Erasmus+ project „The Children of Sheherazade"

Quantitative Tests – SDQ – Germany and Belgium

The SDQ served as the main element of the study from day one. It was used to categorise all pupils. On the basis of the individual subscales (prosocial behaviour, emotional problems, behavioural problems, behavioural problems with peers, hyperactivity and the total problem score (all subscales except prosocial behaviour) and the corresponding cut-off values, pupils were divided into groups (normal, borderline, conspicuous).

A table follows so that this is clear to the reader in the rest of the report:

SDQ (Subscale)	Maximum value	Cutoff normal	Cutoff borderline	Cutoff conspicuous
Prosocial behaviour	10	6-10	5	0-4
Emotional problems	10	0-3	4	5-10
Behavioural problems	10	0-2	3	7-10
Behavioural problems with peers	10	0-2	3	4-10
Hyperactivity	10	0-5	6	7-10
Total problem score (all subscales except prosocial behaviour)	40	0-11	12-15	16-40

These scores help to assess children's behaviour and emotional health and identify whether they may need support. As soon as the overall difficulties or one of the subscales reaches the 'conspicuous' range, there is a need for closer observation. The same applies to the other areas (normal, borderline).

The following results should be read and understood in conjunction with the previous explanation. Due to the different student populations, an average value was determined in all tests and a 'pooled standard deviation' was calculated. Some graphs appear (at first glance) to be effective or critical, but it is important to consider the **effect sizes**.

Cohen's d is a statistical measure that describes the size of an effect or the difference between two groups. It is usually used to show how large the difference between two average values is, relative to the standard deviation of the data. Cohen's d belongs to the family of effect size measures and is particularly common in psychology, medicine and social sciences to illustrate the practical significance of results, not just their statistical significance.

- 0.2: small effect (small difference between the groups)
- 0.5: medium effect
- 0.8: large effect (large difference between the groups)

This means that the greater the value of Cohen's d, the greater the difference between the two groups in relation to the spread of the data.

It is often used to go beyond pure statistical significance and to assess whether a difference is practically relevant or significant.

Quantitative Tests – SDQ – Germany and Belgium
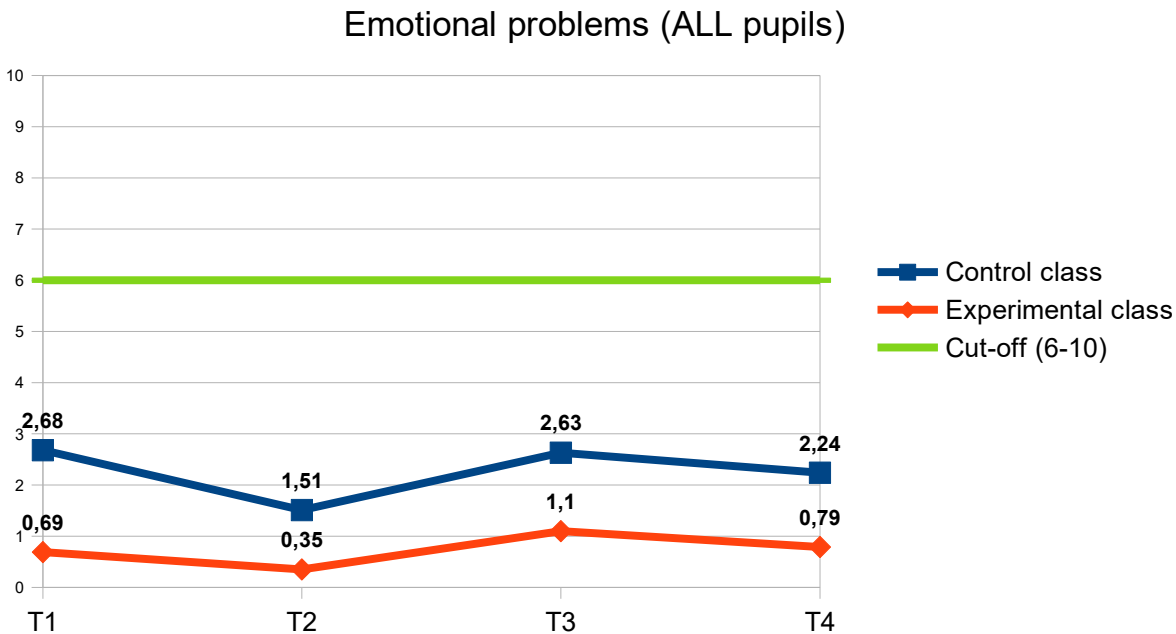
Prosocial behavior (ALL Pupils)
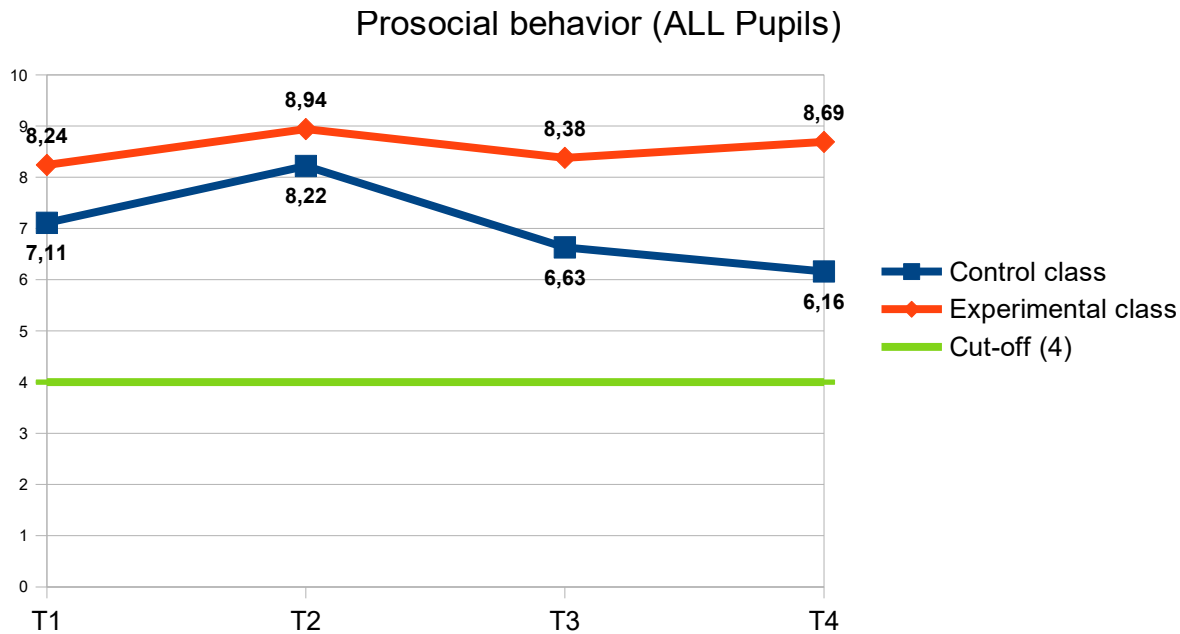
Prosocial behaviour refers to the strengths of individual pupils and should be considered separately from the other categories. The maximum value is 10 points and represents the maximum. If the value falls below the cut-off of 4 points, it is considered ʻconspicuousʼ. The higher the score, the better the prosocial behaviour.

As can be seen from the graph, the control classes start less well in comparison, but also deteriorate significantly. The experimental classes start better and improved in T4. The following values were obtained for the effect size:

Control group: d= 0.4 = negative small/medium effect

Experimental group: d= -0.22 = positive small effect

2021-1-BE03-KA220-SCH-000024196 – The Children of Sheherazade – A research and evaluation study of a storytelling project in primary schools

2

Quantitative Tests – SDQ – Germany and Belgium
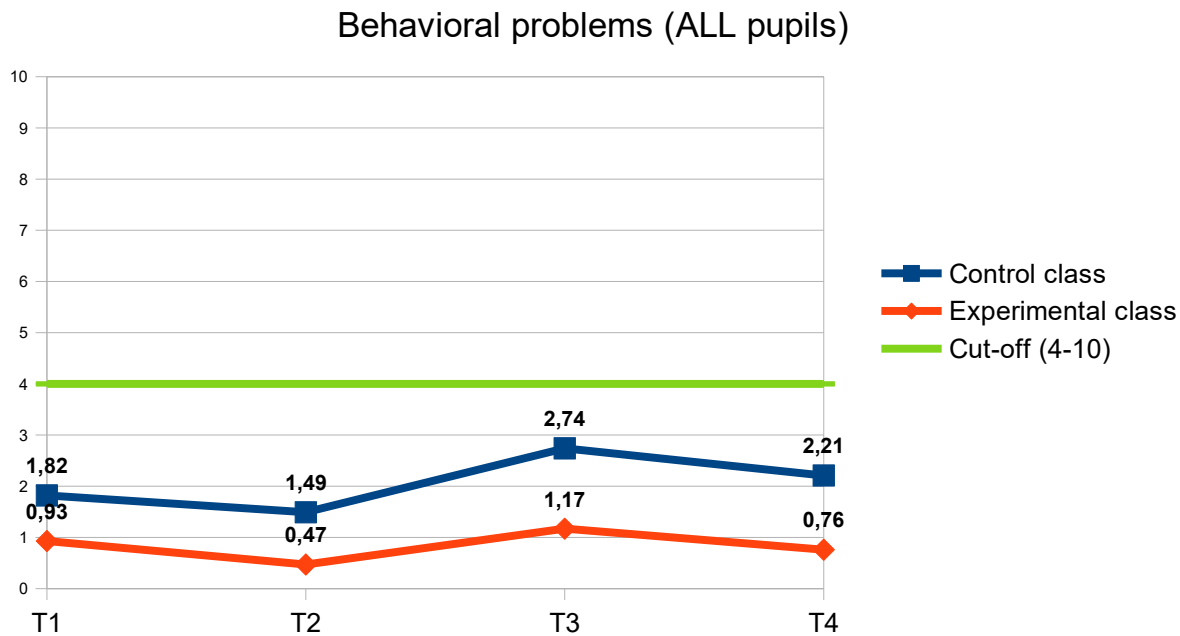
Emotional problems (ALL pupils)

The category behavioural problems with peers shows a maximum value of 10 points.
If the value is above the cut-off of 5 points, it is considered 'conspicuous'.

The control classes start out significantly more stressed due to the teachers'
assessment compared to the experimental classes, but the average value
weakens in T4, which represents a positive trend. The trend for the
experimental classes is almost the same.

Control group: d= 0.17 = minimal/small effect

Experimental group: d= -0.03 = no effect

Quantitative Tests – SDQ – Germany and Belgium

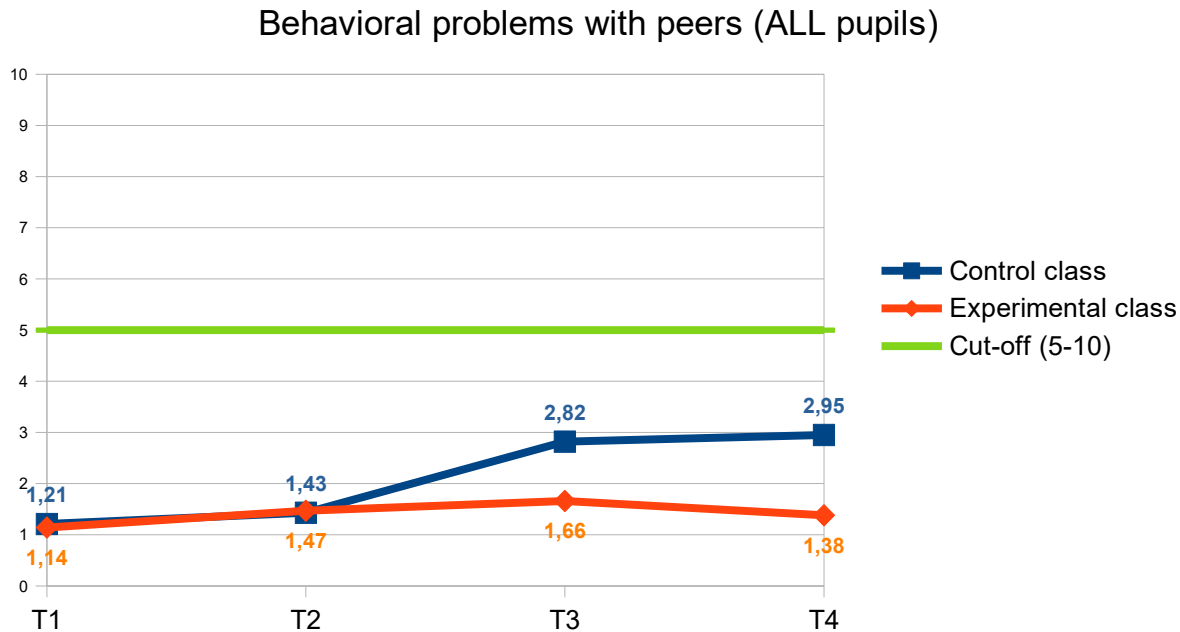### Behavioral problems (ALL pupils)

The category behavioural problems with peers shows a maximum value of 10 points.
If the value is above the cut-off of 5 points, it is considered 'conspicuous'.

The control classes start out more stressed due to the teachers' assessment
compared to the experimental classes and increase in value (behavioral problems),
which has a negative effect on development. In the experimental groups,
the value decreases (positive).

Control group: d= 0.2 = small effect (negative)

Experimental group: d= -0.04 = no effect/equal

Quantitative Tests – SDQ – Germany and Belgium

Behavioral problems with peers (ALL pupils)

The category behavioural problems with peers shows a maximum value of 10 points.
If the value is above the cut-off of 5 points, it is considered 'conspicuous'.

Both groups are classified in the same way according to the average value
of the teachers' assessment, but develop extremely markedly.
While the experimental classes show a minimal increase in scores
from T1 to T4, the control groups show a strong increase.
A look at the effect sizes shows:

Control group: d= -0.67 = large effect (negative)

Experimental group: d= -0.14 = small effect

Quantitative Tests – SDQ – Germany and Belgium

### Hyperactivity (ALL pupils)
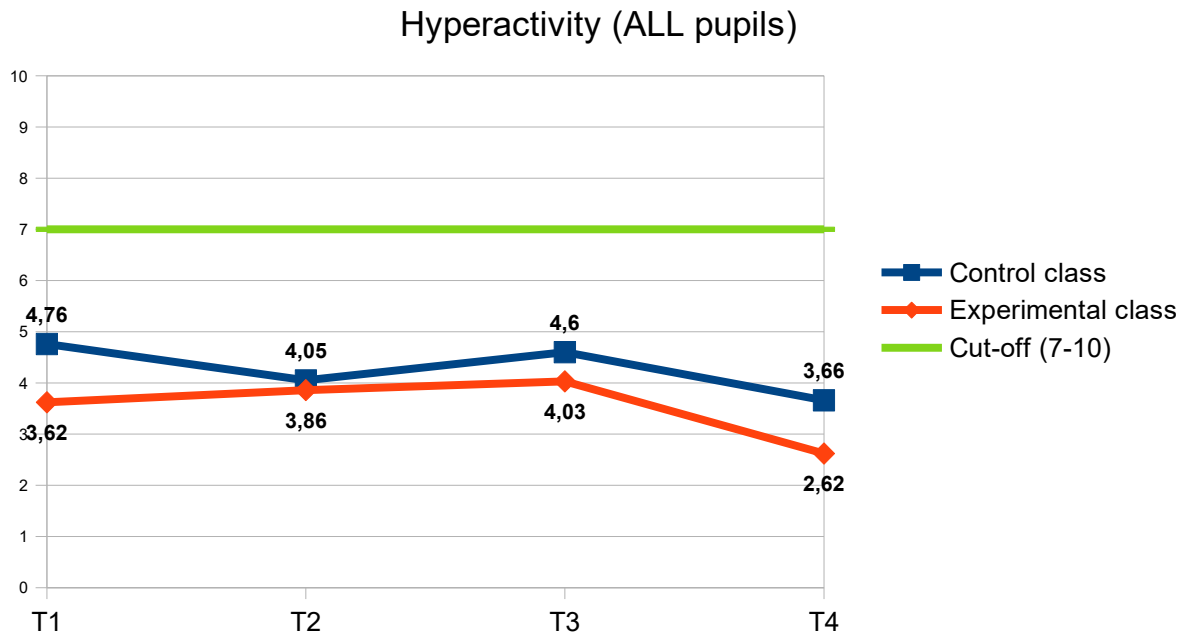
The hyperactivity category applies to all students in the tested classes. The maximum value is 10 points.
If the value is above the cut-off of 7 points, it is considered 'conspicuous'.

The graphs show a reduction in the value, which initially speaks positively for both groups. If we look at the pooled standard deviation, differences can be seen, so that the narrative classes perform significantly better over the three-year period.

Control group: d= 0.61 = clearly negative effect

Experimental group: d= -0.22 = clearly positive effect

Quantitative Tests – SDQ – Germany and Belgium

Overall problem value (all pupils)
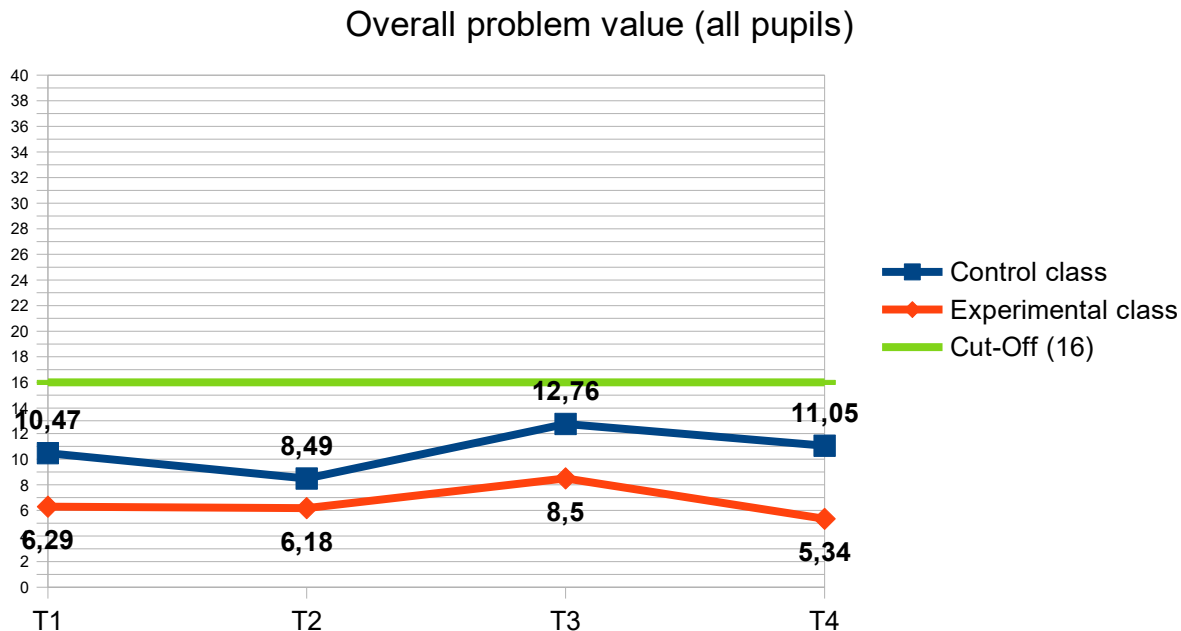
The overall problem value is calculated from the summed values of the subscales (emotional problems, behavioural problems, behavioural problems with peers/peers and hyperactivity – prosocial behaviour is excluded). The highest and therefore worst value is 40 – the lower the value, the better the result of the overall group.
The cut-off is 16; therefore, a score of 16 is to be understood as 'conspicuous'.

As can be seen from the graphs, the control classes started the project with a significantly higher overall problem score (as assessed by the teachers) at the start of the project (T1). If T4 is taken into account, the classes have deteriorated. The situation is different for the experimental classes, which started the project with a significantly lower problem score but also deteriorated positively in their development. According to Cohens-d, the following effect sizes can be determined on the basis of the pooled standard deviation:

Control groups: d= -0.08 = negative minimal effect

Experimental group: d= -0.15 = positive minimal effect