Problems in school reported by children in regular and hospital schools

Judith Janschewski (TU Dortmund) and Philipp Berens (Uni T?bingen)

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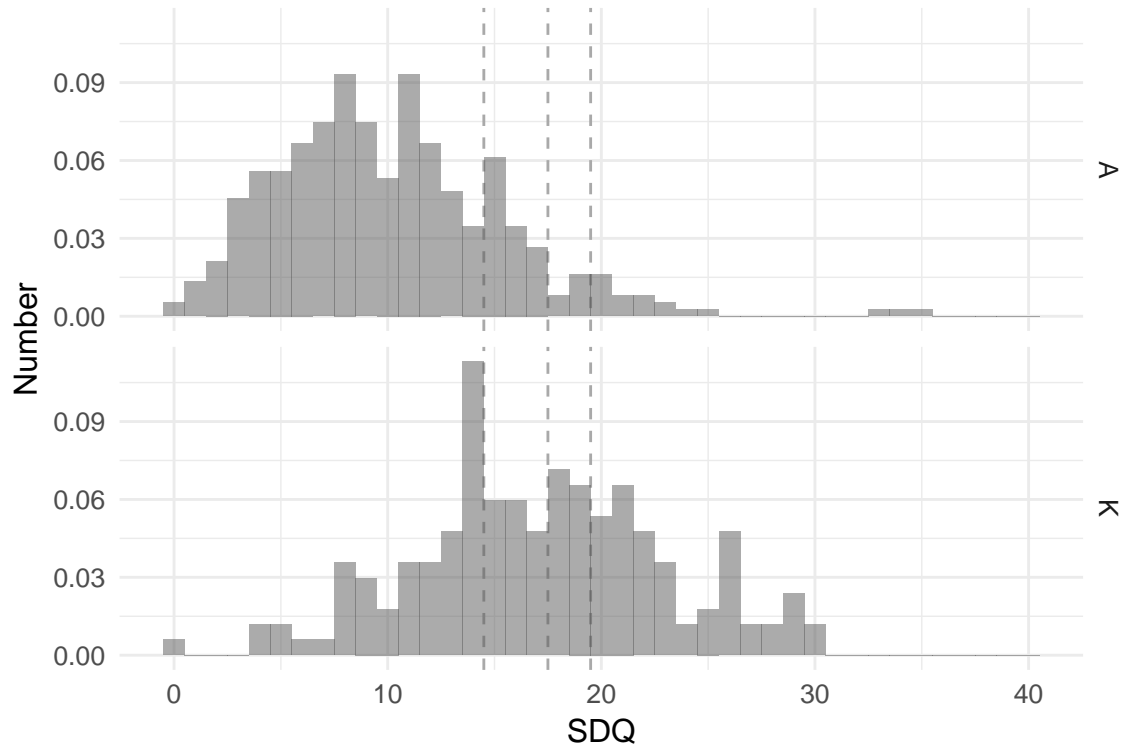
The cumulative SDQ distributions of general (A) and hosptial (K) schools show a clear shift between the two student populations:
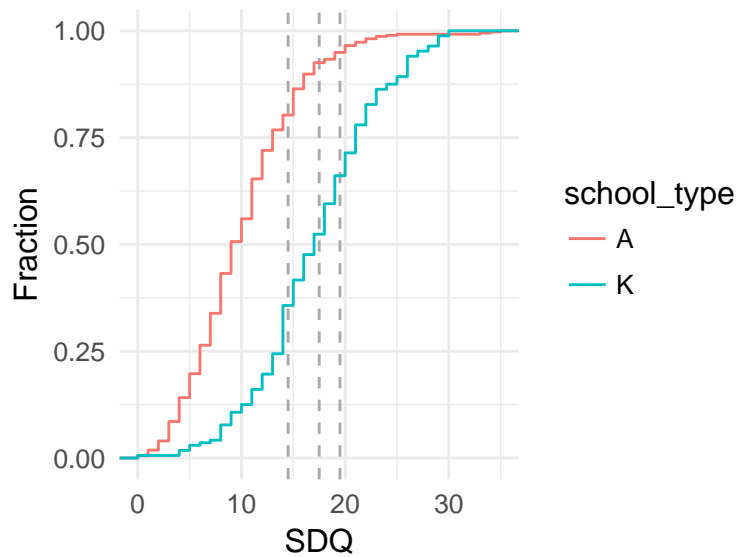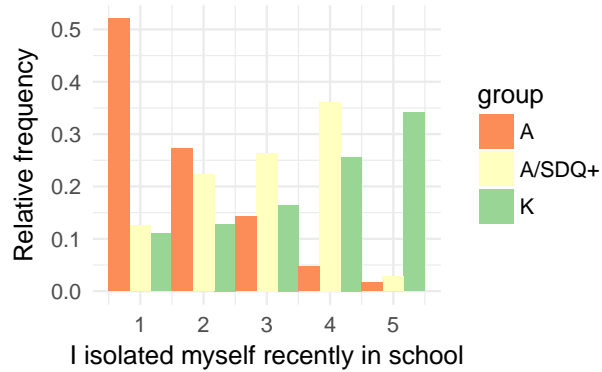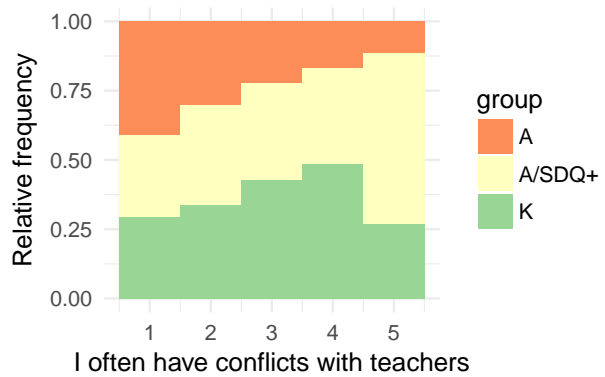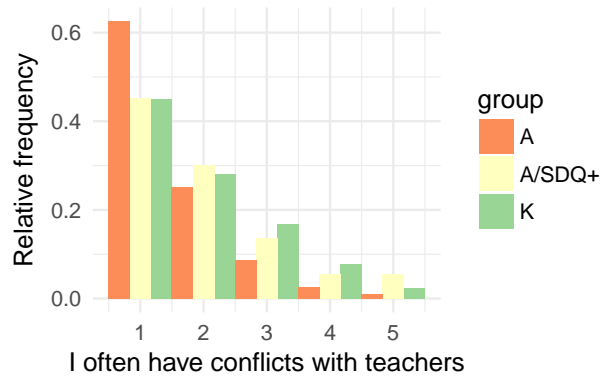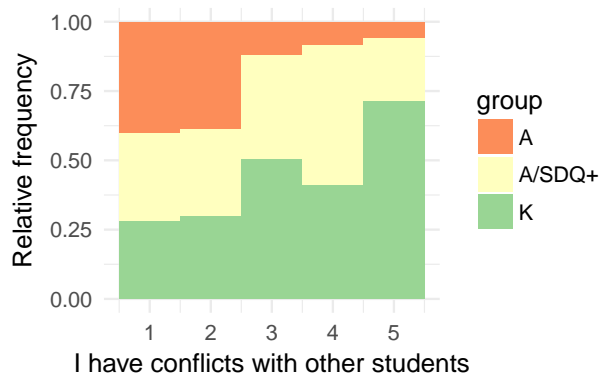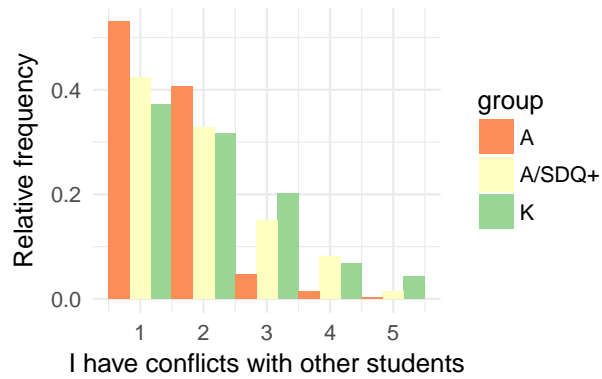

The percentage of students with raised SDQ in general schools corresponds approximately to that expected:

| school_type	N	fraction_raised	fraction_vhigh
A	388	0.1973333	0.0506667
K	185	0.6428571	0.3392857

3 Overview of item responses

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

3.1 Social problems

I have conflicts with other students

I often have conflicts with teachers

I isolated myself recently in school
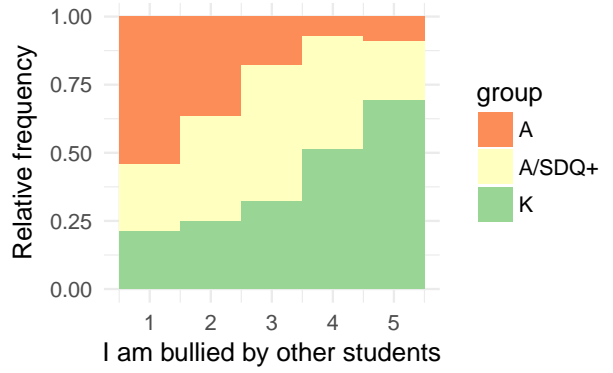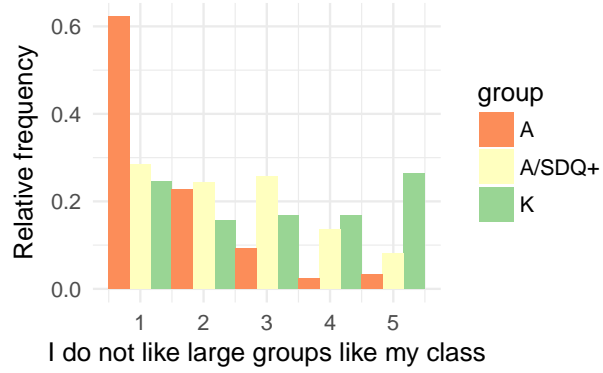
group	N	conflict_stud	conflict_teach	isolation	large_groups
A	301	1.553691	1.545151	1.768708	1.619048
A/SDQ+	74	1.931507	1.958904	2.944444	2.486486
K	168	2.091463	1.946108	3.591463	3.047904
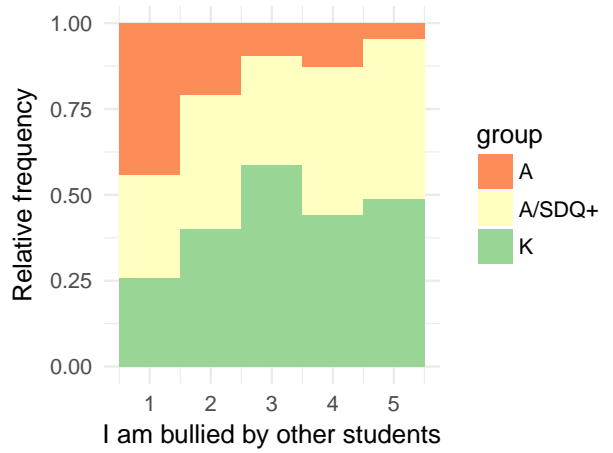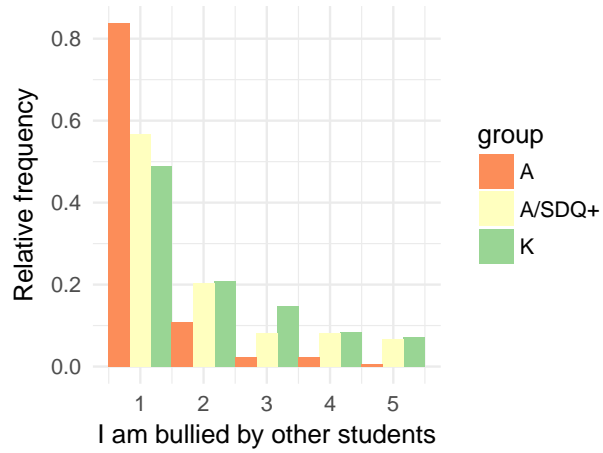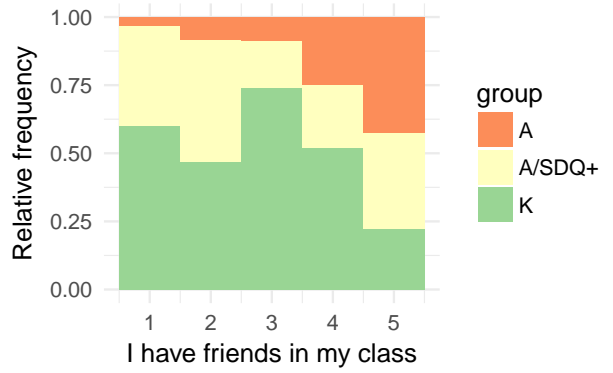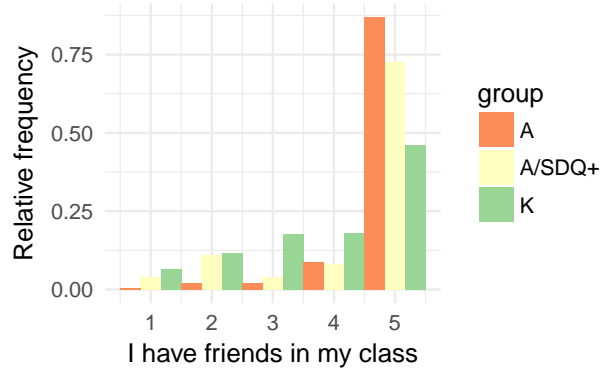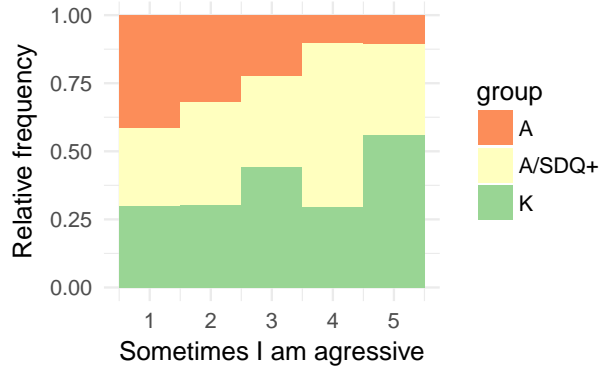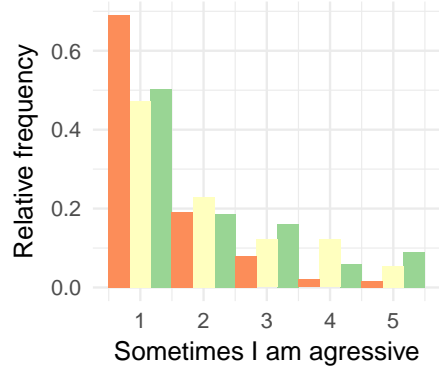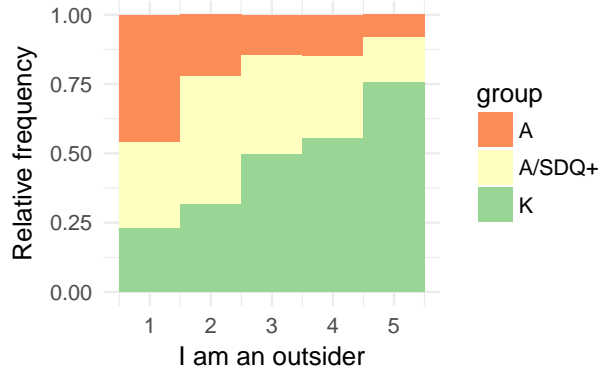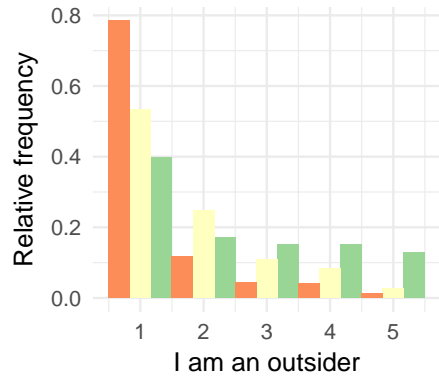
group	N	friends	bullying	outsider	agressive
A	301	4.800000	1.253378	1.380471	1.481481
A/SDQ+	74	4.342466	1.878378	1.821918	2.054054
K	168	3.854545	2.041667	2.445122	2.047904

3.2 School performance
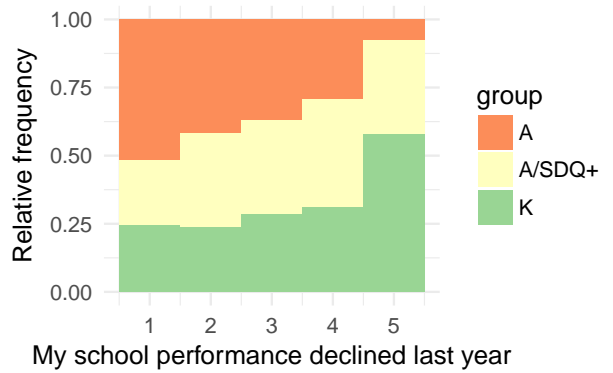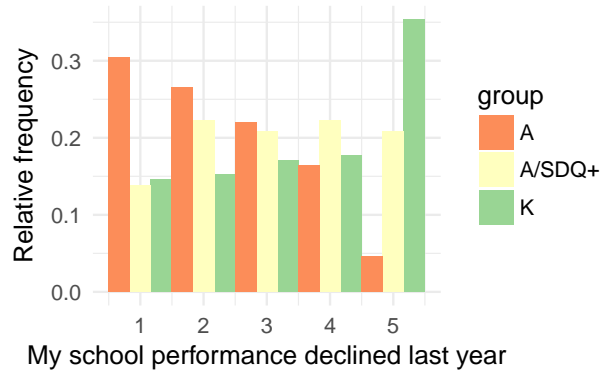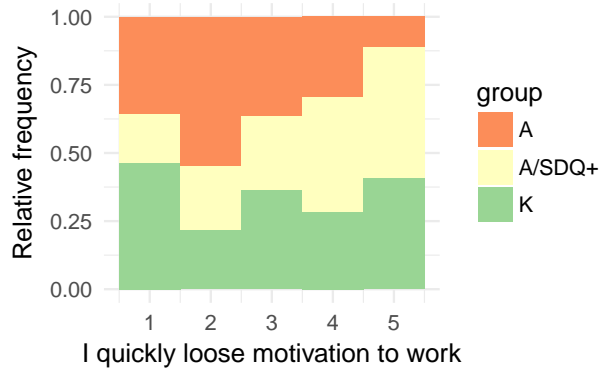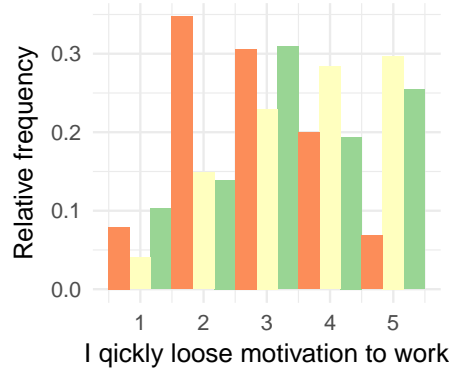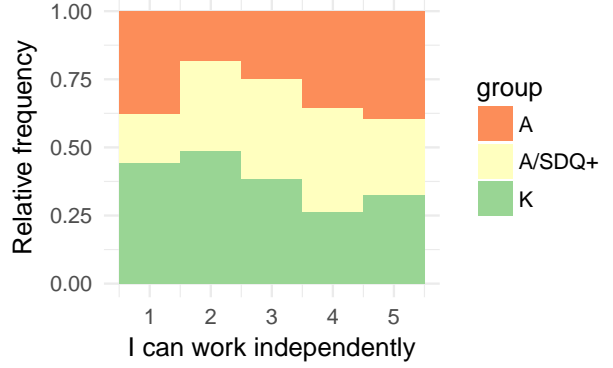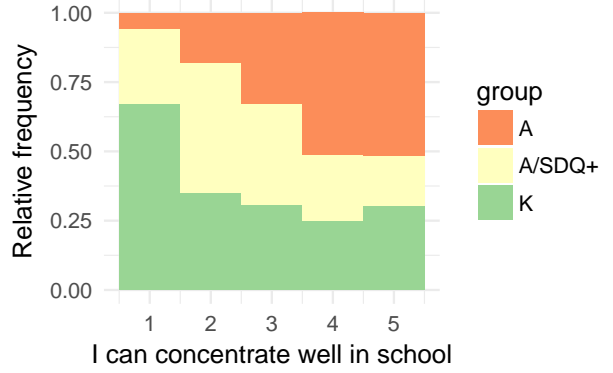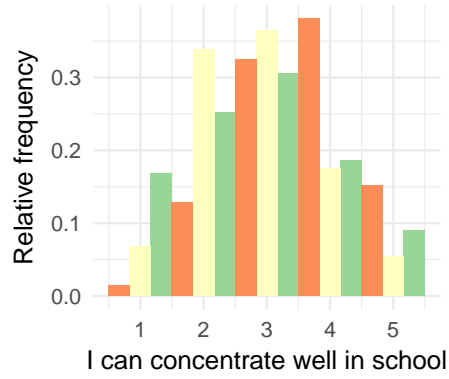
group	N	perf_down	concentrate	independent	loose_mot
A	301	2.381119	3.529412	4.078231	2.831615
A/SDQ+	74	3.138889	2.810811	3.857143	3.648649
K	168	3.439024	2.778443	3.753012	3.357576

3.3 Wellbeing/attitude
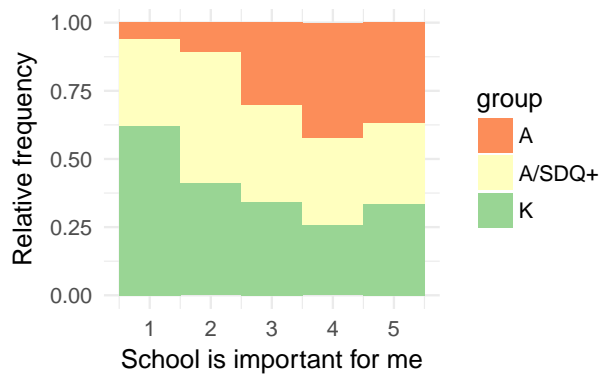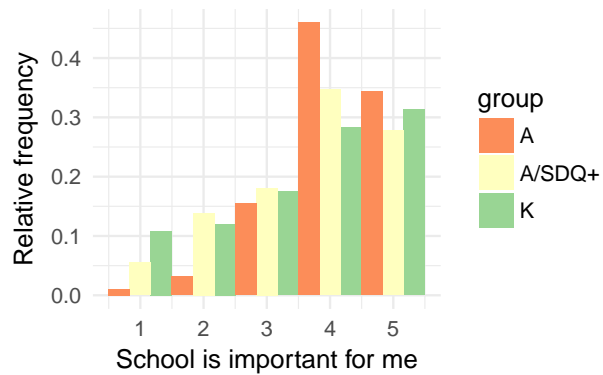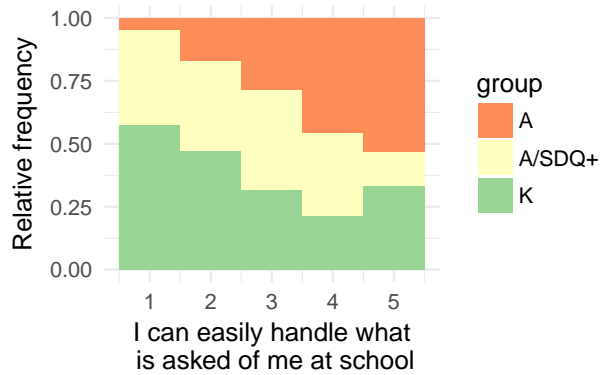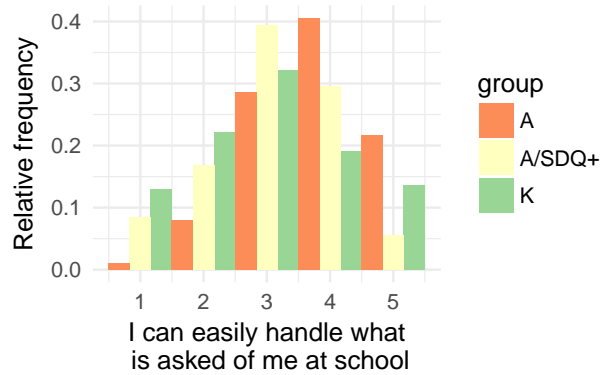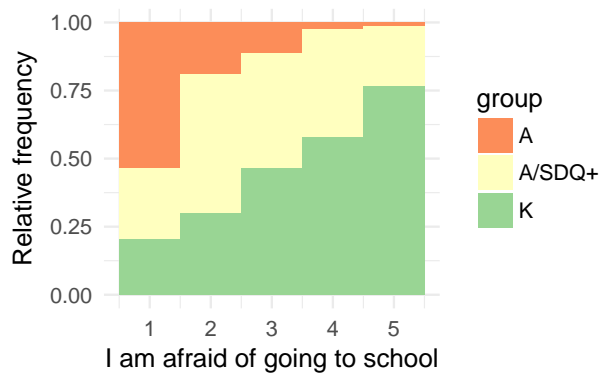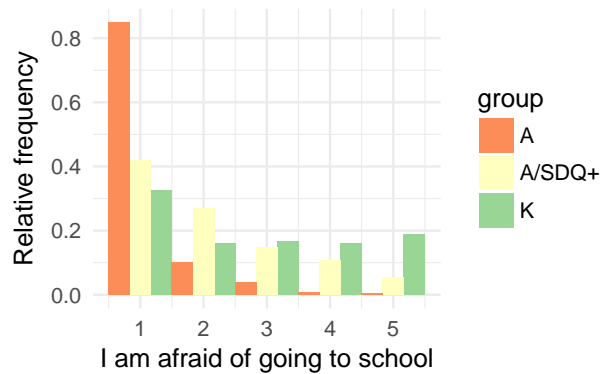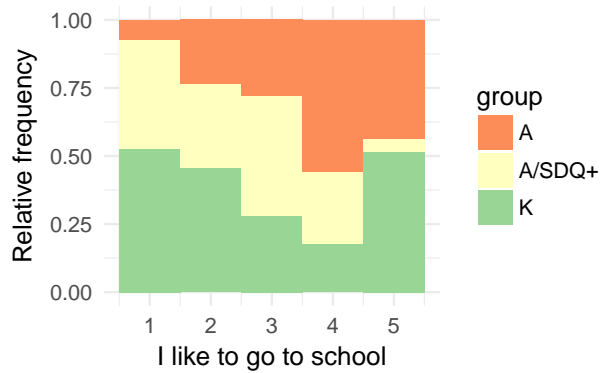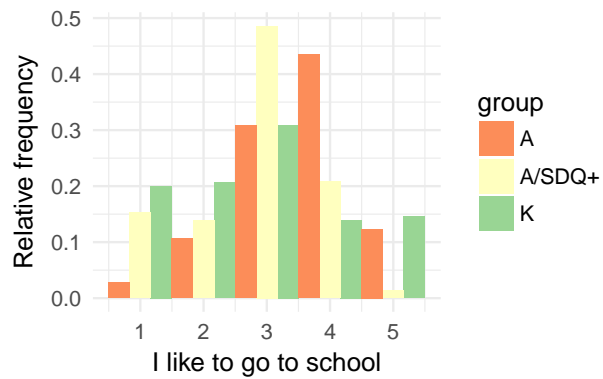
group	N	like_school	afraid	handle	important
A	301	3.520548	1.214765	3.737762	4.094737
A/SDQ+	74	2.791667	2.108108	3.070422	3.652778
K	168	2.824242	2.730061	2.981482	3.572289

3.4 Other items
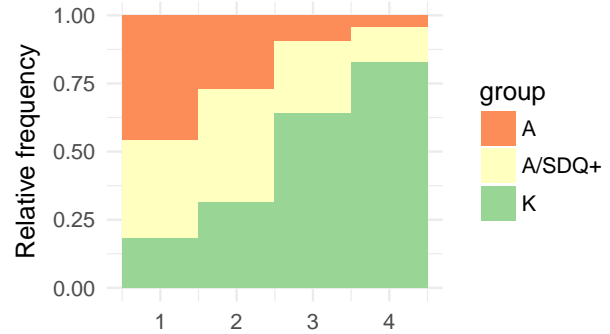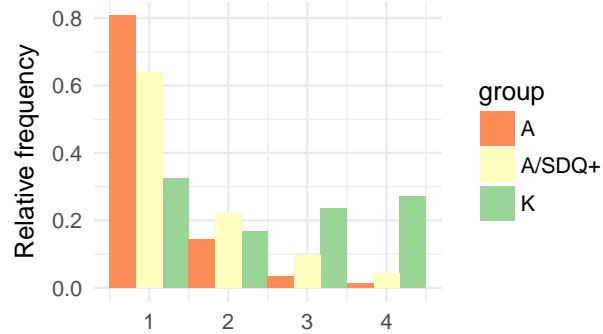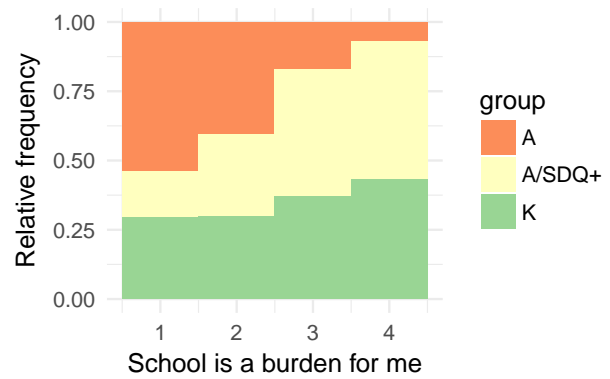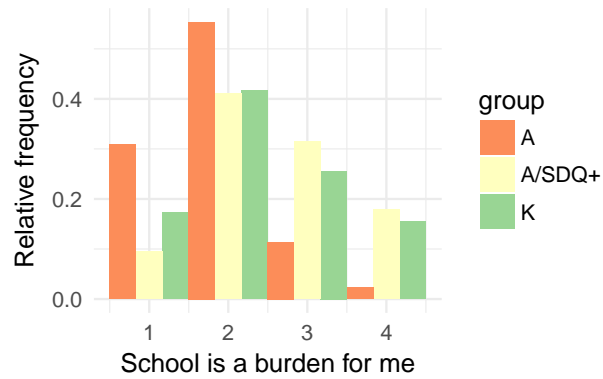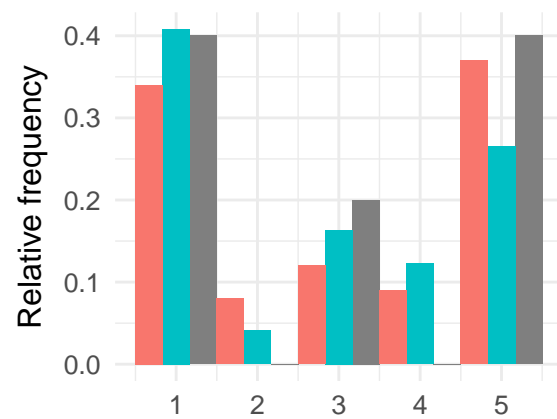
group	N	burden	absence	goback
A	301	1.553691	1.252525	NaN
A/SDQ+	74	1.931507	1.541667	NaN

group	N	burden	absence	goback
K	168	2.091463	2.451807	2.98052

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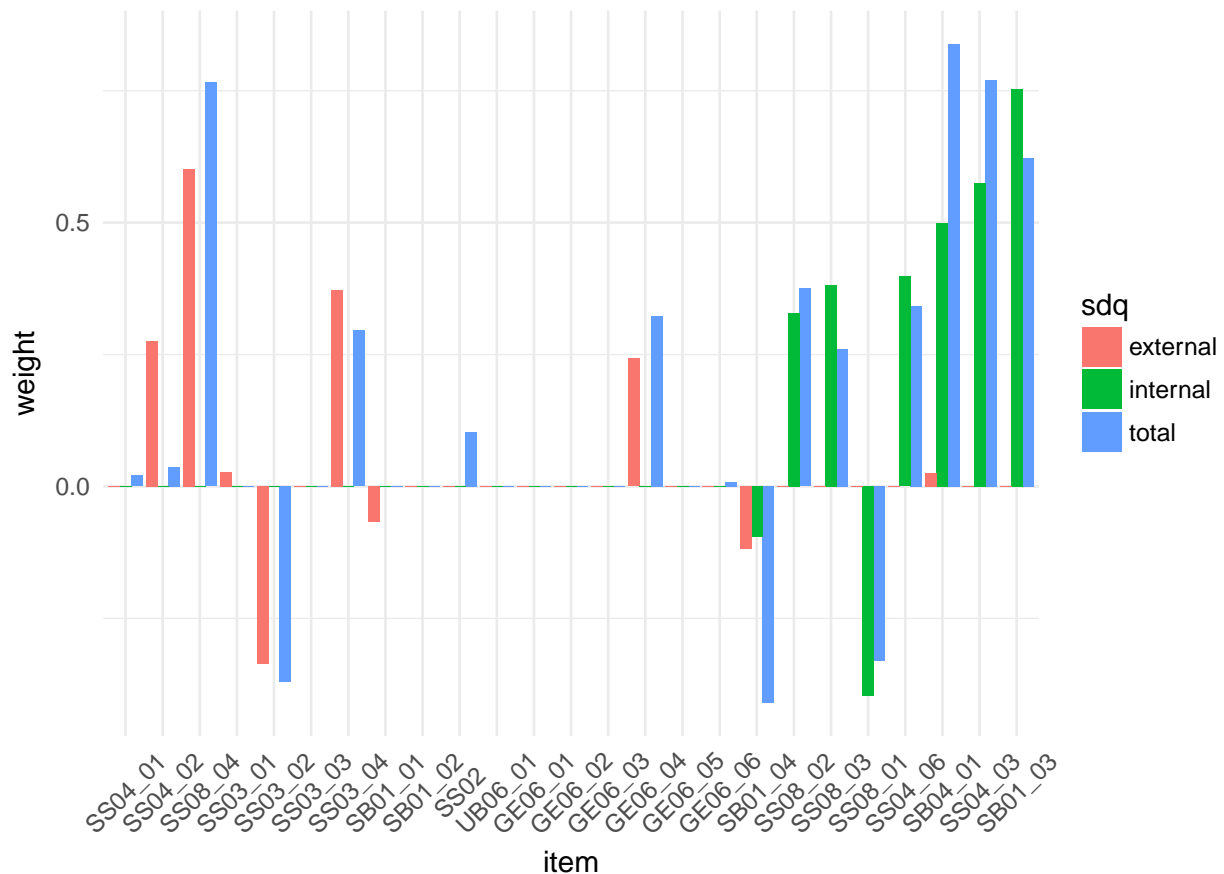
This indicates that the items predictive for a high score of the SDQ internalizing scale are disjunct from those predictive for a high score on the externalizing SDQ scale.

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 preditive for one are also predictive for the total.
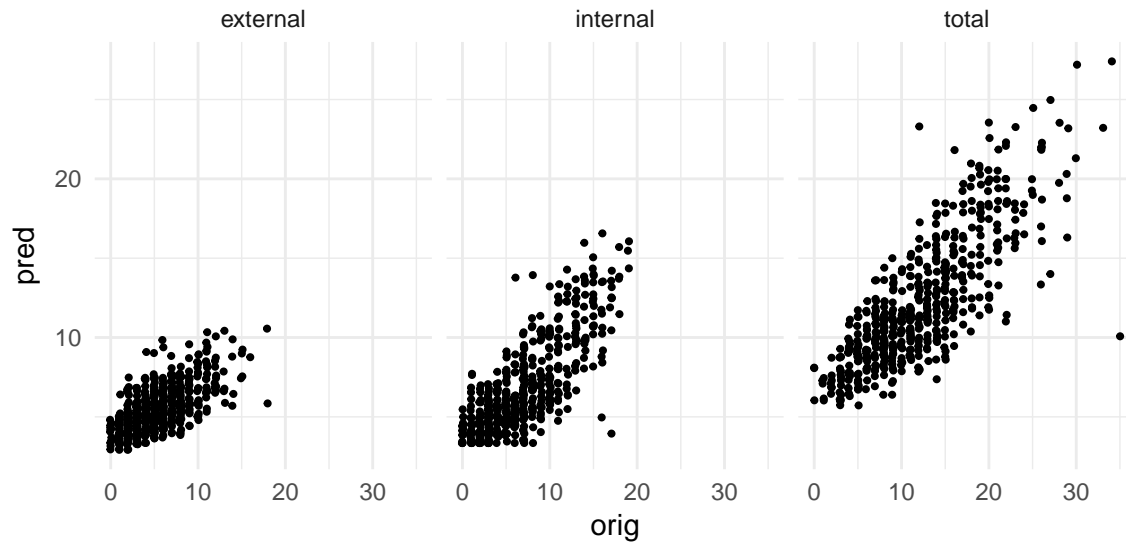
The correlation between the original values and the predictions is high (total: 0.8045056, internal: 0.8010078, external: 0.6467093), indicating good prediction.

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:

Factor Analysis
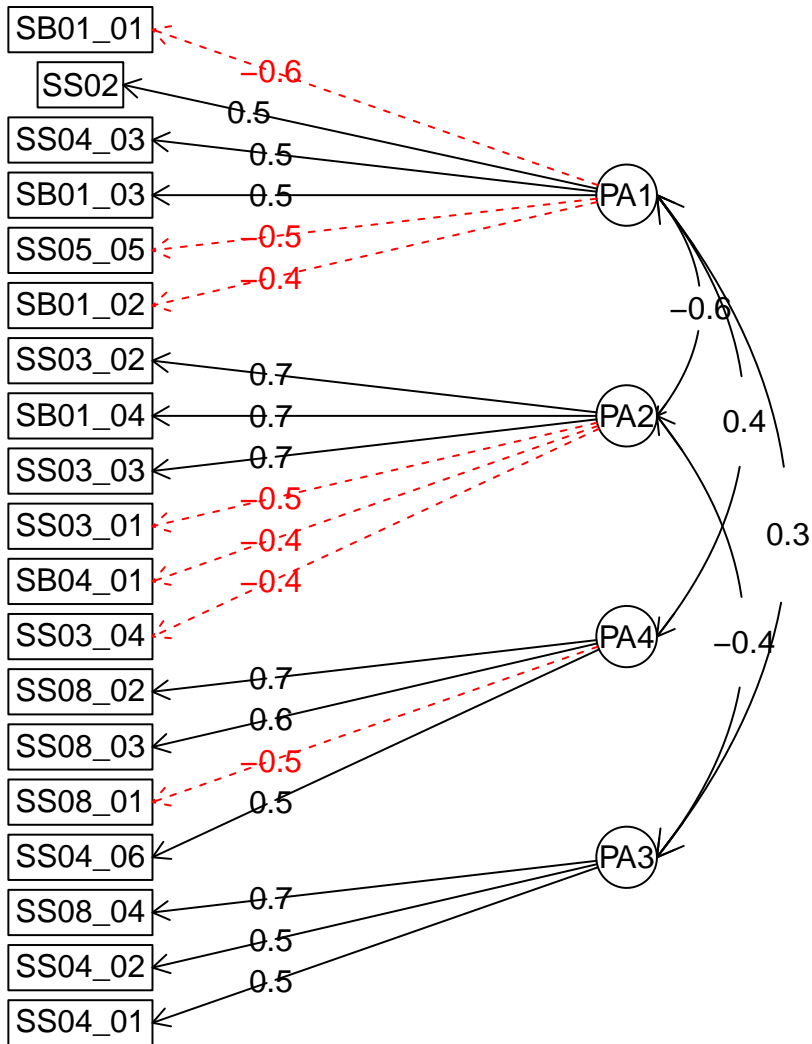
A provisional interpretation of the factors could be:

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

6 Support through teachers

group	N	nans	fraction
A	301	198	0.6578073

group	N	nans	fraction
A/SDQ+	74	23	0.3108108
K	168	44	0.2619048

The spearman correlation is -0.3630675 for total SDQ, -0.3363001 for internalizing SDQ and -0.2026263 for externalizing SDQ.
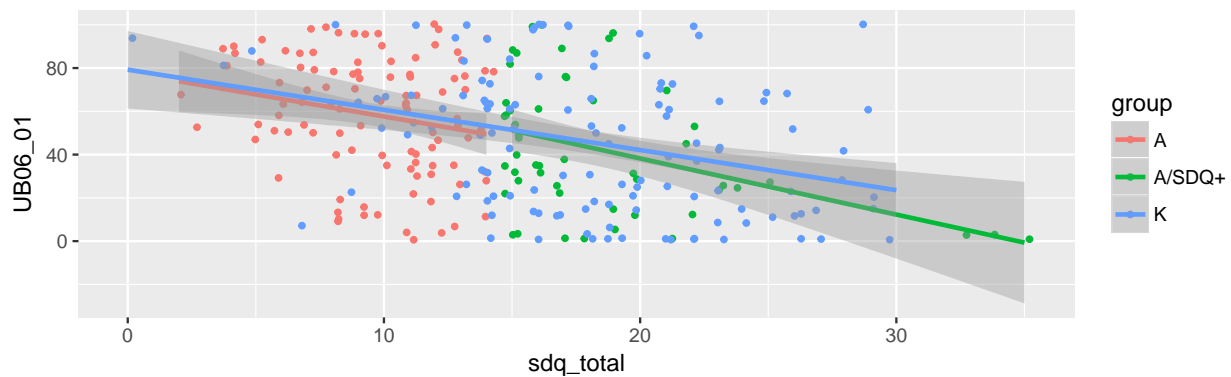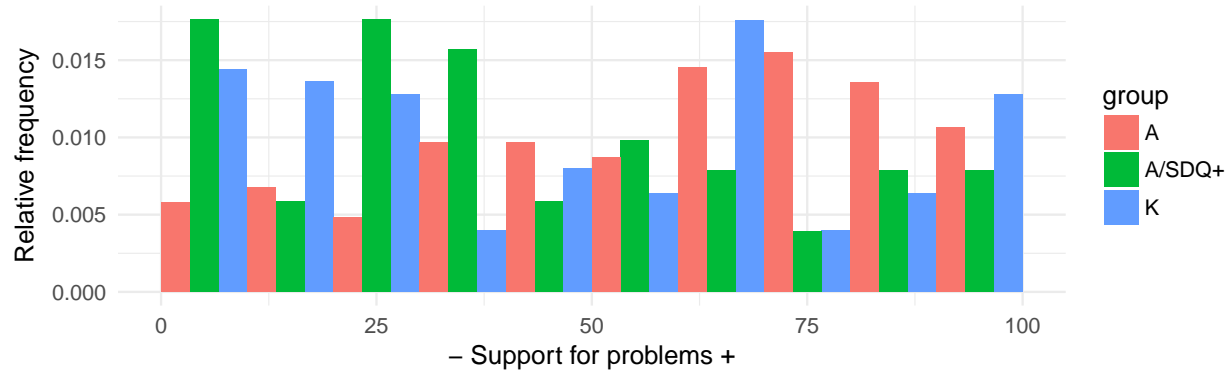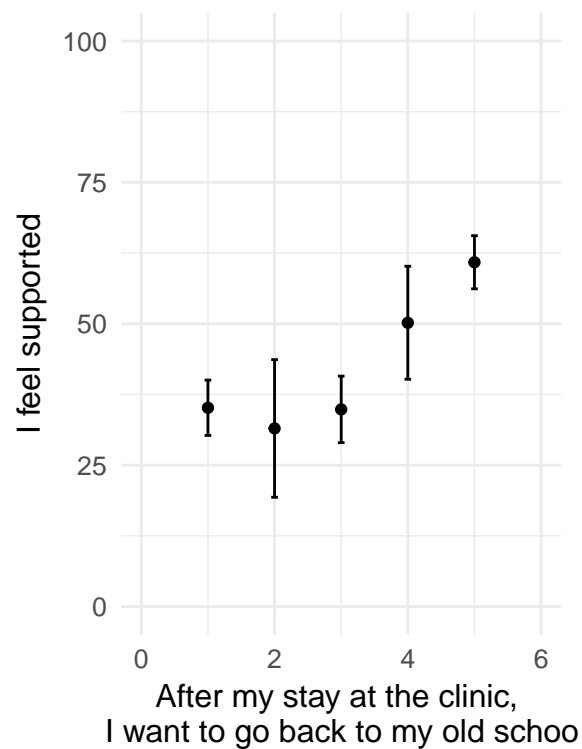
```
##
## Call:
## lm(formula = UB06_01 ~ sdq_total + school_type, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -60.451 -20.886  -0.053  20.278  75.791
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    77.674      4.395  17.674  < 2e-16 ***
## sdq_total      -1.966      0.301  -6.530 3.16e-10 ***
## school_typeK    3.536      3.805   0.929    0.354
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 28.34 on 275 degrees of freedom
##   (295 observations deleted due to missingness)
## Multiple R-squared:  0.1467, Adjusted R-squared:  0.1405
## F-statistic: 23.64 on 2 and 275 DF,  p-value: 3.373e-10
```

```
##
## Call:
## lm(formula = UB06_01 ~ sdq_total + school_type, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -60.451 -20.886  -0.053  20.278  75.791
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     77.674      4.395  17.674  < 2e-16 ***
## sdq_total       -1.966      0.301  -6.530 3.16e-10 ***
## school_typeK     3.536      3.805   0.929    0.354
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 28.34 on 275 degrees of freedom
##   (295 observations deleted due to missingness)
## Multiple R-squared:  0.1467, Adjusted R-squared:  0.1405
## F-statistic: 23.64 on 2 and 275 DF,  p-value: 3.373e-10
```


I want to go back to school vs I feel supported by my teachers

Predict I feel supported from positive and negative properties