

# Prevalence and Time Trends of Self-Reported Mental Health Problems Among Children and Adolescents Between 11 and 17 Years in the KiGGS Study

Franz Baumgarten<sup>1</sup>, Stephan Junker<sup>2</sup> , and Robert Schlack<sup>2</sup>

<sup>1</sup> Hochschule des Bundes für öffentliche Verwaltung, Fachbereich Sozialversicherung, Berlin, Deutschland

<sup>2</sup> Robert Koch Institute, Berlin, Deutschland

**Abstract:** *Objective:* Knowledge about the prevalence of mental disorders in childhood and adolescence is important for clinicians and policy-makers. This study examines the prevalence and trends in self-reported mental health problems among 11- to 17-year-olds in Germany. *Method:* We evaluated data from the self-report version of the Strength and Difficulties Questionnaire (SDQ) of 6,725 children and adolescents from the baseline of the German Health Interview and Examination Survey (KiGGS, 2003–2006) and 6,145 from its second wave (KiGGS wave 2, 2014–2017). *Results:* According to the SDQ total difficulties score, the prevalence estimates did not vary significantly between the study waves, neither regarding the category “abnormal” (9.3 % vs. 9.4 %) nor the pooled categories “borderline/abnormal” (16.9 % vs. 15.4 %). We confirmed the results by linear regression analyses using mean values instead of the SDQ categories. Analyses of the SDQ subscales revealed gender and age-specific time trends. *Conclusions:* These findings differ from those based on the SDQ parent report, which suggests significant declines in symptom load between the study waves. The results indicate the importance of integrating youth self-reports when measuring mental health problems, at least as part of a multi-informant approach.

**Keywords:** mental health problems, children and adolescents, KiGGS, Strength and Difficulties Questionnaire, prevalence and time trends

## Prävalenz und zeitliche Trends selbstberichteter psychischer Probleme bei Kindern und Jugendlichen im Alter zwischen 11 und 17 Jahren in der KiGGS-Studie

**Zusammenfassung:** *Fragestellung:* Für Behandelnde und politische Entscheidungsträger ist das Wissen über die Prävalenz psychischer Störungen im Kindes- und Jugendalter von Bedeutung. Das Ziel dieser Studie ist es, die Prävalenz und Trends bei selbstberichteten psychischen Problemen von 11- bis 17-Jährigen in Deutschland abzubilden. *Methodik:* Ausgewertet wurden die Daten des selbstberichteten Strength and Difficulties Questionnaire (SDQ) von 6725 Kindern und Jugendlichen aus der KiGGS-Basiserhebung (2003–2006) und 6145 Kindern und Jugendlichen aus KiGGS Welle 2 (2014–2017). *Ergebnisse:* Die Prävalenzen des SDQ-Gesamtschwierigkeitsscore zwischen KiGGS-Basis und KiGGS-Welle 2 unterschieden sich weder in der Kategorie „auffällig“ (9,3 % vs. 9,4 %) noch in der Kategorie „grenzwertig/auffällig“ (16,9 % vs. 15,4 %). Die Ergebnisse wurden durch Regressionsanalysen bestätigt. Die Analyse der SDQ-Subskalen ergab geschlechts- und altersspezifische Unterschiede. *Schlussfolgerung:* Die Ergebnisse unterscheiden sich vom SDQ-Elternbericht, der auf eine signifikante Abnahme der Symptombelastung zwischen den beiden Wellen hinweist. Die Studie verdeutlicht, wie wichtig im Sinne verschiedener Informationsquellen die Berücksichtigung von Selbstberichten von Kindern und Jugendlichen bei der Messung von psychischen Problemen ist.

**Schlüsselwörter:** psychische Probleme, Kinder und Jugendliche, KiGGS, Strength and Difficulties Questionnaire, Prävalenz und Trends

## Introduction

The prevalence of mental health problems among children and adolescents in Germany is estimated to be between 15 % and 20 % (Barkmann & Schulte-Markwort, 2012; Klipker et al., 2018). Studies conducted to obtain preva-

lence rates have faced different methodological challenges. In addition to varying populations, instruments, or sample selection procedures, one important criterion to differentiate between such approaches is the type of informant reporting the mental health of the children and adolescents. Clinicians and other practitioners rely mainly on symptoms

of mental health reported by parents, teachers, or other caregivers rather than by the children or adolescents themselves (Barkmann & Schulte-Markwort, 2012; Bentley et al., 2019; Patalay et al., 2014). However, in pediatrics, there is a growing sense that the voice of children and adolescents should be given greater importance, in particular with regard to their mental health (Bentley et al., 2019; Deighton et al., 2014; Huang et al., 2014).

One way to approach self-reported mental health at the population level is to gather the data from the children and adolescents themselves within the framework of representative population-based studies such as the German Health Interview and Examination Survey for Children and Adolescents (KiGGS study), conducted by the Robert Koch Institute. The KiGGS study assessed the mental health of children and adolescents using, among others, the parent-rated and the self-rated version of the internationally established Strength and Difficulties Questionnaire (SDQ) (Goodman, 1997). The SDQ is a brief instrument for assessing mental health problems in children and adolescents which is internationally proven and suitable for large-scale studies. Previous publications on prevalence rates of mental health problems in children and adolescents in Germany were based on parent-reported SDQ. In particular, data indicated that one in five children currently suffers from mental health problems (Hölling et al., 2014). Moreover, data suggested a significant decrease in parent-reported mental health problems, from about 20% at KiGGS baseline to about 17% at KiGGS wave 2 (conducted between 2014 and 2017) (Klipker et al., 2018).

For several reasons, we cannot necessarily assume comparable trends between parent- and self-reported mental health problems cannot. In general, there is little to moderate correspondence between these different informants (Achenbach et al., 1987; de los Reyes et al., 2015). The correlation between parent-reported and self-reported SDQ in community samples ranges from  $r = .14$  to  $r = .58$  (Aebi et al., 2017; Goodman et al., 1998; van Roy et al., 2010; van Widenfelt et al., 2003; Vugteveen et al., 2018). The poor agreement between the child ratings and other sources of information can be attributed to a variety of causes. Children and adolescents may feel or behave differently in varying contexts. For example, a child may display conduct problems in school (e.g., aggressive behavior toward classmates) but not at home (e.g., affectionate interaction with siblings) (Aebi et al., 2017; Vaz et al., 2016). Even if feelings or behaviors are comparable across situations, the perception and awareness of the informants may differ. For example, parents may not easily gain insight into the psychological conditions of a socially withdrawn child with emotional problems. Also, van Roy et al. (2010) argued that differences between parent- and self-reports might lie in the fact that children and adolescents are likely

to notice and communicate problems that are less evident for and appear minor to parents. In addition, children and adolescents might rather report their momentary condition, whereas parents resonate a long-term perspective (Aebi et al., 2017). The lack of correspondence between parent- and self-report may also be a methodological issue. Although the parent- and self-report versions of the SDQ are composed of almost identical items that are summed up to the same scales, they might not measure exactly the same constructs or not in exactly the same way. Using data from KiGGS wave 1, Rogge et al. (2018) found configural but only partially scalar invariance between the parent- and self-report versions of the SDQ. This means that both versions of the SDQ measure the same construct (configural invariance) but on partly different scales (scalar invariance), making direct comparisons difficult.

In sum, information about the mental health of children and adolescents – particularly as provided by the SDQ – differs depending on whether the youth themselves or their parents are surveyed. To gain a complete understanding of the mental health of children and adolescents, it appears important to consider the perspective of the youth themselves. Thus, this study aims to report on the prevalence and time trends of self-reported mental health problems among children and adolescents in Germany between 11 and 17 years described in the KiGGS study.

## Methods

### Study Design and Sample

The Robert Koch Institute conducted the KiGGS study as part of the German health monitoring system, to date consisting of three waves: the KiGGS baseline study (2003–2006), KiGGS wave 1 (2009–2012), and KiGGS wave 2 (2014–2017). To ensure the representativeness of the KiGGS study and its three waves, we selected a total of 167 sample points, representing the settlement structure of the Federal Republic of Germany by state and municipality type. A random selection of addresses of children and adolescents was made at these sample points from the population registers of the relevant municipalities (Hoffmann et al., 2018).

Both the KiGGS baseline study and KiGGS wave 2 were conducted as combined examination and interview surveys, whereas KiGGS wave 1 was conducted as a telephone survey. Since it has been shown that the SDQ is sensitive to different survey modes (Erhart et al., 2009), we decided not to include KiGGS wave 1 in the analysis. The KiGGS baseline study provides nationally representative cross-sectional data for 17,641 German children and adolescents

(49.1% female; 50.9% male) aged between 0 to 17 years (the response rate of the KiGGS baseline study was 66.6%). KiGGS wave 2 also provides representative cross-sectional data for 15,023 children and adolescents (50.2% female; 49.8% male) in the same age range (the response rate of KiGGS wave 2 was 40.1%). Detailed information on the objectives, study design, and implementation of the KiGGS study can be found elsewhere (Kurth et al., 2008; Mauz et al., 2017). For this article, we analyzed data on 6,725 children and adolescents aged between 11 and 17 years from the KiGGS baseline study (49.9% female; 50.1% male) and data on 6,145 children and adolescents in the same age range from KiGGS wave 2 (47.7% female; 52.3% male).

The Ethics Committee of the Charité – Universitätsmedizin Berlin reviewed the KiGGS baseline survey (No. 101/2000), and the Ethics Committee of the Hannover Medical School reviewed KiGGS wave 2 (No. 2275–2014) from an ethical point of view and approved the studies. Participation in the studies was voluntary. We informed the parents with custody about the aims and contents of the studies as well as about data protection and received their written informed consent.

## Measures

The KiGGS study assessed the mental health problems of children and adolescents using the SDQ (Goodman, 1997). The self-report version of the SDQ was completed by children and adolescents aged between 11 and 17 years. The SDQ captures mental health problems on five subscales (5 items each): emotional symptoms (e.g., “I worry a lot”), peer relationship problems (e.g., “Other children or young people pick on me or bully me”), conduct problems (e.g., “I am often accused of lying or cheating”), hyperactivity/inattention (e.g., “I am constantly fidgeting or squirming”), and prosocial behavior (e.g., “I am kind to younger children”). Four of these subscales (prosocial behavior excluded) can be added to generate a total difficulties score.

Population-based normative data of the distribution of the SDQ total difficulties score and the subscale scores allow analysts to classify children and adolescents into the groups “abnormal,” “borderline,” and “normal” (Goodman, 1997). Becker et al. (2018) employed this classification to calculate the SDQ norm values for children and adolescents in Germany based on representative data from the KiGGS baseline. These norm values were used in the present study for the SDQ total score to classify children and adolescents into the categories “abnormal” and the pooled category “borderline/abnormal.” In addition, we computed the mean values of the SDQ total difficulties score and the SDQ subscales.

The analysis included gender (female, male) and age as additional variables. Age was taken into account in years or age groups (11- to 13-year-olds and 14- to 17-year-olds) depending on the type of analysis (continuous or categorical).

## Statistical Analysis

The available data were analyzed using two different approaches. In a first step, the proportions of children and adolescents in the category “abnormal” and the combined category “borderline/abnormal” as well as the corresponding 95% confidence intervals were calculated for the self-reported SDQ total difficulties score based on German norm data. We examined the prevalence separately for gender and the age groups 11 to 13 years and 14 to 17 years in the KiGGS baseline study and in KiGGS wave 2. We examined the differences between the waves of the KiGGS study for gender and age group using the Rao-Scott adjusted Pearson chi-square test. This categorical approach to analyze the data serves to establish comparability with publications on parent-reported mental health problems (Klipker et al., 2018) and to provide a tangible indicator for policymakers and other public stakeholders.

In the second step, we calculated the means, standard deviations, and 95% confidence intervals for the SDQ total difficulties score and each SDQ subscale. Further, linear regression analysis for the SDQ total difficulties score and the SDQ subscales as the dependent variable and gender, age (in years), and their interaction as an independent variable were calculated separately for the KiGGS baseline study and KiGGS wave 2. We report the unstandardized coefficients; thus, they can be interpreted as the associated change in the SDQ scales by 1 year increase in age or, respectively, as the difference between the genders. The coefficients of every single regression were then tested for significant differences between the KiGGS baseline study and KiGGS wave 2 by applying a Wald test. The model-based predictions for the SDQ and its subscales were plotted and used to illustrate the differences for gender and age between the KiGGS baseline study and KiGGS wave 2.

This continuous approach to analyzing the data can be used as a valid indicator of population-based mental health as well (Goodman & Goodman, 2011) and should overcome potential problems associated with categorizing SDQ scores using bandings. For instance, the percentile-based categorization into “normal,” “borderline,” and “abnormal” proposed by Goodman (1997) is based on an approximate estimation of real prevalence of mental health problems in populations. Not knowing these rates exactly can lead to an over- or underestimation. In addition, the computation of categorical variables from inter-

val-scaled variables is associated with a loss of information (Altman & Royston, 2006; Lydersen, 2015). This is especially true for the SDQ, as in most cases, when norm-based bandings are calculated, the raw values of the subscales are rounded to the nearest integer (Janitza et al., 2020). Analyzing the mean values of the SDQ also considers the ongoing debate of whether mental health is represented in distinguishable categories or reflects a continuum of experiences and behaviors (Helzer et al., 2006; Keyes, 2002).

To adjust possible deviations of the sample from the population structure in terms of age, gender, state, German citizenship, and distribution of education among parents (Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder, 2017), all analyses were carried out with a weighting factor. The calculations for trends between the KiGGS baseline study and KiGGS wave 2 were performed based on age-standardized prevalences (population on 31 December 2015). A difference between groups or waves was assumed to be statistically significant if the  $p$ -values were lower than .05. The statistical computations and the creation of graphs was done using R (R Core Team, 2019) in R Studio (RStudio Team, 2018).

## Results

### Categorical Analysis of the SDQ Total Difficulties Score

The proportions of children and adolescents and the respective 95% confidence interval for the SDQ total difficulties score in the categories “abnormal” and “borderline/abnormal” at the KiGGS baseline study and KiGGS wave 2 are reported in Table 1. The overall prevalence for the SDQ total difficulties score did not vary significantly between the two waves, neither for “abnormal” (9.3% at the KiGGS baseline study vs. 9.4% at KiGGS wave 2) nor for “borderline/abnormal” (16.9% at the KiGGS baseline study vs. 15.4% at KiGGS wave 2). However, there were significant differences by gender and age group in terms of a substantial decline of the symptom prevalence for girls aged between 11 and 13 years and boys aged between 14 and 17 years, but only for “borderline/abnormal.”

### Continuous Analysis of the SDQ Total Difficulties Score and the SDQ Subscales

The mean values, standard deviations, and 95% confidence intervals of the SDQ total difficulties score as well as the SDQ subscales in the KiGGS baseline study and KiGGS

wave 2 for gender and age are presented in the electronic supplementary material (ESM 1, Table) to this article. Figure 1 shows the predicted values based on the regression analysis for the SDQ total difficulties score and the SDQ subscales. The predicted values are displayed by age with different lines for the study waves and in separate graphs for girls and boys.

As to the SDQ total difficulties score, the linear regression analysis revealed a significant effect for age in boys at KiGGS wave 2 ( $\beta = -0.12$ ,  $p = .04$ ), whereby increasing age was associated with lower values. More importantly, there was a significant interaction of age and gender in the KiGGS baseline study ( $\beta = 0.19$ ,  $p = .01$ ) and KiGGS wave 2 ( $\beta = 0.33$ ,  $p < .01$ ). In both study waves, increasing age was associated with higher values for girls than boys. There were no significant differences between the KiGGS baseline study and KiGGS wave 2 for age, gender, and the respective interaction.

For emotional symptoms, there was a significant effect for gender in the KiGGS baseline study ( $\beta = 0.48$ ,  $p < .01$ ) and KiGGS wave 2 ( $\beta = 0.66$ ,  $p < .01$ ). In both waves, girls reported higher values than boys. Moreover, the interaction of gender and age was also significant in the KiGGS baseline study ( $\beta = 0.18$ ,  $p < .01$ ) and KiGGS wave 2 ( $\beta = 0.22$ ,  $p < .01$ ). In both study waves, the gender differences were more pronounced with increasing age. The analysis revealed no significant effects between the KiGGS baseline study and KiGGS wave 2.

For peer problems, the analysis revealed a significant effect for age ( $\beta = 0.05$ ,  $p = .03$ ) in boys and the interaction of gender and age ( $\beta = 0.07$ ,  $p = .02$ ) at KiGGS wave 2. Increasing age was associated with higher values, whereby this relationship was more pronounced among girls than boys. These results were only revealed in KiGGS wave 2, leading to a significant difference in the effect of age between the two study waves in boys ( $\beta = 0.07$ ,  $p = .01$ ) and girls ( $0.12$ ,  $p < .01$ ).

For conduct problems, there was a significant effect for gender in KiGGS wave 2 ( $\beta = -0.30$ ,  $p < .01$ ), with boys reporting higher values than girls. The analysis also showed an effect for age in the KiGGS baseline study ( $\beta = 0.04$ ,  $p = .02$ ) and KiGGS wave 2 ( $\beta = -0.05$ ,  $p < .01$ ), both in boys with no significant differences in girls. The effects of age are significantly different between the two study waves ( $\beta = -0.08$ ,  $p < .01$ ) for boys. While in the KiGGS baseline study, increasing age was related to higher values, in the KiGGS wave 2, increasing age was associated with lower values.

For hyperactivity/inattention, the linear regression showed a significant effect for age in the KiGGS baseline study ( $\beta = -0.06$ ,  $p < .01$ ) and KiGGS wave 2 ( $\beta = -0.14$ ,  $p < .01$ ) both for boys with no significant differences for girls. In both waves, increasing age was associated with lower



**Table 1.** Prevalence and trends for the proportions of children and adolescents and the respective 95 % confidence interval for the SDQ total difficulties score in the category “abnormal” and the combined category “borderline/abnormal” in the KiGGS baseline study (2003–2006) and KiGGS wave 2 (2014–2017)

		SDQ „abnormal“					SDQ „borderline/abnormal“				
		KiGGS baseline study		KiGGS wave 2		p-value	KiGGS baseline study		KiGGS wave 2		p-value
		%	95 % – CI	%	95 % – CI		%	95 % – CI	%	95 % – CI	
SDQ Total Difficulties Score	Total	9.3	8.4–10.2	9.4	8.4–10.4	.91	16.9	15.8–18.0	15.4	14.1–16.7	.82
	Girls	10.9	9.6–12.3	11.6	10.2–13.1	.46	20.0	18.3–21.6	18.7	16.7–20.6	.29
	11–13 years	9.8	7.9–11.8	8.7	6.6–10.7	.43	18.4	15.9–20.8	14.3	11.9–16.6	.02
	14–17 years	11.7	9.9–13.5	13.6	11.4–15.7	.17	21.1	18.7–23.4	21.6	19.0–24.2	.75
	Boys	7.7	6.6–8.9	7.2	5.8–8.7	.59	14.1	12.7–15.4	12.3	10.6–14.0	.11
	11–13 years	8.8	7.1–10.5	9.5	7.5–11.5	.60	15.2	13.1–17.4	15.6	13.1–18.1	.81
	14–17 years	7.0	5.5–8.5	5.6	3.9–7.4	.28	13.3	11.5–15.1	10.0	7.9–12.1	.03

values. There was a significant difference for the effect of age between the two study waves for boys ( $\beta = -0.08$ ,  $p = .02$ ) and girls ( $\beta = -0.07$ ,  $p = .048$ ), with this effect becoming more accentuated in the KiGGS wave 2.

For prosocial behavior, the linear regression analysis revealed a significant main effect for gender in the KiGGS baseline study ( $\beta = 0.85$ ,  $p < .01$ ) and KiGGS wave 2 ( $\beta = 0.74$ ,  $p < .01$ ). Girls reported higher values than boys. There were no significant differences between the KiGGS baseline study and KiGGS wave 2.

## Discussion

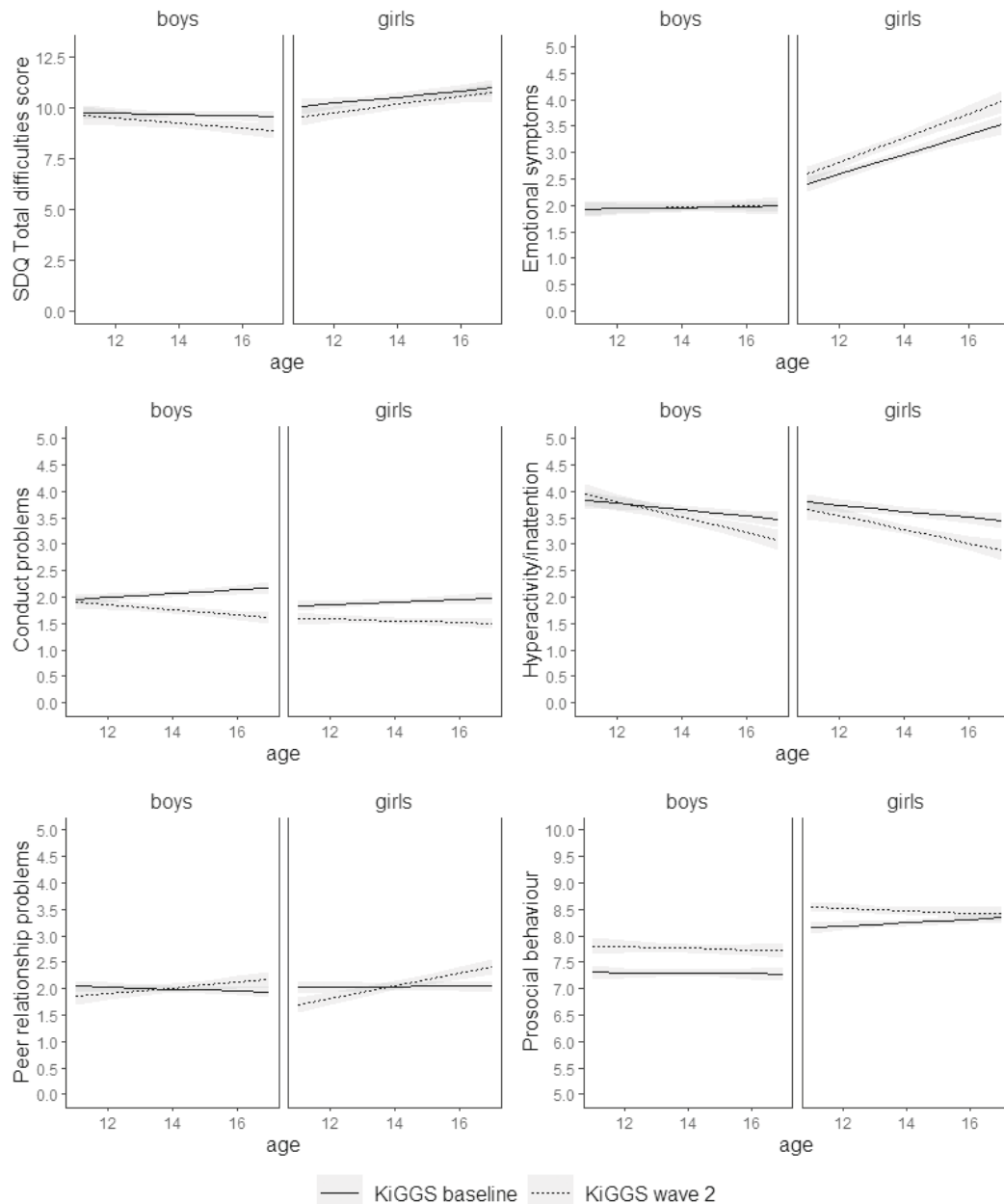
### Summary and Interpretation

This study reports on prevalence rates and time trends of self-reported mental health problems in children and adolescents aged between 11 and 17 years in Germany. Applying the categorical SDQ bandings “normal” and “borderline/abnormal,” the analysis showed that the overall prevalence of mental health problems according to the children’s self-report did not significantly decline from the KiGGS baseline study to KiGGS wave 2. However, there was a significant decline in specific age groups as a function of gender in the pooled category “borderline/abnormal.” The proportion of girls aged 11 to 13 classified as “borderline/abnormal” decreased by 4.1 percentage points between the two waves. Similarly, the prevalence of self-reported mental health problems in boys aged 14

to 17 declined by 3.3 percentage points. There were no significant declines by age group and gender for the category “abnormal” only. Thus, a decline in self-reported mental health problems between the KiGGS baseline study and KiGGS wave 2 seems not only to be determined by age and gender but also to be linked to the severity of symptoms. Goodman (1997) proposed applying the “borderline” category only in high-risk samples where false positives are less relevant. Considering that the population-based KiGGS study comprises a low-risk sample, and given the relatively small decline in the combined category “borderline/abnormal” as a function of gender and age, it does not seem appropriate to assume a substantial decline with respect to the self-reported SDQ total difficulties score among children and adolescents in Germany.

The continuous analysis strengthens the assumption that there are no time trends between the two waves, since no significant differences between the mean values of the SDQ total difficulties score in the KiGGS baseline study and KiGGS wave 2 were revealed. However, the regression-based evaluation of the SDQ subscales can provide a more detailed understanding of age and gender effects of mental health problems among children and adolescents in the KiGGS study.

Our analysis showed that gender effects were more evident on the subscales representing “internalizing” mental health problems (i.e., emotional symptoms and peer problems) and were broadly comparable between the two study waves. In the KiGGS baseline study as well as in KiGGS wave 2, girls reported more emotional symptoms than boys.



**Figure 1.** Predictive margins of the linear regression analysis for the SDQ total difficulties score and the SDQ subscales for boys and girls in the KiGGS baseline study (2003–2006) and KiGGS wave 2 (2014–2017).

In both study waves, the gender difference increased with age. With increasing age, girls also showed higher values in peer problems in KiGGS wave 2 than boys. Furthermore, our results illustrated that age effects were more likely to be observed on the subscales representing “externalizing” problems (i.e., conduct problems and hyperactivity/inattention), and at the same time, were also more different between the study waves. In the KiGGS baseline study, the mean values for conduct problems increased with rising age, whereas the same values decreased with increasing age

at KiGGS wave 2. Further, the decrease in mean values for the subscale hyperactivity/inattention with increasing age was significantly more evident in KiGGS wave 2.

Comparing KiGGS baseline and KiGGS wave 2, the analysis of the subscales in the continuous analyses of our study may indicate why there are fewer children and adolescents in the category “borderline/abnormal” of the SDQ total difficulties score only in the age groups of 11- to 13-year-old girls and 14- to 17-year-old boys. Despite the observed trend of a slight overall decline in the subscales

conduct disorder and hyperactivity/attention, the scores for emotional problems among girls aged 14 to 17 years are too high to cause a decrease in the SDQ total difficulties scores in KiGGS wave 2 that in turn would result in a declining category “borderline/abnormal.” Similarly, boys aged 11 to 13 years were not affected as much by the decline in the subscales conduct disorder and hyperactivity/attention observed in KiGGS wave 2, which resulted in an SDQ total difficulties score comparable to KiGGS baseline in this age group as well.

While this interpretation combines both statistical approaches of this paper, one result is independent of the type of analysis: Girls consistently reported more mental health problems than boys across both study waves. This gender gap becomes more evident with increasing age. On the one hand, the categorial analysis of the SDQ total difficulties score showed that the differences between girls and boys aged 14 to 17 years were larger than between girls and boys aged 11 to 13 years. When we applied the continuous analysis on the other hand, we found that, compared to boys, the SDQ total difficulties score for girls increased with age. Considering the time trends between the study waves, one result should be emphasized more strongly. The emotional problems of girls at a higher age increased in KiGGS wave 2. For example, the mean value of emotional problems for 16-year-old girls was 3.37 in the KiGGS baseline study (boys: 2.07) and 3.85 in KiGGS Wave 2 (boys: 2.08). This finding is in line with the results of other international studies, which describe an increase in internalizing psychological problems in girls within comparable periods (Aebi et al., 2017; Fink et al., 2015; Fleming et al., 2014; Gariépy & Elgar, 2016; Potrebny et al., 2019; van Vuuren et al., 2018).

A notable finding is that the results on time trends as well as on gender and age effects of self-reported mental health problems in children and adolescents differ from the data on parent-reported mental health problems from the KiGGS study (Klipker et al., 2018). The findings align with previous research results, which assume small to moderate agreement between parent- and self-reports (Achenbach et al., 1987; de los Reyes et al., 2015). Although our study does not allow a quantifiable comparison with previous publications on parent-reported mental health problems because of differences in methodological approaches, general differences in trend analysis between parent- and self-reports in the KiGGS study can be identified. Regarding the SDQ total difficulties score, there is no significant decline in self-reported psychological problems between the KiGGS baseline study and KiGGS wave 2. Further, boys showed more mental health problems in almost all age groups according to the parent-reported SDQ, whereas girls displayed higher problem scores in the SDQ self-report.

## Limitations

There are some limitations to consider in the present study. First, it is important to mention that mental health problems, as operationalized with the SDQ, are not equivalent to mental disorders. Rather, the questionnaire-based assessment results indicate who potentially belongs to a risk group in which mental disorders are more likely to occur (Hölling et al., 2014). Although using standardized clinical interviews in epidemiological research is costly, it could be of better use than merely symptom-based questionnaires to estimate the prevalence of mental disorders in children and adolescents at the population level.

Second, a further limitation of the current analysis is that it is impossible to identify further reasons for the observed trends and gender as well as age differences in self-reported mental health. For example, girls might perceive more challenges in modern society and have more difficulties developing appropriate coping strategies. It is also possible that merely the response behaviors between girls and boys are different. Girls might be more sensitive toward questions addressing their inner feelings or toward items reflecting their relationship with others.

Third, note that the regression analysis in this study assumed a linear relationship between the variables used. The variation of the mean values over different age groups suggested more complicated relationships that differed between the subscales. Thus, to prevent overfitting, we assumed a linear relationship. Nevertheless, the quality of other possible relations between mental health, age, and gender (e.g., quadratic, curvilinear) should be tested in future studies.

## Future Research and Conclusion

In addition, we would like to point out some directions for future research. Forthcoming research projects could focus on further explanations for the observed age and gender effects concerning self-reported mental health. One possibility would be to include additional sociodemographic variables (e.g., socioeconomic status) in regression-based longitudinal data analyses (Yoon et al., 2022). Moreover, different age- and gender-related response behaviors could be studied more intensively. A promising study by Patalay et al. (2018) already tapped into the understanding of items in self-reported mental health questionnaires as a function of age. Developing new instruments or adapting current questionnaires about mental health could also help entangle different response behaviors of children and their parents (Aebi et al., 2017). For example, items for parent-reported internalizing mental

health problems could rather focus on observable behavior (e.g., social isolation). Future research should also further investigate the gender- and age-related measurement invariance of the self-reported SDQ. Using the parent report, a study with a British sample has already found a gender-related measurement variance only for some age groups (Murray et al., 2022).

To obtain more detailed information on the prevalence and time trends of parent- and self-reported mental health in Germany, we need further analyses within the framework of the KiGGS study. For example, no study has yet compared parental and self-assessment data in the same age groups or considered the same definition of risk groups.

So far, the findings in our study lead to the question as to which type of reports should be preferred. Several studies indicate that parent reports are more informative about externalizing disorders, and that self-reports are more strongly associated with internalizing disorders (Aebi et al., 2017; Vugteveen et al., 2018). The reason for this is thought to be that children and adolescents have better access to their own and less observable emotional difficulties, while they themselves minimize or repress the problems given more obvious and socially undesirable behavioral problems.

In light of these results, it should not be a question of *which* measure is better suited for detecting mental health problems in children and adolescents. The arguments for why the two perspectives are different seem to be too comprehensible and, at the same time, hardly changeable. Rather, it is important to further implement a multi-informant approach, which has already been suggested as the best possible standard (Kaurin et al., 2016; Martel et al., 2017). This could, for example, ensure that research reports about mental health problems not only integrate the parents' perspective but also consider the children's perspective.

## Electronic Supplementary Material

The electronic supplementary material (ESM) is available with the online version of the article at <https://doi.org/10.1024/1422-4917/a000936>

**ESM 1.** Means and standard deviation of the SDQ scores in KiGGS baseline and KIGGS wave 2 (Table)

## References

Achenbach, T.M., McConaughy, S.H., & Howell, C.T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101(2), 213–232.

- Aebi, M., Kuhn, C., Banaschewski, T., Grimmer, Y., Poustka, L., Steinhausen, H.-C., & Goodman, R. (2017). The contribution of parent and youth information to identify mental health disorders or problems in adolescents. *Child and Adolescent Psychiatry and Mental Health*, 11(1), 23. <https://doi.org/10.1186/s13034-017-0160-9>
- Altman, D.G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, 332(7549), 1080. <https://doi.org/10.1136/bmj.332.7549.1080>
- Barkmann, C., & Schulte-Markwort, M. (2012). Prevalence of emotional and behavioural disorders in German children and adolescents: A meta-analysis. *Journal of Epidemiology and Community Health*, 66(3), 194–203. <https://doi.org/10.1136/jech.2009.102467>
- Becker, A., Wang, B., Kunze, B., Otto, C., Schlack, R., Hölling, H. et al. (2018). Normative data of the self-report version of the German Strengths and Difficulties Questionnaire in an epidemiological setting. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 46, 523–533. <https://doi.org/10.1024/1422-4917/a000589>
- Bentley, N., Hartley, S., & Bucci, S. (2019). Systematic review of self-report measures of general mental health and well-being in adolescent mental health. *Clinical Child and Family Psychology Review*, 22(2), 225–252. <https://doi.org/10.1007/s10567-018-00273-x>
- de los Reyes, A., Augenstein, T.M., Wang, M., Thomas, S.A., Drabick, D.A.G., Burgers, D.E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, 141(4), 858–900. <https://doi.org/10.1037/a0038498>
- Deighton, J., Croudace, T., Fonagy, P., Brown, J., Patalay, P., & Wolpert, M. (2014). Measuring mental health and well-being outcomes for children and adolescents to inform practice and policy: A review of child self-report measures. *Child and Adolescent Psychiatry and Mental Health*, 8, 14. <https://doi.org/10.1186/1753-2000-8-14>
- Erhart, M., Wetzel, R.M., Krügel, A. & Ravens-Sieberer, U. (2009). Effects of phone versus mail survey methods on the measurement of health-related quality of life and emotional and behavioural problems in adolescents. *BMC Public Health*, 30(9), 491. <https://doi.org/10.1186/1471-2458-9-491>
- Fink, E., Patalay, P., Sharpe, H., Holley, S., Deighton, J., & Wolpert, M. (2015). Mental health difficulties in early adolescence: A comparison of two cross-sectional studies in England from 2009 to 2014. *The Journal of Adolescent Health*, 56(5), 502–507. <https://doi.org/10.1016/j.jadohealth.2015.01.023>
- Fleming, T.M., Clark, T., Denny, S., Bullen, P., Crengle, S., Peiris-John, R., Robinson, E., Rossen, F.V., Sheridan, J., & Lucassen, M. (2014). Stability and change in the mental health of New Zealand secondary school students 2007–2012: Results from the national adolescent health surveys. *The Australian and New Zealand Journal of Psychiatry*, 48(5), 472–480. <https://doi.org/10.1177/0004867413514489>
- Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder. (2017). *Mikrozensus, 2013* [Microcensus 2013]. Retrieved from <http://www.forschungsdatenzentrum.de/besta/mikrozensus/>
- Gariépy, G., & Elgar, F.J. (2016). Trends in psychological symptoms among Canadian adolescents from 2002 to 2014: Gender and socioeconomic differences. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, 61(12), 797–802. <https://doi.org/10.1177/0706743716670130>
- Goodman, A., & Goodman, R. (2011). Population mean scores predict child mental disorder rates: Validating SDQ prevalence estimators in Britain. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 52(1), 100–108. <https://doi.org/10.1111/j.1469-7610.2010.02278.x>



- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 38(5), 581–586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>
- Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child & Adolescent Psychiatry*, 7(3), 125–130. <https://doi.org/10.1007/s007870050057>
- Helzer, J.E., Kraemer, H.C., & Krueger, R.F. (2006). The feasibility and need for dimensional psychiatric diagnoses. *Psychological Medicine*, 36(12), 1671–1680. <https://doi.org/10.1017/S003329170600821X>
- Hoffmann, R., Lange, M., Butschalowsky, H., Houben, R., Schmich, P., Allen, J., Kuhnert, R., Schaffrath Rosario, A., & Gößwald, A. (2018). KiGGS Wave 2 cross-sectional study: Participant acquisition, response rates and representativeness. *Journal of Health Monitoring*, 3(1), 78–91. <https://doi.org/10.17886/RKI-GBE-2018-032>
- Hölling, H., Schlack, R., Petermann, F., Ravens-Sieberer, U., Mauz, E., & KiGGS Study Group. (2014). Psychische Auffälligkeiten und psychosoziale Beeinträchtigungen bei Kindern und Jugendlichen im Alter von 3 bis 17 Jahren in Deutschland – Prävalenz und zeitliche Trends zu 2 Erhebungszeitpunkten (2003–2006 und 2009–2012) [Psychopathological problems and psychosocial impairment in children and adolescents aged 3–17 years in the German population: prevalence and time trends at two measurement points (2003–2006 and 2009–2012)]. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 57(7), 807–819. <https://doi.org/10.1007/s00103-014-1979-3>
- Huang, I.-C., Revicki, D.A., & Schwartz, C.E. (2014). Measuring pediatric patient-reported outcomes: Good progress but a long way to go. *Quality of Life Research*, 23(3), 747–750. <https://doi.org/10.1007/s11136-013-0607-2>
- Janitza, S., Klipker, K., & Hölling, H. (2020). Age-specific norms and validation of the German SDQ parent version based on a nationally representative sample (KiGGS). *European Child & Adolescent Psychiatry*, 29(2), 123–136. <https://doi.org/10.1007/s00787-019-01337-1>
- Kaurin, A., Egloff, B., Stringaris, A., & Wessa, M. (2016). Only complementary voices tell the truth: A reevaluation of validity in multi-informant approaches of child and adolescent clinical assessments. *Journal of Neural Transmission (Vienna, Austria: 1996)*, 123(8), 981–990. <https://doi.org/10.1007/s00702-016-1543-4>
- Keyes, C.L.M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior*, 43(2), 207–222.
- Klipker, K., Baumgarten, F., Göbel, K., Lampert, T., & Hölling, H. (2018). Psychische Auffälligkeiten bei Kindern und Jugendlichen in Deutschland – Querschnittergebnisse aus KiGGS Welle 2 und Trends [Mental health problems in children and adolescents in Germany – Results of the cross-sectional KiGGS Wave 2 study and trend]. *Journal of Health Monitoring*, 3(3), 37–45. <https://doi.org/10.17886/RKI-GBE-2018-077>
- Kurth, B.-M., Kamtsiuris, P., Hölling, H., Schlaud, M., Dölle, R., Ellert, U., Kahl, H., Knopf, H., Lange, M., Mensink, G.B.M., Neuhauser, H., Rosario, A.S., Scheidt-Nave, C., Schenk, L., Schlack, R., Stolzenberg, H., Thamm, M., Thierfelder, W., & Wolf, U. (2008). The challenge of comprehensively mapping children's health in a nation-wide health survey: Design of the German KiGGS-Study. *BMC Public Health*, 8(1), 196. <https://doi.org/10.1186/1471-2458-8-196>
- Lydersen, S. (2015). Statistical review: Frequently given comments. *Annals of the Rheumatic Diseases*, 74(2), 323–325. <https://doi.org/10.1136/annrheumdis-2014-206186>
- Martel, M.M., Markon, K., & Smith, G.T. (2017). Research Review: Multi-informant integration in child and adolescent psychopathology diagnosis. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 58(2), 116–128. <https://doi.org/10.1111/jcpp.12611>
- Mauz, E., Gößwald, A., Kamtsiuris, P., Hoffmann, R., Lange, M., Schenck, U. von, Allen, J., Butschalowsky, H., Frank, L., Hölling, H., Houben, R., Krause, L., Kuhnert, R., Lange, C., Müters, S., Neuhauser, H., Poethko-Müller, C., Richter, A., Rosario, A.S., ... Kurth, B.-M. (2017). New data for action: Data collection for KiGGS Wave 2 has been completed. In *Journal of Health Monitoring* (Vol. 2). Robert Koch-Institut, Epidemiologie und Gesundheitsberichterstattung. <https://doi.org/10.17886/RKI-GBE-2017-105>
- Murray, A.L., Speyer, L.G., Hall, H.A., Valdebenito, S., & Hughes, C. (2022). A longitudinal and gender invariance analysis of the Strengths and Difficulties Questionnaire across ages 3, 5, 7, 11, 14, and 17 in a large U.K.-representative sample. *Assessment*, 29(6), 1248–1261. <https://doi.org/10.1177/10731911211009312>
- Patalay, P., Deighton, J., Fonagy, P., Vostanis, P., & Wolpert, M. (2014). Clinical validity of the Me and My School questionnaire: A self-report mental health measure for children and adolescents. *Child and Adolescent Psychiatry and Mental Health*, 8(1), 17. <https://doi.org/10.1186/1753-2000-8-17>
- Patalay, P., Hayes, D., & Wolpert, M. (2018). Assessing the readability of the self-reported Strengths and Difficulties Questionnaire. *BJPsych Open*, 4(2), 55–57. <https://doi.org/10.1192/bjo.2017.13>
- Potrebny, T., Wiium, N., Haugstvedt, A., Sollesnes, R., Torsheim, T., Wold, B., & Thuen, F. (2019). Health complaints among adolescents in Norway: A twenty-year perspective on trends. *PLoS One*, 14(1), e0210509. <https://doi.org/10.1371/journal.pone.0210509>
- R Core Team. (2019). *A language and environment for statistical computing*. [Computer software]. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Rogge, J., Speck, K., Hölling, H., Minnaert, A., Koglin, U., & Schlack, R. (2018). Messinvarianz zwischen Eltern- und Jugendversion des Strengths and Difficulties Questionnaire (SDQ)? [Measurement Invariance Across Youth and Parent SDQ Ratings?]. *Diagnostica*, 64(3), 121–132. <https://doi.org/10.1026/0012-1924/a000200>
- RStudio Team. (2018). *RStudio: Integrated development for R* [Computer software]. Author. Retrieved from <http://www.rstudio.com/>
- van Roy, B., Groholt, B., Heyerdahl, S., & Clench-Aas, J. (2010). Understanding discrepancies in parent-child reporting of emotional and behavioural problems: Effects of relational and socio-demographic factors. *BMC Psychiatry*, 10(1), 56. <https://doi.org/10.1186/1471-244X-10-56>
- van Vuuren, C.L., Uitenbroek, D.G., van der Wal, Marcel, F., & Chinapaw, M.J.M. (2018). Sociodemographic differences in 10-year time trends of emotional and behavioural problems among adolescents attending secondary schools in Amsterdam, The Netherlands. *European Child & Adolescent Psychiatry*, 27(12), 1621–1631. <https://doi.org/10.1007/s00787-018-1157-5>
- van Widenfelt, B.M., Goedhart, A.W., Treffers, P.D.A., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child & Adolescent Psychiatry*, 12(6), 281–289. <https://doi.org/10.1007/s00787-003-0341-3>
- Vaz, S., Cordier, R., Boyes, M., Parsons, R., Joosten, A., Ciccirelli, M., Falkmer, M., & Falkmer, T. (2016). Is using the Strengths and Difficulties Questionnaire in a community sample the optimal way to assess mental health functioning? *PLoS One*, 11(1), e0144039. <https://doi.org/10.1371/journal.pone.0144039>
- Vugtveen, J., Bildt, A. de, Hartman, C.A., & Timmerman, M.E. (2018). Using the Dutch multi-informant Strengths and Difficulties Questionnaire (SDQ) to predict adolescent psychiatric diag-

noses. *European Child & Adolescent Psychiatry*, 27(10), 1347–1359. <https://doi.org/10.1007/s00787-018-1127-y>  
Yoon, Y., Eisenstadt, M., Lereya, S.T., & Deighton, J. (2022). Gender difference in the change of adolescents' mental health and subjective well-being trajectories. *European Child & Adolescent Psychiatry*. <https://doi.org/10.1007/s00787-022-01961-4>

### History

Manuscript submitted: 26.08.2022

Accepted after revision: 19.04.2023

Published online: 07.07.2023

### Conflicts of interests

No conflicts of interest exist.

### Funding

Open access publication enabled by Robert Koch Institute.

### ORCID

Stephan Junker

 <https://orcid.org/0000-0001-8612-9347>

### Dr. Robert Schlack

Robert Koch-Institut  
General-Pape-Straße 62–66  
12101 Berlin  
Germany

[schlackr@rki.de](mailto:schlackr@rki.de)