

University of Groningen

The tip of the iceberg & beyond

Visser, Marieke

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2011

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Visser, M. (2011). *The tip of the iceberg & beyond: Evaluation of an aggression reduction program for special elementary education.* [s.n.].

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Appendices Chapter 6

Appendix 6A. The Intervention Program

TRAffic 8-12 is a Dutch intervention program for 8 to 12 year old children who experience anger and conflicts on a daily basis. It is a typical social skills program such as is often being used in schools and therapeutic settings on a daily basis. The curriculum of TRAffic 8-12 is based on social-cognitive and social-learning principles. With the use of traffic signs such as the stop sign and the rotary section children learn how to control their anger and solve conflicts in a prosocial way. Imitation, reinforcement and transfer to daily life are important components of the program. Children follow 14 one-hour sessions either in groups of six children with two trainers (Group Trained: GT) or individually with one trainer (Individually-trained: IT). We were not fully able to randomly assign children to the group or individual condition because of practical reasons such as availability of trainers and preference of teachers. However, this did not result in significant differences in aggressive behavior between the GT and IT children at baseline T0 for the whole group ($p=0,21$ for the total score on the Aggressive Behavior Checklist).

The program is designed for children with the psychiatric disorders Attention Deficit Hyperactive Disorder (ADHD) and Pervasive Developmental Disorder-not otherwise specified (PDD-nos). In most social skill programs ADHD and PDD-nos are considered as contraindications. Children with ADHD are assumed to disturb the proper functioning of the intervention group because of their impulsive and hyperactive behavior (Van Manen, 2001). Children with PDD-nos are assumed to experience too much unsafety in an unfamiliar group to be able to learn anything, because of their difficulties with social situations. However, also these children may show aggressive behavior and might benefit from a social skills program. TRAffic 8-12 explicitly pays attention to the problems these children have with encoding and interpreting social information. The use of visual tools plays a very important role in TRAffic 8-12, which is especially important for children with PDD-nos. There are for example 'real' stop signs and rotary sections, an anger thermometer and a DVD with examples of how and how not to react in different situations. TRAffic 8-12 trainers were therapists, teachers, and psychology trainees. They followed a three-day training in which they were taught the theory and basics of TRAffic 8-12.

Appendix 6B. Measures

Aggressive behavior was assessed using the *Agressievragenlijst* (Aggressive Behavior Checklist) completed by the teachers (Krol, 1998). It is a Dutch checklist that measures the frequency of aggressive behavior in a school setting. It contains 26 items and is divided in four subscales: Physical Aggression (5 items; e.g. The child kicks or hits other children), Verbal Aggression (6 items; e.g. The child calls names to other children), Indirect Aggression (4 items; e.g. The child gossips about other children) and Negativism (11 items; e.g. The child annoys other children on purpose). Ratings are given on a 5-point Likert-scale: (almost) never = score 0, once per month = score 1, once per week = score 2, once per one or two days = score 3 and more than once per day = score 4. Scores are determined by summing up the scores per item within each subscale. Because the number and the degree of detail of the items differ considerably among subscales, we used weighted scores (subscale scores are weighted and then summed). Cronbach's α is .84 for Physical Aggression, .89 for Verbal Aggression, .84 for Indirect Aggression and .95 for Negativism.

Appendix 6C. Analysis, Statistical Procedure

For several reasons we decided to use random permutation techniques in our statistical analysis. First of all, some children in certain intervention groups were from the same classroom, making the sample partly dependent. Second, at each assessment we were confronted with missing data and at T3 and T4 the number of assessments became quite small. Third, variations in the sample groups were quite large. All of these constraints make it very difficult if not impossible to use conventional statistical techniques. Random permutation tests are much more flexible, making it possible to work with small and dependent samples with missing data (see for example Toddman & Dugard, 2001; Boosman, van der Meulen, van Geert & Jackson, 2002). A limitation of random permutation tests is the fact that it is a relatively laborious and unfamiliar technique, which is particularly used if the data set is 'messy'. However, the requirements of conventional techniques are hardly met in real developmental studies.

In a random permutation test, the empirical distribution is compared with a random distribution that is determined by randomly reshuffling the empirical data, in accordance with the null hypothesis. This reshuffling is carried out a great number of times (e.g. 10000 times). The resulting random distribution is a close approximation of the 'exact' null hypothesis distribution of the current dataset, given all its

peculiarities, such as small sample size. In the next step of the analysis, the empirical distribution is compared with the random distribution (null hypothesis). If both distributions differ significantly from one another, we can conclude that the empirical distribution is likely to differ from the distribution expected on the basis of the null hypothesis and that the differences between the groups that were compared are meaningful.

Appendix 6D. Analysis of the Study on Intervention Group Composition

The short-term effects of TRAffic 8-12 were determined by using T0 and T1 for the session 1 children and T1 and T2 for the session 2 children (from now on called T_{before} and T_{after} which covers a period of three months). Group-trained (GT) and individually-trained (IT) children were compared with respect to their change scores (T_{after} minus T_{before}). The null hypothesis predicted no difference between GT and IT children in the way they profit from the program, or, to put it differently, no dependence of the child's change scores on the group to which the child belongs (both groups are in fact drawn from the same underlying distribution). In order to determine the long-term results of TRAffic 8-12 T_{before} and T3 (13 months) were used. Here, we only included the children who did not transfer to a regular school at T3 in order to avoid confusion between effects of the intervention program and of the change in classroom environment. Furthermore, TRAffic 8-12 effects were analyzed within groups, comparing the children before and after the program with themselves. The null hypothesis was that the program has no effect and, consequently, that it does not matter whether we measure a child's behavior before or after the program.

We also calculated the effect sizes (Cohen, 1988) because in small intervention groups large difference scores are often not significant, while in fact there might be an effect in terms of improvement of targeted behavior. We calculated the effect size (ES) for the GT and IT children as follows. For each child the aggression score before the intervention was subtracted from the aggression score after the intervention. This difference score was then divided by the pooled standard deviation (weighted average standard deviation based on aggression scores before and after the program of all trained children). The average of these outcomes

per child for the GT and IT children gave us an idea of how the proportion of a standard deviation these two groups changed during the intervention program³⁹.

The results of the analyses described above are discussed in Chapter 6.

Effects of Psychiatric Diagnosis and IQ

We checked the possible differential effects of TRAffic 8-12 due to the psychiatric diagnoses of the children (ADHD and PDD-nos). First, we determined the change scores of the children by subtracting the total aggression score after TRAffic 8-12 from the total aggression score before TRAffic 8-12. By means of random permutation techniques, the difference between the observed change scores of children with ADHD, PDD-nos or a combination and the same scores of children with no diagnosis was compared with the distribution of change scores. The null hypothesis is that both groups (diagnosis versus no diagnosis) have similar change scores. The results are shown in Table 1. Both children with ADHD ($M=-7.83$) and children with PDD-nos ($M=-11.42$) did not differ significantly from children without a psychiatric diagnosis ($M=1.52$) with respect to their change scores ($p=0.48$ and 0.43 respectively).

We did not further test the significance of the difference in change scores between children with a combination of ADHD and PDD-nos and children without a diagnosis, because the difference was negligible. To conclude, children with ADHD, PDD-nos or a combination of both did not profit more or less from TRAffic 8-12 than children who do not have these psychiatric problems.

Table 1

Change scores of children with ADHD, PDD-nos and no diagnosis, with p-values, before - after

		change of children without a psychiatric diagnosis (s) ($N=19$)	p
change of children with a psychiatric diagnosis (s)	ADHD ($N=12$)	-7.83 (17.82)	0.48
	PDD-nos ($N=8$)	-11.42 (25.78)	
		1.52 (41.67)	0.43

³⁹ In recent years one has come to an understanding of the meaning of the level of effect sizes (ES). Generally, an ES smaller than 0.20 is considered negligible, an ES between 0.20 and 0.49 is called small, between 0.50 and 0.79 an ES is called mid-high, and an ES above 0.80 is considered high.

By means of calculating correlations, we also tested for differential effects of TRAffic 8-12 due to the IQ of the children. A high IQ might make children benefit more from TRAffic 8-12 than children with a low IQ. The correlation between change scores (after – before TRAffic 8-12) for the whole intervention group and IQ was -0.11. For group-trained and individually-trained children separately the correlation was -0.11 and -0.10 respectively. This means that there is no relation between IQ and change in aggressive behavior; a higher IQ is not accompanied by a higher reduction in aggressive behavior.

Appendix 6E. Analysis of the Study on Classroom Composition

We investigated whether the individual trends of development of aggressive behavior change when children transferred to 1) a school of regular education or 2) another Cluster 4 school. Option 2 was included in the analysis to check whether the observed changes are due to the school transition itself (irrespective of which type of school the child was referred to), or in particular to the transition to a school of regular education. The long-term assessments were carried out later in the school year in order to make sure that the behavior the children were showing was not temporarily adjusted behavior. In our analysis we used the total score on the Aggressive Behavior Checklist. Only if results were questionable, subscales were included in the analysis.

We wanted to test whether the trend of change in aggressive behavior, calculated for the period before school transition in Cluster 4 education, showed a downward direction after a school transition. It is important to note that such trends can show considerable individual differences: some children may show a downward trend, others an upward trend and still others are likely to be constant. Thus, our test focused on the effect of the school transition on the direction and magnitude of the trend. For instance, in a child with an upward trend (before the school transition) we expected to find at least a *decrease* in the upward trend. In a child with an already downward trend, we expected to find at least an *increase* in the downward trend. In order to check the changes in observed trend before the school transition, we proceeded as follows.

First we determined the linear model of aggressive behavior for each child during the Cluster 4 education period. To test whether the direction changed following the transition, we calculated whether the aggression after transition was lower than what could be expected on the basis of the child's own trend of

aggression change. The null hypothesis was that the new environment did not add anything to the trend already initiated, and thus, that the data points after transition would organize randomly around the trend line calculated on the basis of the data points before transition. In Figure 1 we show an example of how (the signs of) the residuals before and after transition are determined for one child. This was done for each child in the sample, after which the average of the residuals after transition was calculated.

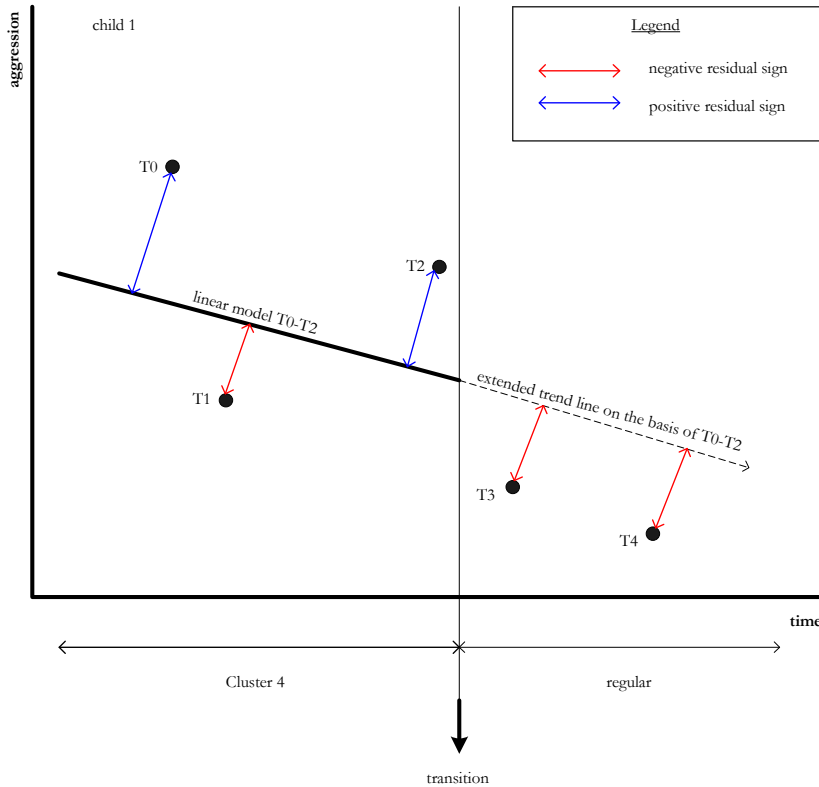


Figure 1. Graphical illustration of the method of analysis.

Null hypothesis: the data points after transition organize randomly around the trend line calculated on the basis of the data points before transition, in other words the trajectory of aggressive behavior does not change after transition to a regular school.

Alternative hypothesis: the data points after transition organize significantly more below the trend line calculated on the basis of the data points before transition compared to the data points before transition, in other words the ‘trajectory’ of aggressive behavior shows a significant descent after transition.

Under the assumption that there would be an equal chance of positive and negative residuals after transition compared to before, the signs of the residuals before transition were randomly permuted and each time multiplied with the absolute residuals after transition. The average of these residuals was compared with the average of the empirical distribution of residuals (after transition) with the use of random permutation tests. We repeated the analysis with an alternative null hypothesis, based on a fifty-fifty distribution of signs (there is an equal chance of a positive or a negative residual sign after school transition). We did this because the empirical distribution of signs before transition is slightly negatively biased. A fifty-fifty distribution gave us a more correct testing of the null hypothesis that the observed scores after the transition have an equal probability of falling above or below the observed trend.

The results of the analyses described above are discussed in Chapter 6.

Control Tests

Two control tests were performed using the same technique as we used in Study 1. These control tests are relevant because they may help to explain eventual changes in trends.

To control for initial differences between children who stayed in Cluster 4 education and children who transferred to a regular school, we tested whether the slopes before transition of children who transferred to a regular school differed from the slopes of children who stayed in Cluster 4 education. We justify the use of the slope as a statistical indicator of the trajectory on grounds of the fact that we had only few measurements (2 to 3) preceding the transition. With this number of measurements we must confine ourselves to describing linear regression models, which is characterized by a slope and an intercept. Instead of the intercept, we took the observed degree of aggressive behavior at baseline as an estimation of the child's initial level of aggression.

For the Stayers we calculated the slopes based on all assessments, for the Regulars and the Changers we calculated the slopes based on the assessments before transition. The random permutation test showed that the average of the slopes of the total scores on aggressive behavior did not differ significantly ($p=0.23$) between the Stayers ($M=2.46$) and the Regulars ($M=-4.45$). This finding implies that the children who transferred to a regular school did not show a significantly different trajectory before transition compared to the children who stayed in the same Cluster 4

education school, at least with respect to the average amount of decrease or increase in aggressive behaviors. The Changers ($M=-4.38$) also did not differ significantly from the Stayers with respect to the slope before transition ($p=0.13$).

A second control test was performed by testing whether the degree of aggressive behavior differed between the two groups at baseline T0. Combined with the information about the slopes, the second control test tells us if children who transferred to a regular school showed significantly less aggression before transition compared to children who stayed in the Cluster 4 school.

For each group we determined the average aggressive behavior score at T0 (baseline). It would have been more obvious to compare the aggressive behavior scores just before transition. However, because children transferred to other schools at different time points (T3 or T4) it was difficult to determine which values of the Stayers to compare with. The random permutation test revealed no significant difference ($p=0.06$) in aggressive behavior at T0 between the Stayers ($M=72.42$) and the Regulars ($M=54.98$). However, the result is close to significant, the Regulars obviously seemed somewhat less aggressive at baseline. The Changers ($M=79.38$) did not differ significantly from the Stayers ($p=0.47$).

It is important to note that, analytically speaking, these control tests are not necessary, because in our analysis we test whether there is a downward change after transition *relative* to the trajectory during the Cluster 4 period, irrespective of whether this trajectory is itself upward, flat or downward. On the other hand, it is still important to know if, with respect to children who transferred to regular education, we had to do with children who showed more improvement in behavior before their transition compared to children who stayed in Cluster 4 education. The results of the control tests showed that this was not the case.

Effects of Psychiatric Diagnosis and IQ

We tested whether the children who transferred to a regular school differed from the children who stayed in Cluster 4 education with respect to their psychiatric diagnoses (by means of crosstabulations) and IQ (by means of random permutation testing).

The children who transferred to a regular school did not seem to have less psychiatric diagnoses than the children who stayed in Cluster 4 education (see Table 2).

Table 2

Numbers and expected numbers of children with ADHD, PDD-nos, a combination and no diagnosis within the Stayers, Changers and Regulars groups

			Stayers	Changers	Regulars
diagnosis	ADHD	N	8	4	2
		exp N	8.2	3.3	3.4
	PDD-nos	N	6	2	3
		exp N	6.0	2.5	2.5
	combination	N	6	1	4
		exp N	6.0	2.5	2.5
	no diagnosis	N	10	6	4
		exp	10.9	4.6	4.6

$$\chi^2 = 3.36, p=0.76$$

As for IQ, the random permutation test revealed that the Regulars ($M=107$) had a significantly higher IQ than the Stayers ($M=96, p=0.06$) and the Changers ($M=92, p=0.04$). The Stayers and the Changers did not differ significantly in their IQ ($p=0.56$).

To conclude, the children who transferred to regular education had a higher IQ, but the same profile of psychiatric diagnoses as the children who stayed in Cluster 4 education.

