

Problems in school reported by children in regular and hospital schools

Judith Janschewski (TU Dortmund) and Philipp Berens (Uni Tübingen)

1 Dataset

The dataset contains questionnaire responses of 573 students from regular and hospital schools. The following tables provides an overview:

gender	N
female	288
male	217
NA	68

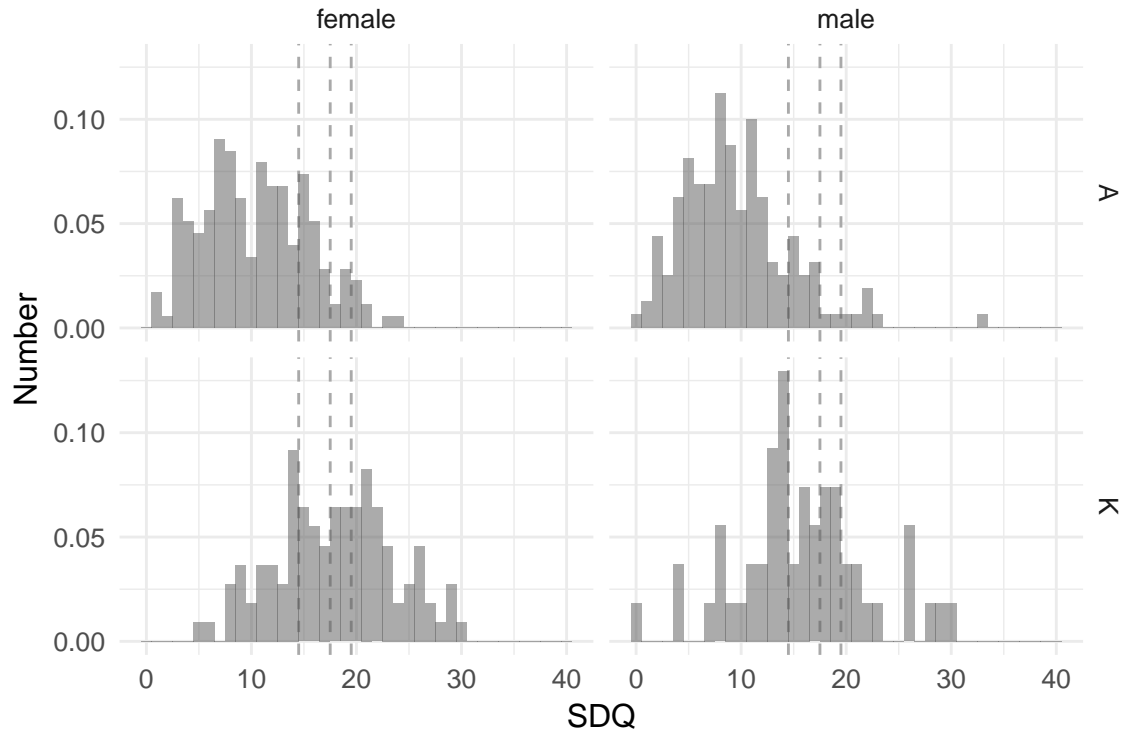
school_type	N
A	388
K	185

gender	school_type	N
female	A	179
female	K	109
male	A	162
male	K	55
NA	A	47
NA	K	21

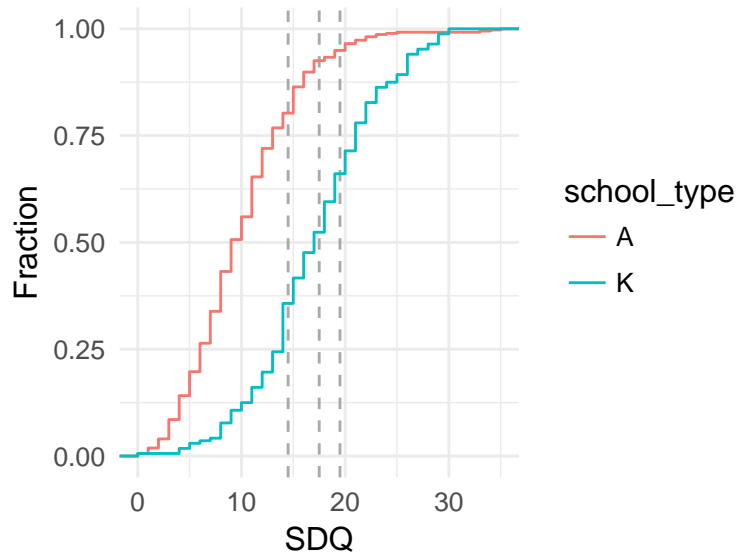
age	N
11-12	60
13-14	205
15-16	167
17-18	110
NA	31

2 Validating the SDQ results

We explore the distribution of SDQ scores for regular (A) and hospital (K) schools separately. The thin lines indicate the four scale levels for a *slightly raised* score (15-17), a *high* score (18-19) and a *very high* score (>20).



The cumulative SDQ distributions of general (A) and hospital (K) schools show a clear shift between the two student populations:



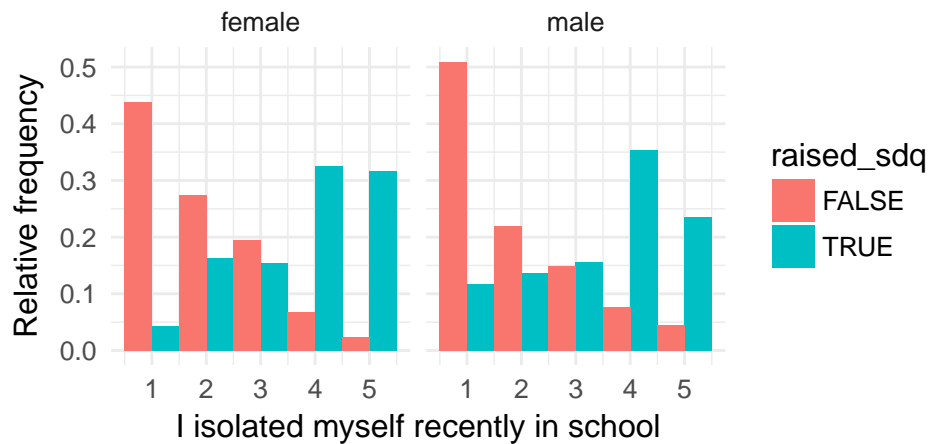
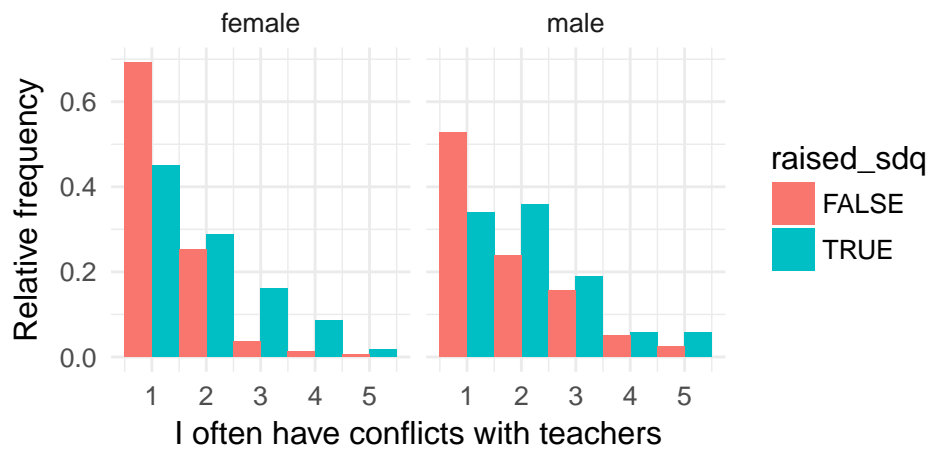
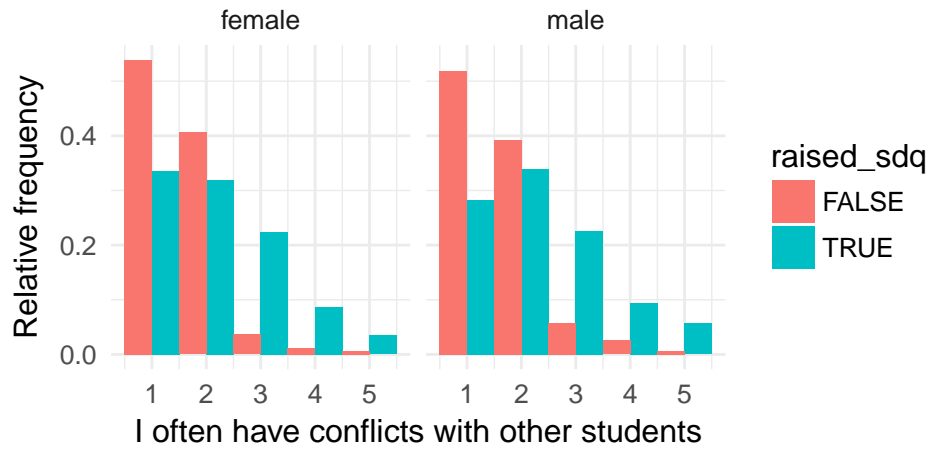
The percentage of students with raised SDQ in general schools corresponds approximately to that expected:

raised_sdq	N	fraction_raised	fraction_vhigh
FALSE	361	0	0.0000000
TRUE	182	1	0.4175824
NA	30	NaN	NaN

3 Overview of item responses

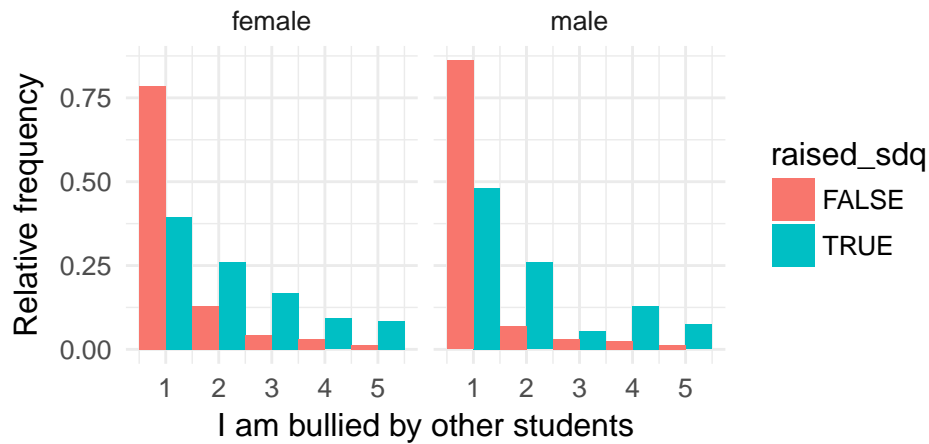
We first provide an overview of several of the response items.

3.1 Social problems





raised_sdq	gender	N	conflict_stud	conflict_teach	isolation	large_groups
FALSE	female	167	1.539394	1.385542	1.963415	1.809816
FALSE	male	160	1.607595	1.805031	1.929032	1.566879
TRUE	female	119	2.163793	1.932203	3.709402	3.294118
TRUE	male	54	2.301887	2.132076	3.450980	2.888889



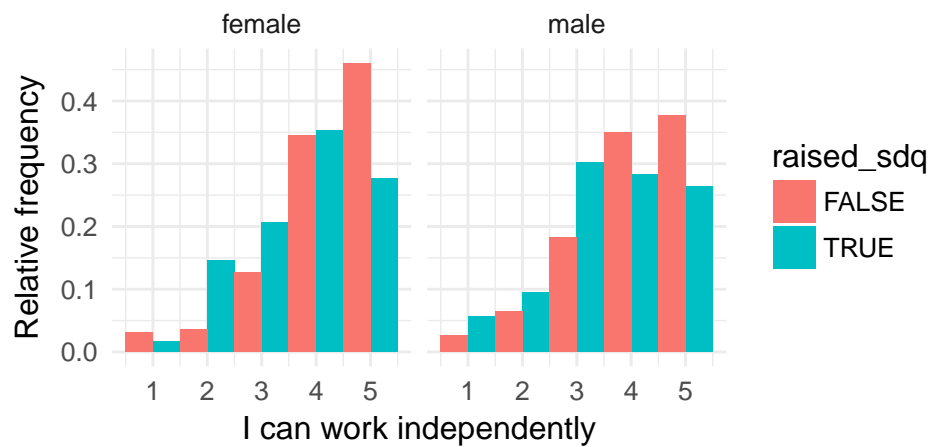
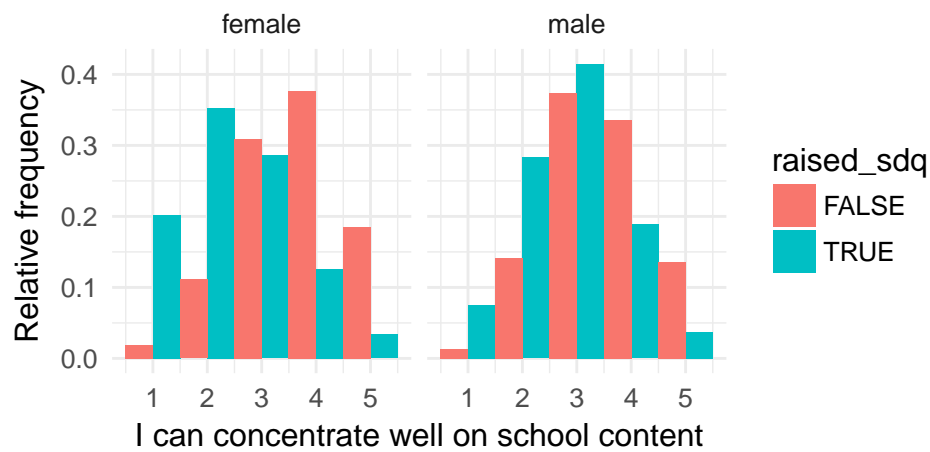


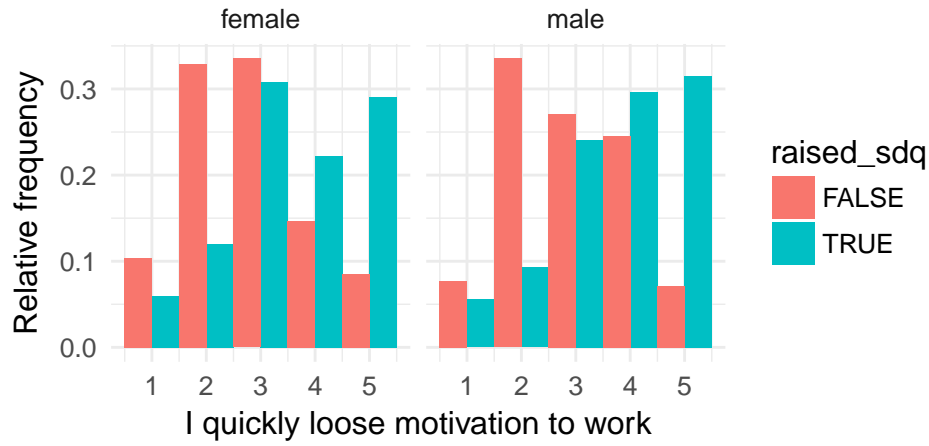
raised_sdq	gender	N	friends	bullying	outsider	agressive
FALSE	female	167	4.722892	1.355828	1.503068	1.409639
FALSE	male	160	4.779874	1.257862	1.484076	1.592357
TRUE	female	119	3.788136	2.210084	2.410256	2.050420
TRUE	male	54	3.846154	2.055556	2.240741	2.471698

3.2 School performance



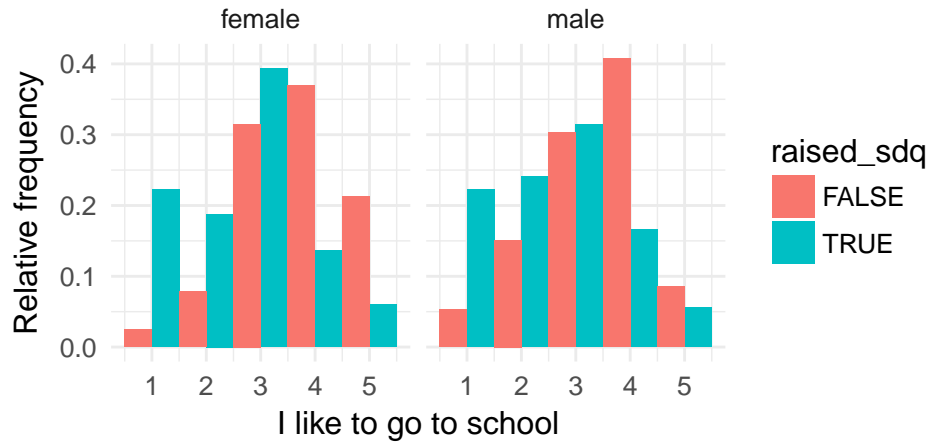
The next item shows a pretty centered, almost Normal distribution. This is interesting - I would imagine this may indicate that the students did not understand the question properly.

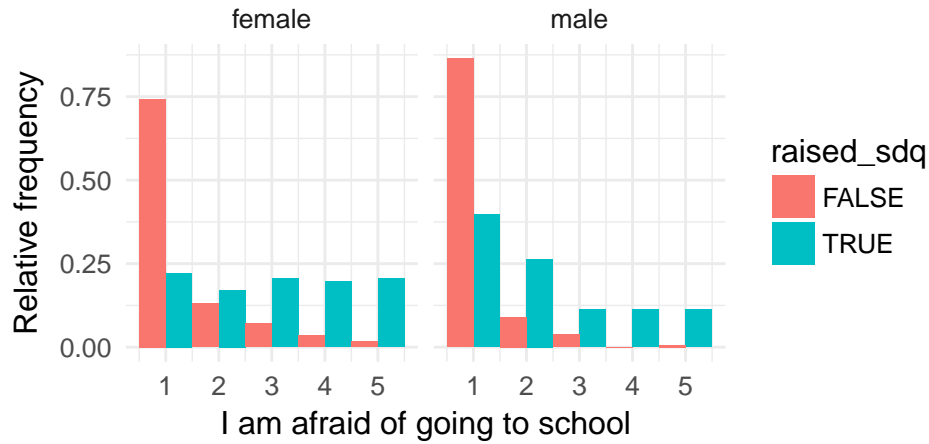




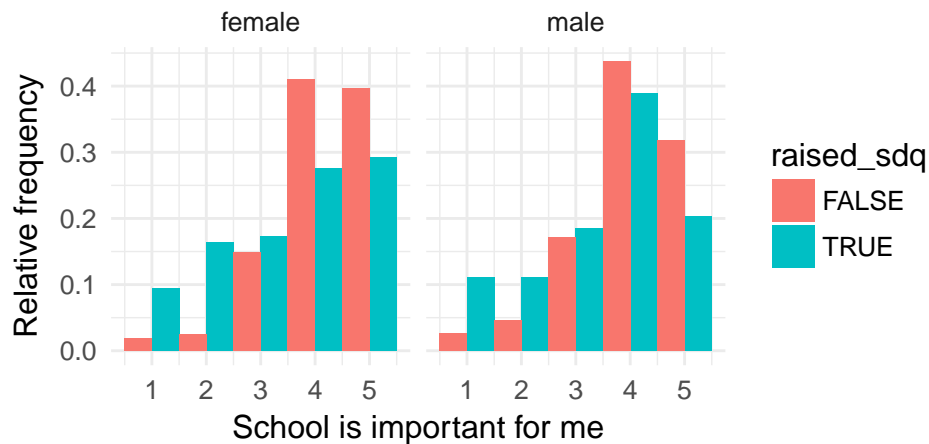
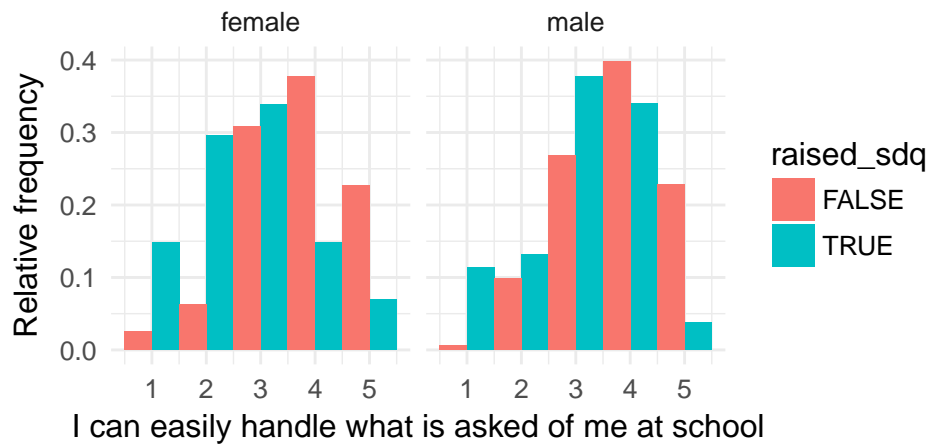
raised_sdq	gender	N	perf_down	concentrate	independent	loose_mot
FALSE	female	167	2.481250	3.598765	4.169697	2.780488
FALSE	male	160	2.592105	3.438710	3.987013	2.896774
TRUE	female	119	3.469565	2.436975	3.724138	3.564103
TRUE	male	54	3.333333	2.830189	3.603774	3.722222

3.3 Wellbeing/attitude





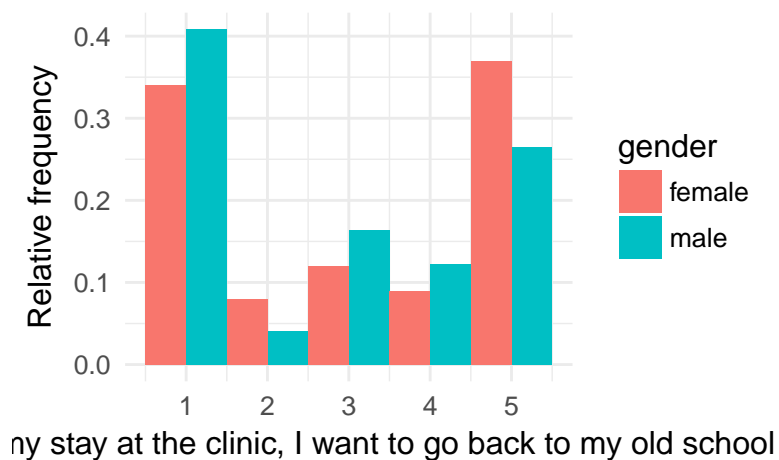
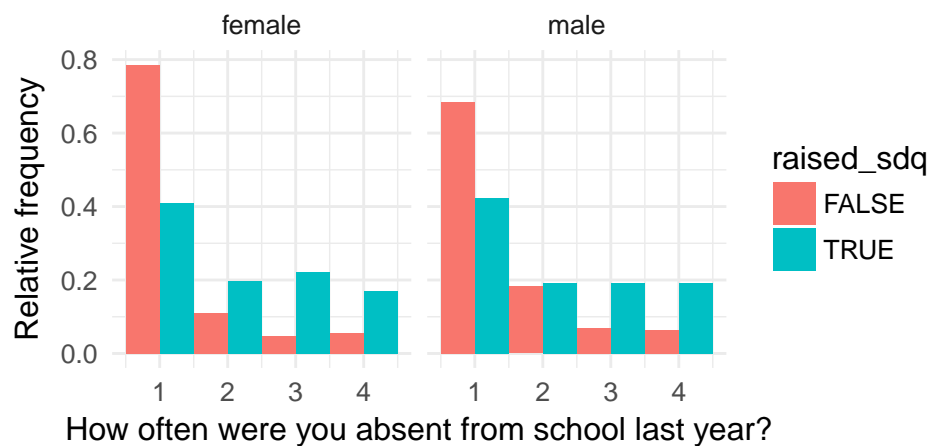
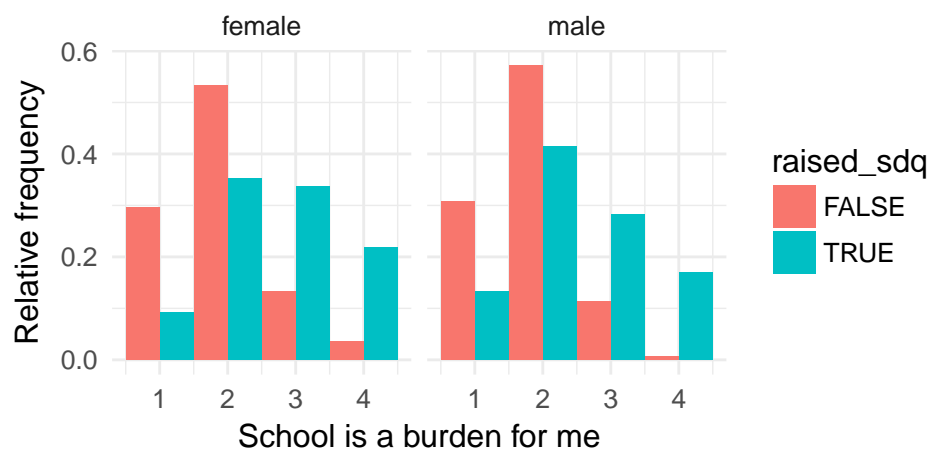
The next item again may not have been properly understood:



raised_sdq	gender	N	like_school	afraid	handle	important
FALSE	female	167	3.666667	1.455090	3.716981	4.142857
FALSE	male	160	3.322368	1.193548	3.745098	3.973510
TRUE	female	119	2.623932	2.991453	2.695652	3.508621
TRUE	male	54	2.592593	2.283019	3.056604	3.462963

raised_sdq	gender	N	like_school	afraid	handle	important
------------	--------	---	-------------	--------	--------	-----------

3.4 Other items



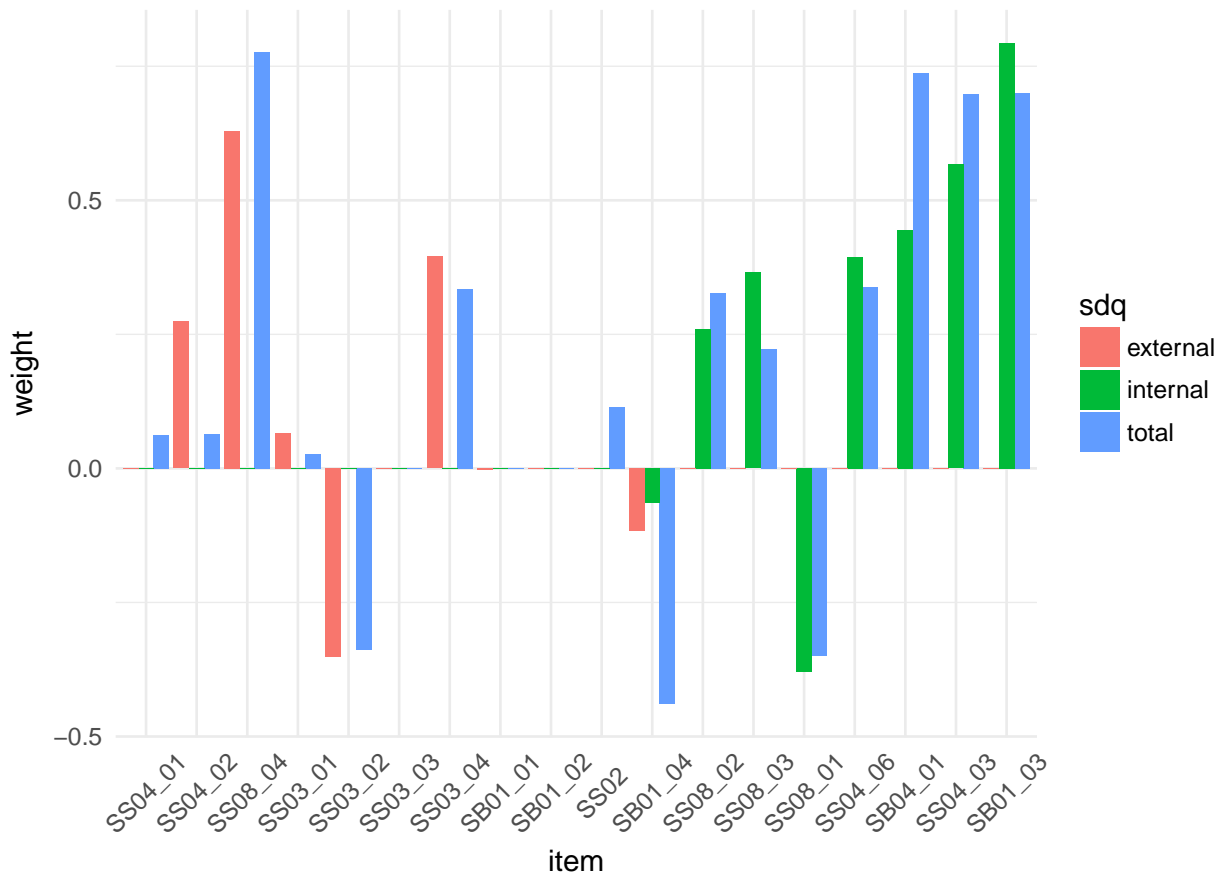
raised_sdq	gender	N	burden	absence	goback
FALSE	female	167	1.539394	1.371951	3.838710
FALSE	male	160	1.607595	1.509434	3.363636
TRUE	female	119	2.163793	2.153846	2.724638
TRUE	male	54	2.301887	2.153846	2.333333

4 Classification analysis

We have quite a few missing values in the predictor variables. For now, we impute them using the mice package link.

We try to predict total, internal and external SDQ score from the response items using the glmnet package link.

We can plot the weights to visualize influential items:



This indicates that the items predictive for a high score of the SDQ internalizing scale are disjunct from those predictive for a high score on the externalizing SDQ scale.

Predictive of a high score on the internalizing SDQ scale are (in order of decreasing weight):

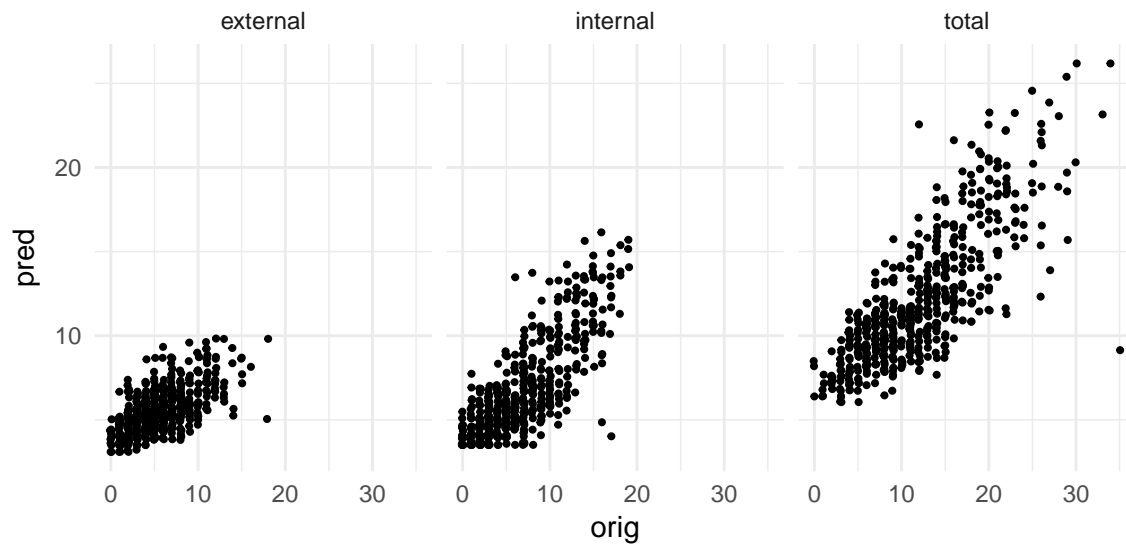
- I am afraid to go to school
- I isolated myself recently
- School is a burden for me
- I don't like large groups like my class

- I have friends in my class
- I am an outsider
- I am bullied by other students

Predictive of a high score an the externalizing SDQ scale are:

- I have conflicts with teachers
- I can work independently
- My school performance dropped during the last year
- I have conflicts with other students

Interesting is “I can easily handle what is asked of me at school” - it is not strongly predictive for either internalizing/externalizing, but highly predictive of the total score. Most other items that are predictive for one are also predictive for the total.

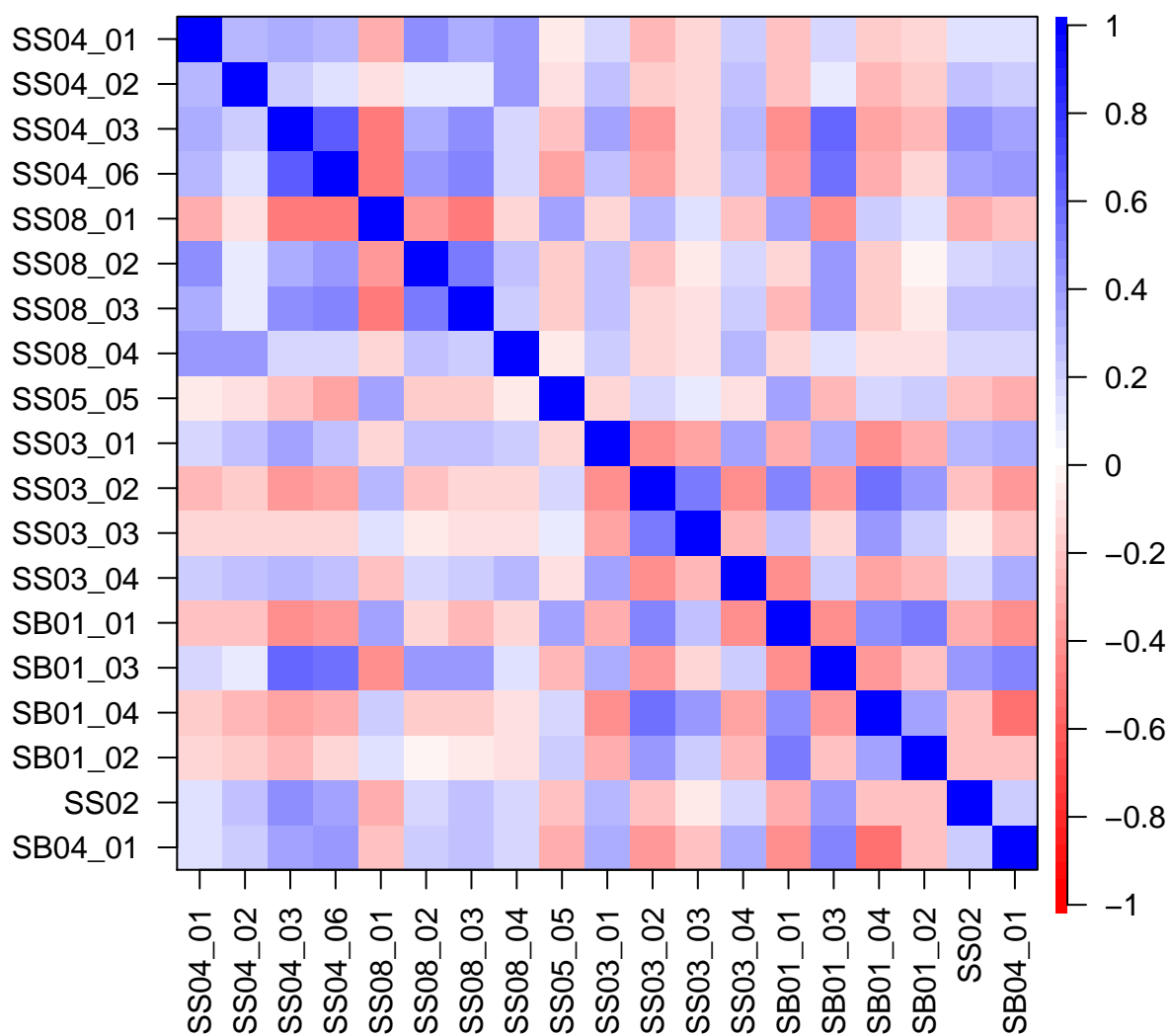


The correlation between the original values and the predictions is high (total: 0.7954679, internal: 0.8001, external: 0.6314871), indicating good prediction.

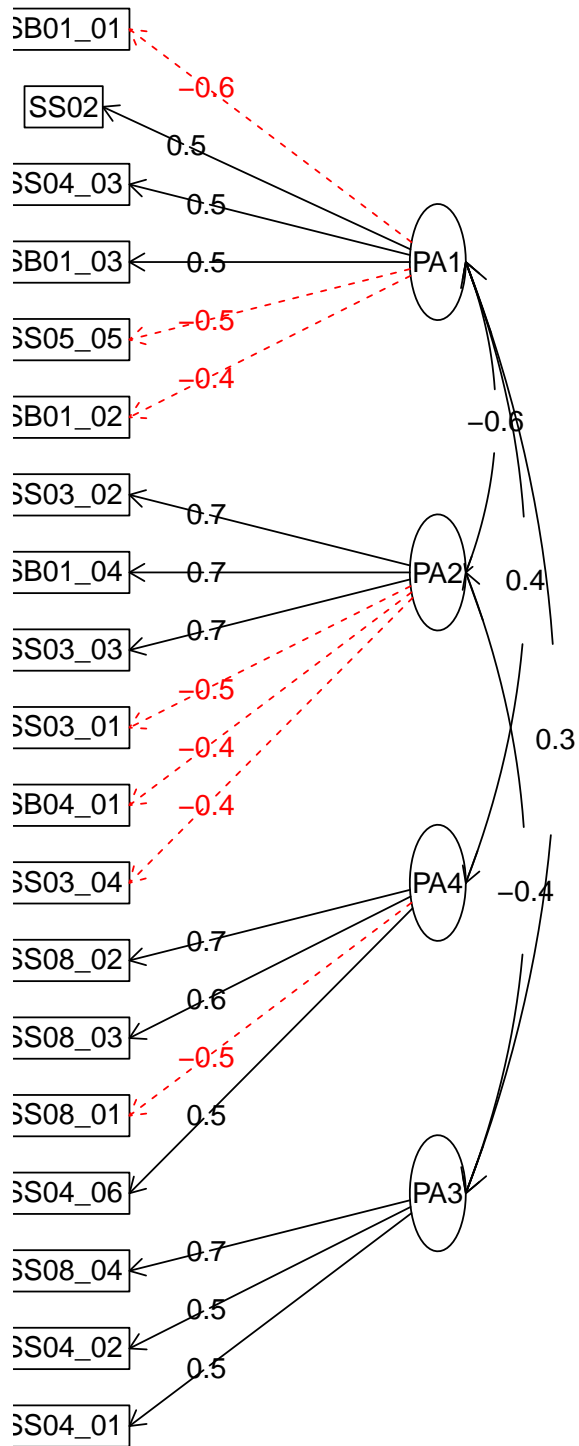
5 Exploratory factor analysis

We use exploratory factor analysis to identify latent factors in the data. We use the psych package. The correlation matrix of all considered items looks like this:

Correlation plot



Factor Analysis



A provisional interpretation of the factors could be:

- PA1: Affinity to school and a well-being at school
- PA2: Ability to cope with school demands
- PA3: Social integration
- PA4: Aggressive Behavior

5.1 Reliability analysis