Social and emotional training in Swedish schools for the promotion of mental health: an effectiveness study of 5 years of intervention

Birgitta Kimber¹*, Rolf Sandell² and Sven Bremberg¹

Abstract

The school is an obvious arena for interventions designed to promote mental health among children. A set of educational techniques named social and emotional learning, which focuses on students’ self-control, social competence, empathy, motivation and self-awareness, has shown promising results in the United States. This is a study of the application of a similar method in Sweden (referred to as social and emotional training) for school years 2000/2001 through to 2004/2005. It is an effectiveness rather than an efficacy study, largely administered by school personnel, which relates duration of the training (1–5 years) to a set of outcomes previously found to be associated with mental health. Positive and significant effects were found on five of seven variables: internalizing problems, externalizing problems, mastery (reflecting self-efficacy or hopelessness), self-image and self-esteem and contentment in school. Effect sizes were medium. Somewhat surprisingly, no relationship was found between the intervention and the promotion of social skills. Nor was there any detectable long-term impact on bullying. Controlling for student gender did not moderate any of the effects.

Introduction

According to the World Health Organization, mental ill health is—alongside cardiovascular disease—one of the two most important public health problems worldwide. Among people aged 1–44 years, mental ill-health, which includes depression, aggressive behaviour, feeling down and alcohol and drug misuse, is the greatest health problem in high-income countries. Internalizing problems, such as anxiety and depression, have been shown to impose a greater burden on mental health than externalizing problems [1].

It is hard to predict who will develop mental disorders among the young, and this has implications for effective health promotion [2]. Since nearly all children go to school, the school is an obvious arena for health-promoting activities. A set of educational techniques, social emotional learning (SEL), based on the use by teachers of cognitive and behavioural methods, which concerns students’ self-control, social competence, empathy, motivation and self-awareness, has shown promising results [3–8]. As a mechanism, it may be worth linking the SEL approach to the increasingly used concept of ‘mentalization’, which, inter alia, stresses sense of agency, social reciprocity, self-regulation of affects, toleration of frustration, goal setting and capacity to symbolize [9].

Meta-analyses, e.g. that of Greenberg et al. [7], suggest that there are some shortcomings to school-based intervention programmes and studies in this field. Only a few include a broad range of school grades, although it has been claimed that ‘short-term preventive interventions produce time-limited benefits, at best, with at-risk groups whereas multi-year
programs are more likely to foster enduring benefits’ (p. 32). Further, most studies report on efficacy trials, undertaken with a research team in charge, rather than effectiveness trials in a community setting [10, 11]. Historically, externalizing problems have received greater attention, but the literature is rapidly growing on the prevention of internalizing problems [12–16].

This study of a Swedish programme (the SET programme, standing for social and emotional training) attempts to address these three issues. First, it considers the impact of the duration of SET (i.e., over a number of years). Second, it is an effectiveness rather than an efficacy study, since the programme has been implemented by school personnel in a real-life setting. Third, it includes internalizing mental health aspects as well as externalizing ones.

A longitudinal study [17] considered the effects of SET on students of all grades (junior and senior) during the first 2 years of programme implementation. SET was found to promote mental health and reduce detrimental health-related behaviours in some respects, in particular through the promotion of aspects of students’ self-image, including wellbeing and a reduction in externalizing problems and alcohol use. It was concluded that SET has the potential to operate effectively as a preventive intervention during the school period.

The aim of this second study of the programme, based on 5 years of data collection, is to explore the longer-term impact of SET on various aspects of the mental health and health-related behaviours of senior school students (Grades 5–9 at time of measurement). Drug and alcohol use will be considered in a separate paper. Specifically, the current study relates duration of SET to a variety of outcome variables.

---

**Methods**

**Population and sample**

In Sweden, children begin compulsory school at age 7 in Grade 1 and end at age 16 in Grade 9. The intervention and the study were carried out in Botkyrka Municipality in the Stockholm metropol-
they had an opportunity to try out the relevant exercises themselves and test them in their classes. They were encouraged to raise methodological and technical issues and discuss remaining problems. The teachers were supervised once a month during the school year 2000/2001 and offered supervision on a voluntary basis during 2001/2002. In order to enhance programme integrity, the teachers were observed and supervised individually at least four times during the first 2 years of the intervention. The views of teachers on the programme were ascertained in a survey conducted 2 years after programme start, which showed general, albeit not universal, approval of its content and implementation [20].

**Study procedures and instruments**

Assessments, by questionnaire, of a panel consisting of all the students who remained in the classes from outset of the SET programme, which started in August of school year 2000/2001, were made in May of each year, i.e. from 2001 (t1) until 2005 (t5). Questionnaires were administered to participants and controls during school hours by regular class teachers. All the instruments employed are well-established and have documented reliability and validity.

‘Youth self-report (YSR)’ [21], used here in an abbreviated Swedish version [22], measures mental health symptoms and problems. There are 35 items, which are split into two subscales, measuring internalizing problems and externalizing problems. Internalizing problems include anxiety (feeling worthless or inferior, feeling unhappy, sad or depressed and feeling nervous or tense), while externalizing problems encompass aggression (threatening to hurt people or destroying property), assertiveness (stubbornness, hot temper, etc.) and attention-seeking (bragging, showing-off, clowning, etc.). The items are rated on three-step scales: ‘Not True’, ‘Somewhat or Sometimes True’, ‘Very True or Often True’. Scale scores were means across items, with higher scores indicating more problems.

‘Mastery’ [23], in one of its Swedish versions, is a nine-item scale measuring feelings of self-efficacy or hopelessness, defined as the extent to which one regards one’s life chances as being under personal control. Examples include ‘There is really no way I can solve the problems I have’ and ‘I have little control over the things that happen to me’. Items are rated on four-step scales, ranging from ‘Strongly agree’ to ‘Strongly disagree’. The scoring of some of the items was inverted so that higher scores indicate stronger sense of self-efficacy.

‘I think I am’ (ITIA) is a Swedish self-rating instrument, ‘Jag tycker jag är’ [24], with roots in American research [25]. It is designed to assess the young person’s self-image and self-esteem and has subscales for body image, family relations, psychological well-being, relations with others and talents/abilities. There are two versions of the instrument: ITIA-I for younger students (Grades 1–3) and ITIA-II for older ones (Grades 4–9). Only findings from ITIA-II are reported here. In ITIA-II, there are 72 items, such as: ‘I have a nice face’, ‘I like myself’, ‘I am often sad’ and ‘My parents trust me’. Students respond to each statement on a four-point scale, from ‘Exactly like me’ to ‘Not at all like me’. The items are scored −2 to 2, where higher scores indicate a more positive self-image. There are a number of subscales, but in this study we only used the mean score across all items.

‘Contentment in school’, or school satisfaction, by analogy with job satisfaction, refers to a single item, ‘How do you like it in school?’, from a Swedish health-behaviour questionnaire administered annually by the Swedish Council for Information on Alcohol and Other Drugs (CAN) [26]. Contentment was rated on a five-step response scale, ranging from ‘Very good’ to ‘Very bad’. Scoring was inverted so that higher scores indicate greater contentment.

‘Bullying’ is the mean of three items from the CAN questionnaire [26]. The aspects considered are being insulted, being physically assaulted and ‘being sent to Coventry’. A three-step response scale was used: ‘Yes, often’, ‘Yes, sometimes’ and ‘No, seldom or never’. Higher scores indicate fewer problems.

‘The Social Skills Rating System (SSRS)’ [27] consists of 34 items for Grades 4–6 and 7 additional
items for Grades 7–9, all with four-point response scales, ‘Never’ (0), ‘Sometimes’ (1), ‘Often’ (2) and ‘Very often’ (3). The ratings may be scored on four subscales (assertion, empathy, cooperation and self-control), but in this study the mean score across all items was used. Higher mean scores indicate greater social skills.

The instruments and scales and their reliabilities and re-test correlations over the first 2 years of the intervention are shown in Table I. It should be noted that their intercorrelations were strong enough to produce a general factor, accounting for 49% of the total common variance before rotation. Two more factors each accounted for >10% of the variance. We concluded that the battery was essentially multifactorial, with a general adjustment factor (mastery, ITIA, SSRS, YSR internalization), one school adjustment factor (school contentment, bullying, SSRS) and one mental health factor (YSR internalization, YSR externalization, bullying) after rotation. The reliabilities of the mastery and bullying scales were poor, probably due to the relatively small number of items. Nevertheless, we decided to use both scales for exploratory reasons.

### Study design

For the current study, we employed a mixed design, in which there is ‘a mixture of between-group and repeated-measures variables’ [28] (p. 483) to compare students in the SET and No-SET schools according to duration of SET or No-SET (1–5 years), regardless of grade. Given a student’s grade at $t_1$, $t_2$, $t_3$, etc., we formed a variable for duration of the SET programme (number of years). Thus, the SET students in Grade 5 when the questionnaire was administered at $t_1$ had had 1 year of SET; students in Grade 5 at $t_2$ (Grade 4 at $t_1$) had had 2 years of SET, students in Grade 5 at $t_3$ (Grade 3 at $t_1$) had had 3 years and so on. We were able to secure a sizable number of observations (2455 in total), for SET (1857) and No-SET (598), distributed across years. We then compared the mean trajectories on each outcome measure between students in the SET schools and the No-SET schools as a function of the number of years that the programme had been running.

### Statistical analyses

Differences between the groups (SET and No-SET) in their development from $t_1$ to $t_5$ were tested in three different ways. SPSS version 12 was used for the statistical calculations.

(i) For each of the outcome variables, a linear regression was performed for each student group, which provides measures of the linear trends as effects of the intervention. Thus, each outcome variable was regressed on number of years (at $t_1$, $t_2$, ..., $t_5$) among the SET students and the No-SET students separately. This permitted direct tests of the trend-across-years differences between the groups,
based on estimates of the mean change rates in the groups provided by the unstandardized slope parameters, \( b_s \). The standardized (\( z \)-transformed) regression coefficients, \( \hat{b}_s \), were taken as within-conditions (or within-groups) effect-size estimates, and the differences between the unstandardized regression coefficients, \( b_s \), were tested according to Cohen [29]:

\[
\hat{z} = \left( \frac{b_1 - b_2}{\sqrt{SE_1^2 + SE_2^2}} \right) \frac{1}{\sqrt{2}}.
\]

(ii) Adopting Becker’s approach [30], between-groups effect sizes were computed for each dependent variable. A within-group \( d \) was first computed for each intervention group by dividing the \( t_5 - t_1 \) mean difference by the standard deviation at \( t_1 \) and then subtracting the \( d \) in the No-SET group from that in the SET group. This gives a change effect size parameter (Becker’s \( \Delta \)). Cohen’s classification of effect sizes (small = 0.2, medium = 0.5, large = 0.8) was employed [29].

(iii) Analyses of variance (ANOVAs) (or multivariate analyses of variance, MANOVAs, when we analysed an instrument with subscales, such as the YSR and the ITIA) were run on the outcome scale (or subscales), with intervention or not (SET or No-SET), number of years (\( t_1, t_2 \ldots t_5 \)) and student gender as independent variables. It was presumed that there is no SET/No-SET by study years or grade interaction, which assumes that duration of SET acts independently of age of student. Although the design was mixed, partly within subjects and partly between subjects, duration (number of years) was analysed as a between-subjects factor, which introduced a conservative bias to the tests. Given significantly different mean changes of the unstandardized regression coefficients, the critical effect was the intervention-by-years interaction. The generalized linear modelling (GLM) routine of SPSS, version 12, was used.

### Results

**Attrition**

Since the study was based on the responses of a panel, there was progressive sample attrition over the years due to normal turnover. Also, there was variable, temporary absence of students at time of testing, which in some cases resulted in more respondents in one year than the year before. See Table II.

Attrition can be determined by looking at the diagonals in Table II. For example, at \( t_1 \) (May 2001), there were 138 Grade 5 respondents out of the 141 students in the SET classes that year (upper left cell of Table II, upper part: Grade 5 at \( t_1 \)), while

<table>
<thead>
<tr>
<th>Table II. Frequency distribution of respondents across grades and number of years (sample sizes in parentheses(^a))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grade</strong></td>
</tr>
<tr>
<td>SET</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>No-SET</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

The table can be read across rows or along columns, in which cases the comparisons are between different student cohorts (between subjects); reading along the upper left to lower right diagonals enables comparisons between students from the same classes over time (within subjects), which show the rates of progressive panel mortality.

\(^a\)That is, both respondents and non-respondents in the classes under study.
at \( t_5 \) (May 2005), there were 44 respondents out of 68 (lower right cell: Grade 9 at \( t_5 \)); these were the remaining original students (from Grade 5 at \( t_1 \)) who had progressed to Grade 9. Overall, there was gradual dropout of 15–19% each year in the SET group. The overall response rate (rate of return of fully filled-in questionnaires) for the SET group was 88%, and for the No-SET group 77% (figures not shown in Table II). Over the years, response rates for SET ranged from 65 to 100% and for No-SET from 51 to 100% (excluding a single case where there was only one respondent, due to administrative error). Response rates in the No-SET group were more variable than in the SET group.

We needed to establish whether our results were biased by selective sample attrition, i.e. whether any apparent positive or negative effects of SET might instead be attributable to differences in dropout between high-scoring and low-scoring students. Improving mean scores along the diagonals (e.g. from Grade 5 students at \( t_1 \) to Grade 9 students at \( t_5 \), i.e. the same students minus dropouts) would imply greater dropout of low-scoring students (negative attrition), and deteriorating scores greater dropout of high-scoring students (positive attrition).

ANOVA was conducted on five of the outcome scales (YSR total, mastery, ITIA total, contentment in school and social skills) for each of the seven cohorts (diagonals) in the SET group. There were significant changes on only seven of the 35 resulting tests. Of these, four were cases of negative attrition and three of positive attrition. We concluded that attrition in the SET sample had not been biased towards either low- or high-scoring students.

Findings

An overall picture of developments over the years is given in Fig. 1. All relationships were in the expected direction; that is, SET students consistently reported more favourable outcomes over time than did NoSET students. There was a time lag with regard to some of the effects, and there were greater fluctuations in effects according to duration in the No-SET group.

Regression coefficients and effect sizes by outcome variable and group are shown in Table III. Of the \( \beta \)s, differences between the two groups were statistically significant on all but two (bullying and social skills) of the seven outcome variables. On the five significant variables (YSR internalizing, YSR externalizing, mastery, ITIA total and contentment in school) effect sizes (\( \Delta \)) were small to medium on Cohen’s criteria [29]. It should be noted that \( \beta \)s between 0.10 and 0.15 correspond to \( ds \) between 0.20 and 0.31 and that \( \Delta \) is the difference between the \( ds \) of the two groups.

It should be recognized that the intercorrelations among the variables and their sheer number may have generated redundant or spurious significances. A reasonable Bonferroni correction would accept two-starred effects (\( P < 0.01 \)) as significant only at the \( P < 0.05 \) level.

To test the effects of the possible dependencies in the data resulting from the complex design, partly crossed, partly nested, we repeated the same analyses, first, with classes and, finally, with schools as units. These analyses yielded essentially very similar estimates of the unstandardized regression coefficients but substantially larger differences between the standardized regression coefficients in the SET and No-SET groups, in comparison with the original analyses. We concluded that within-group dependencies had not exaggerated the between-group differences.

The results of the ANOVAs are displayed in Table IV. These show the statistical significances of the interactions between SET/No-SET and the durations (1–5 years) in relation to the outcome variables rather than the linear trends (in which there is just one comparison for each variable). All but one of the seven interactions were found to be statistically significant, the exception being social skills. On the YSR, there was a significant SET/No-SET-by-years interaction. The interaction was significant on both subscales, although with a stronger effect on internalizing than on externalizing problems.

The total ITIA score showed a univariately significant SET/No-SET-by-years interaction, but there was a clear differentiation between the groups only at \( t_5 \). It should be noted that there were significant univariate effects on three of the subscales, i.e. body image, \( F (4, 2252) = 3.71, P = 0.005 \); relations with others, \( F (4, 2252) = 3.33, P = 0.010 \) and
well-being, $F(4, 2252) = 2.64, P = 0.032$ (figures not shown in Table IV). Student gender did not complicate the two-way interaction on any of the scales. Contentment in school also showed a significant univariate SET/No-SET-by-years interaction, with no further interaction with student gender.

Fig. 1. Relations between duration of SET/No-SET and the outcome variables, with raw scores on the vertical axes and number of years on the horizontal axes.
Although bullying had a small effect size, it showed a highly significant SET/No-SET-by-years interaction. This is explained by the fact that there was a stable mean level over the years in the SET group and a quite variable one in the No-SET group, which ends up only slightly higher than the SET group at \( t5 \). Student gender did not affect developments in the two groups.

The social skills scale did not differentiate between the SET and No-SET groups. None of the subscales even approached a significant SET/No-SET-by-years interaction.

### Discussion

The impact of an SEL intervention in Sweden (the SET training programme) showed generally favourable results in the longer run, as shown by the comparisons between the SET and No-SET regressions.

Relating duration of social emotional training to various outcomes associated with mental health, significant positive connections were found on five out of the seven dependent variables considered: YSR internalizing, YSR externalizing, mastery, ITIA (total) and contentment in school. Effect sizes were medium.

In the SET schools, bullying was at a continuously low level, whereas in the No-SET schools, the level varied strongly from year to year. Considering duration, it was found that there was no difference in trend between the SET and No-SET groups. SET may offer a means of providing greater continuity in this arena in that peak incidences in the level of bullying are avoided.

SET students may have become more familiar than their No-SET counterparts with the ‘vocabulary’ of the SET questionnaire, and this reflected itself in greater consistency of questionnaire responses. The issue of what is artefactual and what is not is of substantive importance; enhancing capacity to give names to feelings has for long been regarded as a promoter of mental health [31].

The current study reveals significant duration lags on some variables. It now appears that there is a greater beneficial effect of SET on internalizing than externalizing problems, but this only

### Table III. Regression coefficients and effect sizes for SET/No-SET by groups and outcome variables

<table>
<thead>
<tr>
<th>Outcome variables</th>
<th>SET</th>
<th></th>
<th>No-SET</th>
<th></th>
<th>( \Delta )</th>
<th>( z_b )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta )</td>
<td>( d_{1-5} )</td>
<td>( \beta )</td>
<td>( d_{1-5} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YSR, internalizing</td>
<td>0.14</td>
<td>0.37</td>
<td>-0.05</td>
<td>-0.19</td>
<td>0.56</td>
<td>4.02***</td>
</tr>
<tr>
<td>YSR, externalizing</td>
<td>0.13</td>
<td>0.37</td>
<td>0.03</td>
<td>-0.05</td>
<td>0.42</td>
<td>2.16*</td>
</tr>
<tr>
<td>Mastery</td>
<td>0.11</td>
<td>0.42</td>
<td>0.00</td>
<td>0.06</td>
<td>0.36</td>
<td>2.42**</td>
</tr>
<tr>
<td>ITIA, total</td>
<td>0.11</td>
<td>0.44</td>
<td>-0.03</td>
<td>-0.10</td>
<td>0.54</td>
<td>2.74**</td>
</tr>
<tr>
<td>Contentment in school</td>
<td>0.06</td>
<td>0.19</td>
<td>-0.11</td>
<td>-0.41</td>
<td>0.60</td>
<td>3.57***</td>
</tr>
<tr>
<td>Bullying</td>
<td>0.03</td>
<td>0.11</td>
<td>-0.02</td>
<td>-0.24</td>
<td>0.35</td>
<td>0.72</td>
</tr>
<tr>
<td>Social skills, total</td>
<td>0.11</td>
<td>0.26</td>
<td>0.10</td>
<td>0.19</td>
<td>0.07</td>
<td>0.15</td>
</tr>
</tbody>
</table>

All figures in the table are adjusted to show positive values to indicate improvement and negative values to indicate deterioration. \( \beta \) = standardized regression coefficient when the outcome variable is regressed on years \((1–5)\), \( d_{1-5} = (M \text{ year 1 – M \text{ year } 5})/\text{pooled SD} \). \( \Delta = d_{1-5} \text{ in SET} – d_{1-5} \text{ in No-SET} \). \( z_b \) refers to the difference between the SET and No-SET \( \beta \)s.

*\( P < 0.05 \), **\( P < 0.01 \), ***\( P < 0.001 \) (one-tailed tests).

### Table IV. F-tests of interactions between SET/No-SET and number of years by outcome variable

<table>
<thead>
<tr>
<th>Outcome variable</th>
<th>Degrees of freedom</th>
<th>( F )</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>YSR, internalizing</td>
<td>4, 2337</td>
<td>4.86</td>
<td>0.001</td>
</tr>
<tr>
<td>YSR, externalizing</td>
<td>4, 2337</td>
<td>2.50</td>
<td>0.041</td>
</tr>
<tr>
<td>Mastery</td>
<td>4, 2340</td>
<td>3.43</td>
<td>0.008</td>
</tr>
<tr>
<td>ITIA, total</td>
<td>4, 2253</td>
<td>3.48</td>
<td>0.008</td>
</tr>
<tr>
<td>Contentment in school</td>
<td>4, 2312</td>
<td>4.77</td>
<td>0.001</td>
</tr>
<tr>
<td>Bullying</td>
<td>4, 2255</td>
<td>3.86</td>
<td>0.004</td>
</tr>
<tr>
<td>Social skills, total</td>
<td>4, 2356</td>
<td>0.71</td>
<td>0.588</td>
</tr>
</tbody>
</table>
emerged after 3–4 years. In the case of mastery (which, *inter alia*, measures depressive feelings), 3 years of SET seem to have been needed before the programme had a detectable impact and in the case of the ITIA (which measures self-image and self-esteem) 4 years. It seems that SET, as a form of health promotion, requires a long time of regular systematic implementation to be effective. This is in line with earlier studies [7]. It should be re-emphasized that in a real-life effectiveness study, beneficial outcomes may take longer to appear or detect.

Social skills is a remarkable exception to the rest [27]. Maybe, as suggested by Durlak and Wells in their review [6], SEL programmes have a greater impact on emotional than on social skills.

Although the repeated-measures analyses were cross-sectional, the sample on which these analyses were performed was subject to attrition. Obviously, some SET participants and controls did not respond >5 years or even >2 years. We have, however, shown that the differential outcomes between the SET and No-SET groups cannot be explained away by selective attrition within the SET group, i.e. that students with poorer mental health were less likely to respond over longer periods. Also, we know that some positively scoring students moved to a new school (a so-called ‘free’ school, with higher academic ambitions) when it opened in the vicinity of the SET schools. Although there was also some movement from the No-SET schools, which were further away, the recruitment differential might mean that the recorded positive effects of SET were underestimated.

There are some issues with reference to applications for practitioners. If SET has a positive value, should it be a regular part of the school curriculum? If so, how much training of the teachers is needed and how much parental involvement is necessary?

The study revealed some clear positive effects of SET on mental health. Continued analysis of the SET programme is ongoing with regard to drug, alcohol, volatile-substance use (sniffing) and smoking outcomes and to issues of implementation and attitudes of the SET teachers.

---

**Funding**

Funding for the research was provided through grant 2001–1072 from the Swedish Council for Social Research (SFR) and grant 2003–2785 from the Swedish Council for Working Life and Social Research.

---

**References**


Received on October 15, 2007; accepted on June 11, 2008