School psychologists rely heavily upon teachers' evaluation of children's classroom behavior as part of their assessment and intervention processes. However, some evidence suggests that teachers' ratings may fluctuate during the academic year, with higher levels being reported later in the year. This study assessed the comparability of different teachers' ratings of children's externalizing behavior from fall and spring intervention cohorts. Teachers of children from divorced families (N=240, ages 9 to 12) completed pretest, posttest, and six-month follow-up ratings. Two separate ANOVAs revealed that changes in mean levels were not significantly different between the fall and spring cohorts for either the pretest/posttest or pretest/follow-up comparison. Ratings obtained from children's current teachers within the first month of school were compared to those completed by the children’s teachers from the previous school year. Pretest levels of externalizing behavior were higher when reported retrospectively by the previous teachers. However, neither the previous teachers' nor the current teachers' fall ratings were significantly different from the spring teachers' pretest ratings. Thus, school psychologists may elect to obtain ratings from either the previous or current teacher early in the fall of a new academic year in the assessment of children's behavior. (Contains 2 tables and 12 references.) (JDM)
TEACHERS' RATINGS OF CHILDREN'S CLASSROOM BEHAVIOR: TIME OF YEAR EFFECTS?

Poster presented at the 108th Annual Meeting of the American Psychological Association (Washington, DC) in August, 2000 by:

Heather V. A. Hoover, M.A.
Sanford L. Braver, Ph.D.
Sharlene A. Wolchik, Ph.D.
& Irwin N. Sandler, Ph.D.
Arizona State University
Department of Psychology

For more information, please contact:
Heather V. A. Hoover, M.A.
Arizona State University
Clinical Psychology Ph.D. Program
Department of Psychology
PO Box 871104
Tempe, AZ 85287-1104
Phone (480) 965-3326
Fax (480) 965-8544
Email: heather.hoover@asu.edu
Abstract

Teachers' reports may be affected by measurement and interpretation difficulties associated with the potential differences between ratings completed during the fall and spring, particularly if different teachers provide the ratings. This study assessed the comparability of teachers' ratings of children's externalizing behavior from fall and spring intervention cohorts. Teachers of 240 children of divorce aged 9 to 12 years completed pretest, posttest, and six-month follow-up ratings. Children were randomly assigned to one of three treatment conditions (groups for their mothers only, separate groups for mothers and children, or a bibliotherapy control condition). Two separate 3 X 2 X 2 ANOVAs revealed that teachers' ratings were significantly different between the fall and spring follow-up ratings. This study assessed the comparability of ratings obtained from different teachers who provided the ratings. Teachers' reports may be affected by measurement and interpretation difficulties associated with the potential differences between the fall and spring assessments.
A related issue involves the use of teachers' reports to evaluate a student's progress in behavioral domains. When behavior ratings are used to monitor a student's response to an intervention, teachers may be asked to report on the student's behavior in the fall and again in the spring of an academic year. If the intervention period spans more than one academic year, different teachers are likely to provide ratings of the same student's behavior at different time points. The validity of the ratings may be compromised by inter-rater unreliability. On the other hand, when the same teacher provides pre-intervention and post-intervention ratings, the ratings may be less sensitive to detecting change because the teacher's perceptions of the child's behavior are resistant to change (Bryk & Raudenbush, 1992; Raudenbush, 1984). Although assessments are intended to obtain reasonably objective measures of different assessment points, teachers' reports may be particularly affected by measurement and interpretation difficulties associated with the potential differences between fall and spring behavior ratings.

Among the issues that may lead to systematic differences between fall and spring ratings are (1) the length of time a child has been in a particular class with a particular teacher, (2) whether the ratings are completed by the same or different teachers at different assessment points, and (3) whether the sample of observed behavior changes in a different setting at different assessment points. Teachers' reports may be particularly affected by measurement and interpretation differences associated with the potential differences between fall and spring ratings.

The study assessed the comparability of teachers' ratings of children's externalizing behavior from fall and spring intervention cohorts. This study addressed the comparability of semester-specific effects as they affect the evaluation of interventions for school-age children. This study involved a multiple step beginning with the selection of a sample of families (New Beginnings; Wolchik, Sandler, West, & Anderson, 1997). The families were recruited to participate in the study via a multiple step school-based recruitment process. The school on which the child was attending at the time of the pretest was the only school on which the child was attending when the post and follow-up ratings were completed.

The child or children who completed behavior rating scales were part of a preventive intervention study of divorced families. The children on whom the teachers completed behavior ratings were part of a preventive intervention study of divorced families. The children on whom the teachers completed behavior ratings were part of a preventive intervention study of divorced families. The children on whom the teachers completed behavior ratings were part of a preventive intervention study of divorced families. The children on whom the teachers completed behavior ratings were part of a preventive intervention study of divorced families.

Participants
Teachers of 240 children aged nine to twelve years completed pretest, posttest, and six-month follow-up ratings for four different cohorts. The intervention group included two fall and two spring cohorts over a two-year period. The intervention was designed to address the emotional and behavioral needs of children aged nine to twelve years. The intervention was provided in a school-based setting, and the children received individual and group interventions. The intervention consisted of a series of workshops and small group sessions. The intervention was provided by trained teachers and social workers.

Eligibility criteria:
1. The divorce decree was granted within the past two years.
2. The mother was the custodial/primary residential parent.
3. The child had not been involved with a mental health provider in the past two years.
4. The child had not been involved with a mental health provider in the past two years.
5. The child had not been involved with a mental health provider in the past two years.
6. The child had not been involved with a mental health provider in the past two years.
7. The child had not been involved with a mental health provider in the past two years.
8. The child had not been involved with a mental health provider in the past two years.
9. The child had not been involved with a mental health provider in the past two years.
10. The child had not been involved with a mental health provider in the past two years.
Method

Measure

Teacher-Child Rating Scale (T-CRS; Hightower et al., 1986; Primary Mental Health Project, 1995). The Teacher-Child Rating Scale includes a six-item Acting Out subscale that taps externalizing behavior problems. The T-CRS has demonstrated adequate reliability as well as concurrent and discriminant validity to differentiate referred from non-referred children (Hightower et al., 1986).

In the present study, the internal consistency of the Acting Out subscale was excellent. The reliability coefficients (Cronbach's alpha) were calculated on each subject (i.e., pretest, posttest, and follow-up) for the repeated-measure ANOVA with one repeated measure (within-subjects) factor.

Data Collection

In the current investigation, the pretest assessments were administered on each subject. The pretest was completed in early September and followed by the posttest in May and follow-up in November. For all cohorts, the ratings were collected at the end of each semester-long group. The ratings for the spring cohorts were collected as follows: pretest in late February; posttest in May; and follow-up in November. For the fall cohorts, the pretest data were collected from both the previous and current teachers of each child in the study.

Treatment Conditions

For each of the comparisons between fall and spring cohorts, correlation coefficients were computed. Mean externalizing scores were examined to assess the substantive question regarding differences between cohorts.

Control was a random, between-subjects factor. The experimental condition was a two-by-two factorial design with two between-subjects factors: (a) pretest/posttest/follow-up, and (b) cohort in which a child was referred (i.e., referred or non-referred). The repeated measures factor was the time each rating was collected on each subject (i.e., pretest, posttest, and follow-up).

The repeated measures factor was the time each rating was collected on each subject (i.e., pretest, posttest, and follow-up). This design was replicated as another 3 X 2 ANOVA for the pretest/follow-up scenario. This design was replicated in a 3 X 2 ANOVA with one repeated measure (within-subjects) factor (self-report/registration) for the T-CRS acting out subscale.

Design

For each of the comparisons between fall and spring cohorts, correlation coefficients were computed. Mean externalizing scores were examined to assess the substantive question regarding differences between cohorts.
would be obtained with a different sample. A control condition should be a separate and independent treatment with making appropriate comparisons, it is unknown whether the same results could be obtained. First, the instrument used in this study, the Teacher-Child Rating Scale (T-CRS, Hightower et al., 1998), lacks normative data and is not widely used. A second consideration is the presence of sensitivity to detecting treatment effects. Also, the generalization of the findings should be approached with caution for a number of reasons. Results: Two separate 3 x 2 x 2 ANOVAs revealed the changes in mean levels were not significantly different between the fall and spring cohorts for either the pretest/posttest or pretest/follow-up comparison. When different teachers completed the ratings, the pretest/posttest and pretest/follow-up correlations demonstrated stability over nine months (\( r = .55 \) to .79). (In a meta-analysis of thirteen studies that used two teachers as informants on children's behavior, Achenbach, McConaughy, and Howell (1987) found a mean \( r = .64 \) between ratings completed by different raters over a period of nine months. The correlations between ratings over time were remarkably high, even when different teachers provided behavior ratings over a period of nine months. The correlations between ratings were stable.)

Conclusions: Regardless of the underlying reasons, it is clear that teachers' ratings of children's externalizing behavior are quite consistent, even when different teachers provide behavior ratings over a period of nine months. The correlations between ratings over time were remarkably high, even when different teachers provided behavior ratings over a period of nine months. The correlations between ratings were stable.

Limitations of this study include the lack of information about whether the stability of teachers' ratings is actually problematic with respect to detecting treatment effects. Also, the generalization of the findings should be approached with caution for a number of reasons. Results: Two separate 3 x 2 x 2 ANOVAs revealed the changes in mean levels were not significantly different between the fall and spring cohorts for either the pretest/posttest or pretest/follow-up comparison. When different teachers completed the ratings, the pretest/posttest and pretest/follow-up correlations demonstrated stability over nine months (\( r = .55 \) to .79). (In a meta-analysis of thirteen studies that used two teachers as informants on children's behavior, Achenbach, McConaughy, and Howell (1987) found a mean \( r = .64 \) between ratings completed by different raters over a period of nine months. The correlations between ratings over time were remarkably high, even when different teachers provided behavior ratings over a period of nine months. The correlations between ratings were stable.)

Conclusions: Regardless of the underlying reasons, it is clear that teachers' ratings of children's externalizing behavior are quite consistent, even when different teachers provide behavior ratings over a period of nine months. The correlations between ratings over time were remarkably high, even when different teachers provided behavior ratings over a period of nine months. The correlations between ratings were stable.

Limitations of this study include the lack of information about whether the stability of teachers' ratings is actually problematic with respect to detecting treatment effects. Also, the generalization of the findings should be approached with caution for a number of reasons. Results: Two separate 3 x 2 x 2 ANOVAs revealed the changes in mean levels were not significantly different between the fall and spring cohorts for either the pretest/posttest or pretest/follow-up comparison. When different teachers completed the ratings, the pretest/posttest and pretest/follow-up correlations demonstrated stability over nine months (\( r = .55 \) to .79). (In a meta-analysis of thirteen studies that used two teachers as informants on children's behavior, Achenbach, McConaughy, and Howell (1987) found a mean \( r = .64 \) between ratings completed by different raters over a period of nine months. The correlations between ratings over time were remarkably high, even when different teachers provided behavior ratings over a period of nine months. The correlations between ratings were stable.)

Conclusions: Regardless of the underlying reasons, it is clear that teachers' ratings of children's externalizing behavior are quite consistent, even when different teachers provide behavior ratings over a period of nine months. The correlations between ratings over time were remarkably high, even when different teachers provided behavior ratings over a period of nine months. The correlations between ratings were stable.

Limitations of this study include the lack of information about whether the stability of teachers' ratings is actually problematic with respect to detecting treatment effects. Also, the generalization of the findings should be approached with caution for a number of reasons. Results: Two separate 3 x 2 x 2 ANOVAs revealed the changes in mean levels were not significantly different between the fall and spring cohorts for either the pretest/posttest or pretest/follow-up comparison. When different teachers completed the ratings, the pretest/posttest and pretest/follow-up correlations demonstrated stability over nine months (\( r = .55 \) to .79). (In a meta-analysis of thirteen studies that used two teachers as informants on children's behavior, Achenbach, McConaughy, and Howell (1987) found a mean \( r = .64 \) between ratings completed by different raters over a period of nine months. The correlations between ratings over time were remarkably high, even when different teachers provided behavior ratings over a period of nine months. The correlations between ratings were stable.)

Conclusions: Regardless of the underlying reasons, it is clear that teachers' ratings of children's externalizing behavior are quite consistent, even when different teachers provide behavior ratings over a period of nine months. The correlations between ratings over time were remarkably high, even when different teachers provided behavior ratings over a period of nine months. The correlations between ratings were stable.
Future research should examine teachers' ratings of internalizing behaviors and verify the validity of teachers' ratings compared to behavioral observations. Further research could elucidate the mechanisms through which teachers' ratings become stable and determine if teachers and classroom environments themselves become contexts that promote stability in the way children are perceived.

Table 1

<table>
<thead>
<tr>
<th>Design of the Study</th>
<th>Fall cohorts</th>
<th>Pre</th>
<th>Post</th>
<th>FU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1B &amp; T2A</td>
<td>81</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T2B &amp; T2F</td>
<td>79</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T3B &amp; T3A</td>
<td>75</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T4F</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. T = Teacher, and is followed by a number assigned to differentiate whether the teacher is the same or different teacher as the one who completed a rating at another time point. The letters after the numbers indicate when the rating took place: B = Before, A = After, and F = Follow-up.

Table 2

<table>
<thead>
<tr>
<th>Partial Correlations</th>
<th>Pre/Post Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre/Post</td>
<td>1.13</td>
</tr>
<tr>
<td>Pre/Post</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Note. Coefficients in the same column sharing the same superscript are not significantly different from one another. * p < .01; all other values with different superscript letters are significantly different at p < .05.

Table 2

<table>
<thead>
<tr>
<th>Partial Correlations</th>
<th>Pre/Post Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre/Post</td>
<td>1.13</td>
</tr>
<tr>
<td>Pre/Post</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Note. Fall current indicates the pretest data came from each child's current teacher at the time of the rating. Fall previous indicates that the pretest rating was obtained from each child's teacher from the previous school year.
REFERENCES


I. DOCUMENT IDENTIFICATION:

Title: TEACHERS' RATINGS OF CHILDREN'S CLASSROOM BEHAVIOR: TIME OF YEAR EFFECTS?

Author(s): Heather V.A. Hoover, Sanford L. Braiker, Sharlene A. Wolchik, & Irwin N. Sandler

Corporate Source: Publication Date: August 2000

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample stickers shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

The sample stickers shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2A

The sample stickers shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2B

Level 1

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

Level 2A

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only.

Level 2B

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Heather V.A. Hoover
Printed Name/Position/Tite: Heather V.A. Hoover / Ph.D. Candidate
Organization/Address: Arizona State University Psychology Dept. PO Box 871104 Tempe, AZ 85287
Phone: 480 965 2326 Fax: 480 965 8544 E-mail Address: heather.hoover@asu.edu Date: 11/15/01

APA 2000
III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

<table>
<thead>
<tr>
<th>Publisher/Distributor:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Address:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Price:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

<table>
<thead>
<tr>
<th>Name:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Address:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of North Carolina at Greensboro
ERIC/CASS
201 Ferguson Building
PO Box 26171
Greensboro, NC 27402-6171

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)