# Teacher Report of Children's Problem Behavior on The Sutter-Eyberg Student Behavior Inventory—Revised (SESBI-R) in a Norwegian Sample of Preschool and School Children

Bente Kirkhaug [a] , May Britt Drugli [a] , Willy-Tore Mørch [b] & Bjørn Helge Handegård [b]

[a] Norwegian University of Science and Technology

[b] University of Tromsø

Published online: 24 Nov 2011.

PLEASE SCROLL DOWN FOR ARTICLE

Routledge
Taylor & Francis Group

# Teacher Report of Children's Problem Behavior on The Sutter– Eyberg Student Behavior Inventory—Revised (SESBI-R) in a Norwegian Sample of Preschool and School Children

Bente Kirkhaug and May Britt Drugli
Norwegian University of Science and Technology

Willy-Tore Mørch and Bjørn Helge Handegård
University of Tromsø

In the present study, the applicability of The Sutter–Eyberg Student Behavior Inventory— Revised (SESBI-R) was explored within a Norwegian sample of 983 children aged 3–8 years. By using an exploratory factor analysis the same 2-factor solution as the original SESBI-R was supported by Principal Component Analysis. Good internal consistency (alpha > .96) was found for both the SESBI-R Intensity and Problem scales. Teachers in Norway report relatively low levels of behavior problems among children in day-care units and school-aged boys receive higher Intensity and Problem scores than girls. The results show that the SESBI-R is a reliable assessment tool that can be used in the Norwegian population for efficient screening of behavior of conduct-disordered children.

*Keywords:* Sutter–Eyberg Student Behavior Inventory—Revised (SESBI-R), behavior problems, teacher report, Norwegian sample

Behavior problems such as defiance, aggressiveness, noncompliance, and impulsiveness are currently the most common reasons for referring young children to mental health services (Watt, Hoyland, Best, & Dadds, 2007). Increasing evidence suggests that early onset behavior problems are persistent and associated with enhanced risk of impaired functioning, such as dropping out of school, criminal involvement, drug abuse, and violent behavior later in life (Broidy et al., 2003; Odgers et al., 2008). Young children with high levels of behavior problems appear to be at the highest risk of persistent and serious mental health problems (Loeber & Farrington, 2000). It is therefore important to have valid assessment tools to distinguish between children with mild and severe behavior problems as well as to indicate the appropriate approach to intervention. Intervention approaches will not be taken into consideration in this article. Campbell et al. (2006) found that behavior problems addressed at pre-school age may lead to normal mental health and school functioning by the age of nine.

Bente Kirkhaug, Institute of Neuroscience, Norwegian University of Science and Technology; May Britt Drugli, Institute of Neuroscience, Norwegian University of Science and Technology; Willy-Tore Mørch, Institute of Clinical Medicine, University of Tromsø; Bjørn Helge Handegård Institute of Clinical Medicine, University of Tromsø.

Correspondence concerning this article should be addressed to Bente Kirkhaug, Regional Centre for Child and Adolescent Mental Health, Department of Neuroscience, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. E-mail: bente.kirkhaug@ntnu.no

To get a better understanding of child psychopathology, it is important to standardize assessment measures across different countries and cultures (Achenbach & Rescorla, 2001). When data are available from many societies, they can be analyzed together to compare variations between societies and within societies (Rescorla et al., 2007). In studies of child mental health, in particular those using parental evaluations, Scandinavian parents consistently report lower problem scores for children than parents in most other countries, including those from the USA (Ivanova et al., 2007; Reedtz et al., 2008). However, studies based on teacher reports in school samples have shown mixed findings. For example, Rescorla et al. (2007) reported that teachers from Finland rated children as having lower total Problem scores on the Teacher Report Form (TRF) compared with the grand mean in a study conducted in 21 other countries, while teachers from Denmark rated the children at levels very similar to the grand mean. Further, it was found that Norwegian teachers rated their students as having higher total scores on the TRF than teachers from Finland, but lower than teachers from Denmark and the grand mean (Larsson & Drugli, 2011). There are few standardized instruments for assessing child behavior problems among preschool-aged children (Funderburk, Eyberg, Rich, & Behar, 2003). To date, no studies have explored teacher-reported levels of child behavior problems in day-care units in different countries. Taken together, these findings emphasize the importance of developing effective methods to assess children's behavior problems at all stages of development and across a variety of settings, and the importance of investigation of the applicability of an instrument when it is to be used in a new country.

The Sutter–Eyberg Student Behavior Inventory—Revised (SESBI-R) has been used by teachers to evaluate behavior problems in children aged 2–16 and can differentiate among different levels of behavior problems. It is designed to assess the severity of behavior problems and the extent to which teachers find the behavior worrying (Eyberg & Pincus, 1999). The original version of the SESBI was constructed as a companion instrument to the Eyberg Child Behavior Inventory (Eyberg & Pincus, 1999), a parent-rating scale of behavior problems. The SESBI has demonstrated adequate psychometric properties (Funderburk & Eyberg, 1989), but was later revised to assess whether the representation of the SESBI conduct behaviors could be strengthened by closer estimates of the disruptive behavior disorders from the DSM-III-R (American Psychiatric Association, 1987). Eight of the original items were therefore replaced with ten new items.

In a comparison study of the original and revised version, the SESBI-R was found to be better than the SESBI in its item-distribution variability, with a more stable factor structure (Rayfield, Eyberg, & Foote, 1998). Rayfield et al. found that although the SESBI-R is in general a measure of disruptive behavior problems, with the use of only a total-score and no subscales, it has two distinct factors that can be measured: oppositional behavior and attentional difficulties. While significantly lower scores on the teacher-rated SESBI-R Intensity and Problem scales were found for girls than for boys, children's age did not influence the scores. Analyses also showed that scores on the Intensity scale were least affected by individual teacher differences.

The SESBI-R has US norms for children aged 2–16, with available percentile ranks and approved translations into nine different languages: Welsh, Australian, Spanish, German, French, Arabic, Finnish, Chinese, and Norwegian (National Child Traumatic Stress Network, 2010). The SESBI-R may be a useful instrument in research and clinical work focusing on child behavior problems in Norwegian day-care units and schools. However, the applicability of the SESBI-R in the Norwegian context needs to be explored.

The aims of the present study of a Norwegian sample of children aged 3–8 years were to examine: (1) psychometric properties of the SESBI-R in regard to internal consistency (alpha coefficient); (2) its factor structure by means of principal components analysis (PCA); and (3) differences in scores by children's sex and age.

## Methods

### Procedures

The study was conducted in a Norwegian sample of municipal day-care units in the city of Trondheim and municipality schools in the city of Tromsø, including both urban and rural areas from both districts. All day-care units and schools in the municipalities were invited to participate by sending a request to the day-care unit managers and school principals. The subjects were randomly selected within each day-care unit and school classes. The selection of subjects in the day-care units was stratified according to the children's sex and age (3, 4, and 5 years), and the selection of subjects in school classes was stratified according to the children's sex. To reduce dependency in teacher evaluations, each teacher filled in a maximum of six questionnaires. A contact person within each day-care unit and school distributed information about the study and all material needed to the teachers responsible. The teachers distributed information to parents with a request for permission for the child to participate in the study and with the option of withdrawing the child from the study at any time, after informing the teacher. The questionnaires were returned to the research group in a prepaid envelope or through the Internet survey tool Questback (http://www.questback.no/). Brief demographic questions were included regarding each child's age, sex, and ethnicity. A list of randomly selected reserve children was given to the day-care units and schools, in case some children originally selected were refused permission to participate. The teachers were allowed to send one parental reminder to obtain consent. To increase the validity of teacher assessment of the children, the study was conducted in the early spring semester, allowing teachers at least six months of acquaintance with the pupils.

### Participants

A total of 33 day-care unit managers and 16 principals agreed to participate in the study, and yielded a response rate of 60 and 57%, respectively. A total of 737 day-care unit children were invited to participate in the study, and 555 questionnaires were completed (75.3%), 276 boys and 279 girls (Figure 1). A total of 536 school children were invited to participate and 428 questionnaires were completed (79.8%), 212 boys and 216 girls (Figure 1). Of the participants, 95% were ethnic Norwegian and 5% were of non-ethnic Norwegian origins. Because the non-ethnic Norwegian participants constitute such a small amount, no additional analyses were conducted on ethnicity. The only information collected about the non-responding day-care units and schools selected was geographical district, and no significant difference in responses was found in the day-care units and schools of either Trondheim or Tromsø. Complete data on the SESBI-R was obtained for a total of 983 children.

### Instrument

The SESBI-R is constructed to assess current frequency and severity of behavior problems in school and day-care unit settings, as well as the extent to which teachers find the behavior troublesome (Eyberg & Pincus, 1999). The SESBI-R is designed to identify
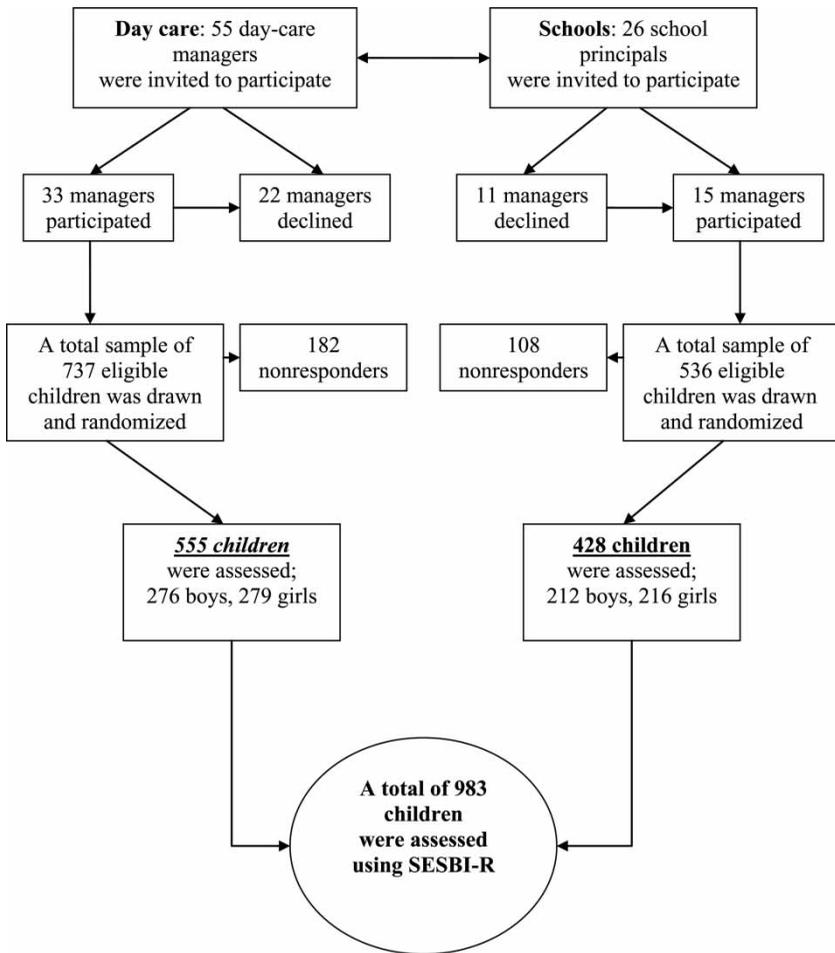
*Figure 1.*   Flowchart of participants in day-care units and schools.

children aged 2–16 years who need treatment for their behavior problems (Querido & Eyberg, 2003) and consists of 38 items. These items represent common behavior problems that are observable by teachers (Table 1). The teacher assesses current frequency of each behavior on a 7-point Intensity scale: 1 = never, 2–3 = seldom, 4 = sometimes, 5–6 = often and 7 = always. In addition, the teacher assesses whether or not the behavior is currently a problem on a yes–no (0–1) scale, the Problem scale. A total score is computed on both the Intensity and Problem scales. The scores range from 38 to 266 on the Intensity scale and from 0 to 38 on the Problem scale. For more detailed information on the scoring process of the SESBI-R, refer to page six of the Professional Manual (Eyberg & Pincus, 1999). The SESBI-R has been found to be a reliable and valid instrument for efficiently screening and tracking the behaviors of conduct-disordered children. Cronbach's alpha for the Intensity scale was found to be .98, with an average corrected item to total correlation of .76, and the Cronbach's alpha for the Problem scale was .96, with an average corrected item to

Table 1

*Corrected Item−Total Statistics for the SESBI-R*

| Item | Teacher reporting item as a problem % | Frequency rating | | Corrected item−total correlation |
|---|---|---|---|---|
| | | M | SD | |
| 1. Has temper tantrums | 8.2 | 2.11 | 1.33 | .70 |
| 2. Pouts | 8.9 | 2.57 | 1.31 | .67 |
| 3. Teases or provokes other students | 12.6 | 2.52 | 1.39 | .76 |
| 4. Lies | 7.4 | 2.00 | 1.15 | .66 |
| 5. Acts frustrated with difficult tasks | 6.8 | 2.46 | 1.36 | .65 |
| 6. Does not obey school rules on his/her own | 10.2 | 2.36 | 1.37 | .70 |
| 7. Demands teacher attention | 11.5 | 3.49 | 1.47 | .60 |
| 8. Dawdles in obeying rules or instructions | 16.2 | 2.95 | 1.47 | .77 |
| 9. Acts bossy with other students | 11.3 | 2.67 | 1.49 | .59 |
| 10. Gets angry when doesn't get his/her own way | 7.9 | 2.52 | 1.38 | .71 |
| 11. Interrupts teacher | 10.0 | 2.56 | 1.38 | .71 |
| 12. Impulsive, acts before thinking | 7.8 | 2.45 | 1.33 | .78 |
| 13. Refuses to obey until threatened with punishment | 6.0 | 1.74 | 1.12 | .72 |
| 14. Has difficulty staying on task | 10.3 | 2.53 | 1.38 | .68 |
| 15. Blames others for problem behaviors | 7.0 | 1.97 | 1.31 | .72 |
| 16. Has difficulty entering groups | 6.7 | 2.17 | 1.22 | .49 |
| 17. Is easily distracted | 13.1 | 2.79 | 1.44 | .68 |
| 18. Has difficulty accepting criticism or correction | 10.7 | 2.61 | 1.40 | .75 |
| 19. Fails to finish tasks or projects | 8.4 | 2.41 | 1.27 | .64 |
| 20. Sasses teacher | 6.6 | 1.77 | 1.17 | .66 |
| 21. Verbally fights with other students | 9.0 | 2.70 | 1.27 | .70 |
| 22. Whines | 6.6 | 2.42 | 1.25 | .61 |
| 23. Is overactive or restless | 8.4 | 2.10 | 1.35 | .74 |
| 24. Physically fights with other students | 6.4 | 1.77 | 1.11 | .68 |
| 25. Makes noise in class | 10.2 | 2.15 | 1.27 | .78 |
| 26. Acts defiant when told to do something | 8.0 | 2.08 | 1.24 | .73 |
| 27. Argues with teachers about rules or instructions | 6.8 | 2.00 | 1.25 | .73 |
| 28. Interrupts other students | 7.4 | 2.44 | 1.19 | .73 |
| 29. Is noisy | 9.4 | 2.35 | 1.25 | .76 |
| 30. Has trouble awaiting turn | 8.1 | 2.33 | 1.34 | .79 |
| 31. Talks excessively | 7.4 | 2.28 | 1.29 | .64 |

(*Continued.*)

Table 1   (*Continued.*)

| Item | Teacher reporting item as a problem % | Frequency rating | | Corrected item–total correlation |
| --- | --- | --- | --- | --- |
| | | M | SD | |
| 32. Loses things needed for school activities | 4.8 | 1.86 | 1.01 | .57 |
| 33. Fidgets or squirms in seat | 10.1 | 2.41 | 1.43 | .77 |
| 34. Fails to listen to instructions | 16.7 | 2.74 | 1.34 | .78 |
| 35. Is touchy or easily annoyed | 6.7 | 2.28 | 1.28 | .63 |
| 36. Bothers others on purpose | 9.9 | 2.01 | 1.23 | .73 |
| 37. Has trouble paying attention | 12.5 | 2.55 | 1.35 | .70 |
| 38. Has difficulty staying seated | 9.3 | 2.29 | 1.32 | .77 |

*Note.* SESBI-R = The Sutter–Eyberg Student Behavior Inventory-Revised; $N = 983$.

total correlation of .65. Further, test-retest correlation was found to be .87 and .93 for the Intensity and Problem scale, respectively (Eyberg & Pincus, 1999). The SESBI-R scores were found to have a high correlation with externalizing scores on the TRF. It was also found to show good predictive validity in terms of significant correlations between SESBI-R scores and number of school suspensions and number of referrals for conduct problems one and two years later (Eyberg & Pincus, 1999).

Before the present study began, the original version of the SESBI-R was translated into Norwegian by the first author, and then back-translated into English by a bilingual scholar. The originator then approved the translation.

### Ethics

The study was approved by the Regional Committee for Medical Research ethics at the University of Tromsø.

### Analyses

Following the instructions in the manual, missing answers were replaced with "never" (1) on the Intensity scale and "not a problem" (0) on the Problem scale. Responses with more than four missing items on either scale were considered invalid and could not be scored (Eyberg & Pincus, 1999). In this sample, four missing items on each of the Intensity and Problem scales were replaced (0.4%), and three responses with missing answers on the Problem scale were removed from the analysis (0.3%).

Descriptive statistics (means, standard deviations, and percentages) were calculated for total scores on both scales. Internal consistency of the SESBI-R total Intensity and Problem scales was established with Cronbach's alpha. Rayfield et al. (1998) did an exploratory factor analysis on SESBI-R, and concluded with a two-factor solution. Their factors were: Oppositional, consisting of 30 items, and Attentional, consisting of eight items. As the first factor was quite general encompassing both oppositional, conduct, and hyperactive behavior, we wanted to explore whether in fact the Norwegian data can give a different factor solution.

Therefore, in the present study we used an exploratory procedure instead of trying to confirm Rayfield's proposed model. A PCA was chosen in this study to estimate the explained variance of the factors and was performed using PASW 17 (SPSS) to explore the factor structure of the SESBI-R. An oblique rotation with delta set to zero, was performed because it allows factors to be correlated (Tabachnick & Fidell, 2007). A Parallel Analysis was used to best determine what number of components to extract from the PCA (Watkins, 2006), using The Monte Carlo PCA for Parallel Analysis program with 100 replications.

To investigate if there was any dependency in teacher ratings of more than six children per class, the intraclass correlation (ICC) was computed using multilevel analysis (Singer & Willett, 2003). The ICC was found to be 0.19, indicating no need for multilevel analysis (Muthen & Satorra, 1995). Two-way ANOVAs were conducted to compare group means of the SESBI-R total Intensity and Problem scores by sex and age group and to evaluate possible interaction effects. Because the scores on the SESBI-R total Intensity and Problem scales were found not to be normally distributed, log-transformed data was used in the analysis.

Effect sizes (ES) were estimated by means of partial eta squared and interpreted according to Cohen's criteria for the percentage of variance accounted for: small effect $= 1\%$, medium effect $= 5.9\%$, and above $13.8\%$, a large effect (Cohen, 1988). An alpha level of $p < .05$ indicated statistical significance.

## Results

### Descriptive Statistics and Internal Consistency

The frequency distribution of the SESBI-R total Intensity score is presented in Figure 2. The mean Intensity score for the total sample was 89.4 (SD $= 34.5$). The mean Problem score for the total sample was 3.5 (SD $= 6.9$). Table 2 shows means, standard deviation, and percentile scores for the SESBI-R Intensity and Problem scales by age group and sex.
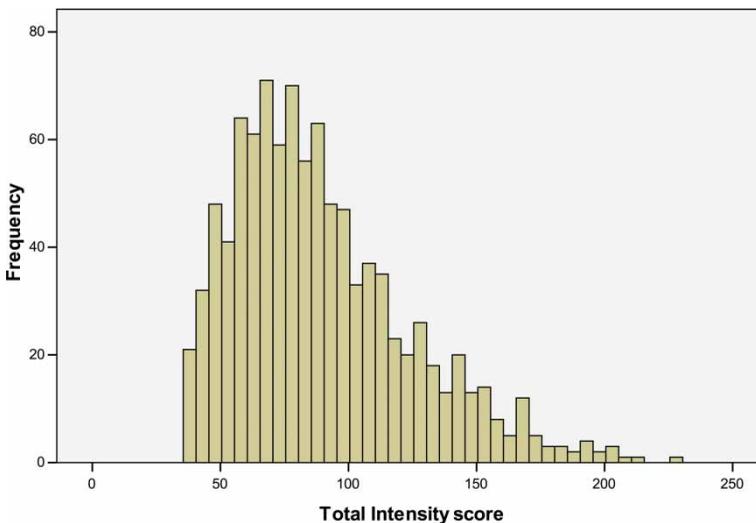


*Figure 2.*    Frequency distribution of the Sutter–Eyberg Student Behavior Inventory-Revised (SESBI-R) Intensity score ($N = 983$).

Table 2

*Means, SD and Percentile Scores for the SESBI-R Intensity and Problem Scales*

| Scale | Age group | Child sex | N | M | (SD) | 90th percentile | 95th percentile | 98th percentile |
|---|---|---|---|---|---|---|---|---|
| Intensity | 3–5 years | Girls | 279 | 87.7 | (27.5) | 124.0 | 143.0 | 159.2 |
| Scale | | Boys | 276 | 99.9 | (31.9) | 145.6 | 156.3 | 169.9 |
| | | Total | 555 | 93.7 | (30.4) | 138.0 | 152.0 | 167.8 |
| | 6–8 years | Girls | 216 | 69.3 | (27.1) | 108.3 | 126.2 | 145.6 |
| | | Boys | 212 | 98.7 | (42.8) | 167.4 | 185.1 | 204.5 |
| | | Total | 428 | 83.8 | (38.6) | 144.0 | 169.6 | 192.0 |
| | Total | | 983 | 89.4 | (34.5) | 140.0 | 156.8 | 177.3 |
| Problem | 3–5 years | Girls | 278 | 2.0 | (4.6) | 7.0 | 11.1 | 19.4 |
| Scale | | Boys | 274 | 3.1 | (5.6) | 11.0 | 17.3 | 23.5 |
| | | Total | 552 | 2.6 | (5.3) | 9.7 | 14.0 | 21.9 |
| | 6–8 years | Girls | 216 | 2.1 | (4.7) | 7.0 | 14.2 | 20.0 |
| | | Boys | 212 | 7.3 | (10.2) | 23.7 | 30.0 | 34.0 |
| | | Total | 428 | 4.6 | (8.3) | 19.0 | 25.0 | 30.0 |
| | Total | | 980 | 3.5 | (6.9) | 12.0 | 20.0 | 28.0 |

*Note.* SESBI-R = The Sutter–Eyberg Student Behavior Inventory-Revised.

For all 983 children in the sample, the mean frequency of occurrence ratings for the 38 items of the SESBI-R Intensity scores ranged from 1.74 to 3.49 (Table 1). The standard deviations of the frequency ratings ranged from 1.10 to 1.48 (Table 1). The items were reported as problems by between 4.8% ("Loses things needed for school activities") and 16.7% ("Fails to listen to instructions") of the teachers (Table 1).

Cronbach's alpha for the SESBI-R Intensity scale was in the present study .97, with an average item-to-total correlation of .69. The lowest item-to-total correlation was for "Has difficulty entering groups" (.49), and the highest was for "Has trouble waiting turn" (.79). Scores on the SESBI-R Intensity scale ranged from 38 to 227.

For the Problem scale, Cronbach's alpha was found to be .96, with an average item-to-total correlation of .49. The lowest corrected item-to-total correlation was for "Acts bossy with other students" (.33), and the highest was for "Refuses to obey until threatened with punishment" (.64). Scores on the SESBI-R Problem scale ranged from 0 to 38.

Both scales were positively skewed. The Intensity scale had a Skewness coefficient of .94 and a coefficient of kurtosis of .68. The Problem scale had a Skewness coefficient of 2.52 and a kurtosis coefficient of 6.23. The Pearson correlation between scores from the two scales of the SESBI-R, the Intensity scores and Problem scores, was $r$ (.71).

## Principal Component Analysis

Prior to performing PCA, the suitability of data for factor analysis was assessed. Inspection of the inter-item correlation matrix revealed that most coefficients had values of .3 and above. The Kaiser–Meyer–Oklin value was .97, exceeding the recommended value of .6, and Bartlett's test of Sphericity reached statistical significance, supporting factorability of the correlation matrix.

Table 3

*SESBI-R Pattern and Structure Matrix for PCA with Oblimin Rotation of Two-Factor Solution*

| Item | Factor 1 | | Factor 2 | | Communalities |
|---|---|---|---|---|---|
| | Pattern coefficient | Structure coefficient | Pattern coefficient | Structure coefficient | |
| 1. Has temper tantrums | **.773** | .761 | .021 | −.454 | .579 |
| 2. Pouts | **.775** | .737 | .061 | −.414 | .546 |
| 3. Teases or provokes other students | **.681** | .776 | −.155 | −.573 | .618 |
| 4. Lies | **.708** | .711 | −.005 | −.440 | .506 |
| 5. Acts frustrated with difficult tasks | .310 | .565 | **−.416** | −.606 | .427 |
| 6. Does not obey school rules on his/her own | **.466** | .661 | −.318 | −.604 | .500 |
| 7. Demands teacher attention | .320 | .523 | −.332 | −.528 | .343 |
| 8. Dawdles in obeying rules or instructions | .319 | .658 | **−.553** | −.749 | .624 |
| 9. Acts bossy with other students | **.787** | .680 | .174 | −.309 | .481 |
| 10. Gets angry when doesn't get his/her own way | **.820** | .777 | .070 | −.433 | .606 |
| 11. Interrupts teacher | **.461** | .658 | −.320 | −.604 | .497 |
| 12. Impulsive, acts before thinking | .393 | .688 | **−.480** | −.722 | .617 |
| 13. Refuses to obey until threatened with punishment | **.669** | .747 | −.126 | −.537 | .567 |
| 14. Has difficulty staying on task | −.139 | .432 | **−.930** | −.845 | .726 |
| 15. Blames others for problem behaviors | **.746** | .763 | −.027 | −.485 | .582 |
| 16. Has difficulty entering groups | .300 | .444 | −.234 | −.419 | .232 |
| 17. Is easily distracted | −.133 | .441 | **−.935** | −.853 | .739 |
| 18. Has difficulty accepting criticism or correction | **.757** | .786 | −.047 | −.512 | .619 |
| 19. Fails to finish tasks or projects | −.085 | .424 | **−.829** | −.777 | .608 |
| 20. Sasses teacher | **.751** | .729 | .035 | −.426 | .533 |
| 21. Verbally fights with other students | **.768** | .759 | .015 | −.456 | .576 |
| 22. Whines | **.631** | .640 | −.015 | −.402 | .410 |
| 23. Is overactive or restless | .167 | .582 | **−.676** | −.779 | .624 |
| 24. Physically fights with other students | **.631** | .703 | −.117 | −.505 | .503 |
| 25. Makes noise in class | .413 | .699 | **−.465** | −.719 | .624 |

(*Continued.*)

Table 3   (*Continued.*)

| Item | Factor 1 | | Factor 2 | | Communalities |
|---|---|---|---|---|---|
| | Pattern coefficient | Structure coefficient | Pattern coefficient | Structure coefficient | |
| 26. Acts defiant when told to do something | **.743** | .775 | −.052 | −.508 | .602 |
| 27. Argues with teachers about rules or instructions | **.824** | .798 | .042 | −.464 | .639 |
| 28. Interrupts other students | **.479** | .684 | −.333 | −.627 | .536 |
| 29. Is noisy | .373 | .675 | **−.491** | −.721 | .606 |
| 30. Has trouble awaiting turn | .361 | .684 | **−.526** | −.748 | .640 |
| 31. Talks excessively | .346 | .568 | **−.363** | −.575 | .405 |
| 32. Loses things needed for school activities | −.002 | .407 | **−.667** | −.665 | .443 |
| 33. Fidgets or squirms in seat | .085 | .577 | **−.801** | −.853 | .733 |
| 34. Fails to listen to instructions | .238 | .639 | **−.653** | −.799 | .674 |
| 35. Is touchy or easily annoyed | **.741** | .698 | .071 | −.384 | .490 |
| 36. Bothers others on purpose | **.659** | .747 | −.144 | −.548 | .571 |
| 37. Has trouble paying attention | −.123 | .458 | **−.947** | −.871 | .768 |
| 38. Has difficulty staying seated | .132 | .593 | **−.751** | −.832 | .704 |

*Note.* SESBI-R = The Sutter–Eyberg Student Behavior Inventory-Revised. Bolded items indicate major loadings for each item. correlation between the two factors from the oblique rotations was −.614.

The results of PCA revealed four components with eigenvalues exceeding 1, explaining 49, 7.5, 4.8, and 3.9% of the variance, respectively. The first five eigenvalues were 18.6, 2.9, 1.8, and 1.5. The first four factors exceeded the criterion value obtained from parallel analysis of eigenvalues (the first four random eigenvalues were, 1.39, 1.35, 1.32, and 1.28, respectively). The scree plot revealed a clear break after the second component, suggesting that two factors should be retained for rotation.

The oblique rotation (oblimin) yielded the two-factor solution presented in Table 3. The first rotated factor accounted for 49% of the variance and contained 20 items with item–factor correlations greater than .40. These items are related to oppositional behavior (e.g., "Has temper tantrums," "Pouts," "Acts defiant when told to do something"). Cronbach's alpha for the first rotated factor was found to be .96.

The second rotated factor accounted for 7.6% of the variance and contained 15 items with item-factor correlations greater than .40. These 15 items are related to attentional difficulties (e.g., "Has difficulty staying on task," "Has trouble paying attention," "Fails to finish tasks or projects"). Cronbach's alpha for the second rotated factor was found to be .95.

### Intensity and Problem Scale Scores by Children's Sex and Age

The results of ANOVAs revealed a significant main effect for sex on the SESBI-R Intensity scale in that boys received higher scores than girls, $F (1, 979) = 103.5$, $p < .001$, ES = 9.6%. A significant main effect was also found for age group (3–5 years vs. 6–8 years). Preschoolers received scores at a significantly higher level than school-aged children, $F (1, 979) = 46.3$, $p < .001$, ES = 4.5%.

The results also showed a significant sex and age interaction effect, $F (1, 979) = 20.9$, $p < .001$, ES = 2.1%. Girls aged 6–8 years received lower scores on the Intensity scale than those aged 3–5 years, while boys showed the same Intensity scale scores across the two age groups.

On the SESBI-R Problem scale a significant main effect for sex was obtained, with boys receiving higher scores than girls, $F (1, 386) = 37.0$, $p < .001$, ES = 8.8%. No significant main effect for age group was found. However, a significant interaction effect was found between sex and age group, $F (1, 386) = 5.8$, $p < .05$, ES = 1.5%. Boys aged 6–8 years received higher scores on the Problem scale than those aged 3–5 years, while girls received stable and low scores across the two age groups on the Problem scale.

### Discussion

In the present study, psychometric properties, factor structure, and differences by age and sex on the SESBI-R (teacher report) were investigated in a sample of 983 Norwegian children aged 3–8 years.

Overall, the mean scores on both the Intensity and Problem scales indicate that Norwegian teachers evaluate their students as having fewer and less severe behavior problems than do teachers in the US (Eyberg & Pincus, 1999). These findings are also in line with previous teacher reports from Scandinavian countries (Ivanova et al., 2007; Larsson & Drugli, 2011), and may indicate that Norwegian children demonstrate fewer behavior problems than children from the USA. However, these differences may also relate to normative differences between the countries. Different age groups in the samples may also have influenced the results, in that the children in the US sample were older. These differences indicate the

need of establishing national norms for the SESBI-R when it is used both of clinicians' and researchers outside the USA.

Good internal consistency was found for both the Intensity and Problem scales of the SESBI-R in the Norwegian sample, indicating a high degree of homogeneity in the scales. Other studies have found similarly high internal consistency for both scales (Querido & Eyberg, 2003; Rayfield et al., 1998). A high correlation between the SESBI-R Intensity scores and Problem scores indicates that the frequency of problem behavior in student and teachers' reporting of behaviors as problems are related dimensions. However, a shared variance of 50% also shows that the two scales measure different aspects of behavior problems. Although the Intensity and Problem scores correlate highly, they may also provide some unique information. Both scales should therefore be used when behavior problems are assessed in a day-care unit or school. Querido and Eyberg (2003) argue that discrepancies between a low Intensity score and a high Problem score may indicate that the teacher is frustrated or angry with the student, or that the teacher may be intolerant of misbehavior in general. A high Intensity score and a low Problem score might, on the other hand, occur when the teacher is unresponsive to the student's problems, or the teacher may be defensive about being able to manage classroom behavior.

Both scales were positively skewed, which implies that few children in this sample were rated with high scores on either scale. The lowest corrected item-to-total correlation was found for "Has difficulty entering groups" compared with the US sample "Acts frustrated with difficult tasks," and the highest for this sample was "Has trouble waiting turn," while for the US sample it was "Refuses to obey until threatened with punishment". For the Problem scale the lowest corrected item-to-total correlation for this sample was for "Acts bossy with other students," and the highest corrected item-to-total correlation was for "Refuses to obey until threatened with punishment," comparable to the US sample the lowest corrected item-to-total correlation was found for "Has difficulty entering groups," and the highest was "Impulsive, acts before thinking". The differences in which behaviors the teachers find troublesome, and to what extent, may be because of differences in culture and in the structure of the day-care units and schools; moreover, the difference in age of the groups may have had an effect. This study used younger children than the US study.

The results of the PCA revealed four components, although the SESBI-R is a general measure of behavior problems. However, the scree plot suggested that the SESBI-R has two separate, measurable factors which represent a valid structure with clinical relevance, and therefore a two-factor solution was retained. The first factor reflects oppositional behavior and the second attentional difficulties. These findings are comparable to those of the US PCA analysis of the SESBI-R (Rayfield et al., 1998). Although the present sample revealed four components and the US analysis revealed three, both analyses show two distinct factors. The results of the present study show that the SESBI-R is a reliable measure of behavior problems in a Norwegian sample of boys and girls in the two age groups of 3–5 and 6–8 years. Our findings support the applicability of the SESBI-R to countries outside the US; however, this should be further investigated. Rayfield et al. (1998) argue that the SESBI-R may prove useful in diagnostic decisions, because the factors could be scored separately and used to decide what position attentional difficulties play in a child's overall behavior problems in the classroom. Further alteration of the SESBI-R should include a confirmatory factor analysis using a new sample to evaluate the present model.

A significant main effect for sex was found with a medium effect size, boys receiving higher Intensity scores than girls in both age groups. This result accords with previous

research using teacher assessment of child behavior (Broidy et al., 2003; Eyberg & Pincus, 1999; Miner & Clarke-Stewart, 2008). A significant main effect was found in regard to the different age groups in that preschoolers received higher Intensity scores than the school-aged children. This result supports the findings of Miner and Clarke-Stewart (2008), but differs from other research using teacher assessments of children (Querido & Eyberg, 2003). However, the small effect size found in the Norwegian sample indicates that the difference may not be clinically significant. On the other hand, this finding is in line with the manner in which parents in Norway assess their children's problem behavior (Reedtz et al., 2008). The interaction effect shows that behavior problems of girls decrease after they begin school, while boys' scores increase. A reason for this may be that the school environment contributes to reduced behavior problems amongst Norwegian girls, or it may be easier for girls than boys to adapt to the school system. No such significant interaction effect was found for sex and age in the US sample (Querido & Eyberg, 2003).

Boys receive higher Problem scores than girls, a finding with medium effect size. This result is in line with previous research on the SESBI-R (Eyberg & Pincus, 1999; Rayfield et al., 1998). Further, a small sex-by-age interaction effect was found, older boys receiving higher scores than younger boys, while girls received stable and low Problem scores across the two age groups. Teachers in schools seem to find the boys' problem behavior more difficult to manage than day-care unit teachers, even if teachers in both contexts rate boys' problem behavior at the same level. Boys' negative behavior in day-care units is experienced differently than negative behavior in schools. The reason for this might be because teachers in schools find boys' negative behavior to be more disturbing than day-care unit teachers do. The school context may be more demanding for young boys than for girls of the same age. This finding needs to be discussed with teachers in Norwegian schools.

One limitation of the present study is the relatively low response rate (60 and 57%), which restricts the generalization of our findings to the general population. The present study was conducted in the spring semester and only one reminder was allowed. It is likely that with the use of more than one reminder, the response rate would be higher. However, the school and day-care unit districts were equally represented from urban and rural areas within the cities of Trondheim and Tromsø, and there was found no significant difference between the responding and non-responding day-care units and schools in Trondheim and Tromsø in regard to school or day-care unit centrality. Another limitation is the lack of a representative sample; however, in Larsson and Drugli's (2011) national sample they found no difference in regard to teachers' assessments and geographical location. We can therefore assume that Trondheim and Tromsø do not diverge significantly from other cities in Norway. Another limitation of the study is the lack of information about the non-responding participants. It is possible that some of the non-responding children were experiencing problems, and therefore would have raised the scores for both genders. Strengths of the present study are the use of a standardized assessment tool within a Norwegian sample across day-care units and schools, the relatively large sample size, and that this study is the only one that has been conducted according applicability of the SESBI-R outside the US.

## Conclusion and Implication for Practitioners

Overall, our findings show that teachers in Norway report relatively low levels of behavior problems among day-care unit and school-aged children, and that Norwegian teachers evaluate their students as having fewer and less severe behavior problems than their US

counterparts. Further, this study indicates that the SESBI-R is a reliable assessment tool that is applicable to the Norwegian population. The study shows that, like the US measure, the SESBI-R has two factors: oppositional behavior and attentional difficulties. This study also shows that boys receive higher Intensity and Problem scores than girls, and that pre-schoolers receive higher Intensity scores than school-aged children.

In future research among Norwegian children in day-care and school, psychometric work needs to be conducted on test-retest reliability, and more evidence about predictive and con-current validity should be gathered. This study is just a beginning, and its results should be examined further in a more representative sample to estimate Norwegian norms on the SESBI-R.

The Sutter–Eyberg Student Behavior Inventory–Revised (SESBI-R) is a straightforward questionnaire that can be used in day-care units and schools to evaluate behavior problems and the extent to which teachers find the behavior troublesome. The SESBI-R can also be used as an efficient screening instrument to evaluate the effect of treatments for external pro-blems. The SESBI-R provides professionals with a reliable and valid instrument for efficient screening of the behaviors of conduct-disordered children.

## References

Achenbach, T.M., & Rescorla, L. (2001). *Manual for the ASEBA school-age forms and profiles: Child behavior checklist for ages 6–18, teacher's report form, youth self-report: An integrated system of multi-informant assessment*. Burlington, VT: ASEBA.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev). Washington, DC: Author.

Broidy, L.M., Nagin, D.S., Tremblay, R.E., Bates, J.E., Brame, B., Dodge, K.A., & et al. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology*, *39*(2), 222–245.

Campbell, S.B., Spieker, S., Burchinal, M., Poe, M.D., & NICHD (2006). Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. *Journal of Child Psychology and Psychiatry*, *47*(8), 791–800.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Eyberg, S.M., & Pincus, D. (1999). *Eyberg Child Behavior Inventory (ECBI) & Sutter-Eyberg Student Behavior Inventory-Revised (SESBI-R)*. Lutz, FL: Psychological Assessment Resources.

Funderburk, B.W., & Eyberg, S. (1989). Psychometrick characteristics of the Sutter-Eyberg Student Behavior Inventory: A school behavior rating scale for use with preschool children. *Behvioral Assessment*, *11,* 297–313.

Funderburk, B.W., Eyberg, S.M., Rich, B.A., & Behar, L. (2003). Further psychometric evaluation of the Eyberg and Behar rating scales for parents and teachers of preschoolers. *Early Education and Development*, *14*(1), 67–81.

Ivanova, M.Y., Achenbach, T.M., Rescorla, L.A., Dumenci, L., Almqvist, F., Bathiche, M., et al. (2007). Testing the Teacher's Report Form Syndromes in 20 societies. *School Psychology Review*, *36*(3), 468–483.

Larsson, B., & Drugli, M.B. (2011). School compentence and emotional/behavioral problems among Norwegian school children as rated by teachers on the Teacher Report Form. *Scandinavian Journal of Psychology*. Advance online publication. doi:10.1111j.1467-9450.2011.00889x.

Loeber, R., & Farrington, D.P. (2000). Young children who commit crime: Epidemiology, develop-mental origins, risk factors, early interventions, and policy implications. *Developement and Psychopathology*, *12,* 737–762.

Miner, J.L., & Clarke-Stewart, KA. (2008). Trajectories of externalizing behavior from age two to age nine: Relations with gender, temperament, ethnicity, parenting, and rater. *Developmental Psychology*, *44*(3), 771–786.

Muthen, B., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25,* 267–316.

National Child Traumatic Stress Network (2010). *Sutter-Eyberg Student Behavior Inventory-Revised.* Retrieved July 30, 2010, from http://www.nctsnet.org/nctsn_assets/pdfs/measure/SESBI-R.pdf

Odgers, C.L., Moffitt, T.E., Broadbent, J.M., Dickson, N., Hancox, R.J., & Harrington, H., et al. (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology*, *20*(2), 673–716.

Querido, J.G., & Eyberg, S.M. (2003). Psychometric properties of the sutter-eyberg student behavior inventory-revised with preschool children. *Behavior Therapy*, *34*(1), 1–15.

Rayfield, A., Eyberg, S.M., & Foote, R. (1998). Revision of the Sutter-Eyberg Student Behavior Inventory: Teacher ratings of conduct problem behavior. *Educational and Psychological Measurement*, *58,* 88–98.

Reedtz, C., Bertelsen, B., Lurie, J.I.M., Handegard, B.H., Clifford, G., & Morch, W.-T. (2008). Eyberg Child Behavior Inventory (ECBI): Norwegian norms to identify conduct problems in children. *Scandinavian Journal of Psychology*, *49*(1), 31–38.

Rescorla, L., Achenbach, T., Ivanova, M.Y., Dumenci, L., Almqvist, F., Bilenberg, N., et al. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and Behavioral Disorders*, *15*(3), 130–142.

Singer, J.D., & Willett, J.B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford: Oxford University Press.

Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics*. Boston: Pearson/Allyn and Bacon.

Watkins, M.W. (2006). Determining parallel analysis criteria. *Journal of Modern Applied Statistical Methods*, *5*(2), 344–346.

Watt, B.D., Hoyland, M., Best, D., & Dadds, M.R. (2007). Treatment participation among children with conduct problems and the role of telephone reminders. *Journal of Child and Family Studies*, *16*(4), 522–530.