(1.40-2.05)	494	1.39(1.13-1.70)	494	1.67(1.38-2.01)
Symptom										
<i>No impair.</i>										
Zero	1866	1 (ref)	1958	1 (ref)	1386	1 (ref)	1138	1(ref)	582	1(ref)
Moderate	122	1.19(0.89-1.58)	40	1.09(0.67-1.77)	627	1.39(1.20-1.61)	627	1.37(1.19-1.57)	1415	1.28(1.09-1.49)
High	38	0.99(0.69-1.62)	25	1.05(0.57-1.93)	10	0.78(0.31-1.94)	19	2.71(1.49-4.92)	29	1.78(0.99-3.20)
<i>Impair</i>										
Zero	328	1 (ref)	357	1 (ref)	116	1(ref)	25	1(ref)	118	1(ref)
Moderate	90	1.09(0.74-1.62)	61	1.23(0.77-1.99)	251	0.73(0.50-1.07)	265	1.73(0.92-3.27)	254	1.19(0.82-1.72)
High	76	1.29(0.84-2.00)	76	0.89(0.59-1.33)	127	0.80(0.52-1.24)	204	2.12(1.11-4.04)	122	0.95(0.62-1.44)

Note. Grand mean (N) reflects overall sample size. Small n's refer to sample size of Nonresponders within each symptom subscale.

Abbreviations: OR=odds ratio. CI=confidence interval. SDQ=Strengths and Difficulties Questionnaire. Total= Total Difficulties. Emotion=Emotional symptoms. Conduct=Conduct problems. Hyper=Hyperactivity. Peer= Peer relationship problems. ASSQ=Autism Spectrum Screening Questionnaire. OCD=Obsessive Compulsive Disorder. SNAP=Swanson. Nolan & Pelson rating scale. ODD=Oppositional Defiance Disorder. Hyper=Hyperactivity. Impair.=impairment

Figure 1

percentiles