TEACHER EVALUATION BIAS TOWARD LD CHILDREN – ATTENUATING EFFECTS OF THE CLASSROOM ECOLOGY

Melinda Dukes and Richard A. Saudargas

***Abstract.* The present study investigated whether classroom context cues in videotapes showing a nonhandicapped and an LD expectancy would attenuate teacher evaluation bias. Participants were 80 elementary-school teachers who were randomly divided between a nonhandicapped and an LD expectancy. The teachers completed a rating scale on a child's behavior before and after viewing a videotape of children engaged in individual seatwork or group instruction. Results showed that at pretest the teachers' ratings differed according to the assigned label. After viewing the videotape, the teachers rated the two expectancy groups as not differing significantly. However, the teachers rated the child's behavior during group instruction as less appropriate concurring with the child's actual behavior. The findings suggest that teachers hold initial biases about the behavior of LD and normal children, but take into account the classroom context when rating actual behavior.**

This study investigated the effect of classroom context variables upon teacher bias created by the learning disabled (LD) label. Research has shown that under certain conditions, the LD label biases teachers' evaluations of students (Fogel & Nelson, 1983; Jacobs, 1978; Sutherland & Algozzine, 1979; Taylor, Smiley, & Ziegler, 1983). However, under other conditions special education labels have been found not to bias teacher evaluation of students (Yoshida & Meyers, 1975).

Foster and Ysseldyke (1976) found that teacher evaluations of students were negatively influenced by special education labels. This bias emerged in a study where teachers viewed a videotape of a normal child who was engaged in assessment tasks and play activities. Participants were divided into two groups — one was informed that the child had been assigned a special education label, subjects in the other group were told the child was normal. Results showed that the teachers in the former category evaluated the child more negatively than those in the latter. These findings have been replicated with other subject samples (Fogel & Nelson, 1983; Foster, Ysseldyke, & Reece, 1975) and using different experimental procedures (Algozzine & Sutherland, 1977; Jacobs, 1978).

In the preceding studies videotapes were used showing a nonhandicapped child involved in play or testing activities. None of the tapes included groups of children engaging in academic work or interacting with the teacher as commonly occurs in the classroom. Therefore, since the tapes lacked an adequate portrayal of a classroom setting, generalization of the findings is limited.

Not all research has shown that labels generate negative biases. Instead, some researchers have suggested that expectancy biases may not result from special education labeling (Claiborn, 1969; Dusek & O'Connell, 1973; Gozali & Meyen, 1970; Jose & Cody, 1971; Yoshida & Meyers, 1975). The hypothesis that extended teacher-student

MELINDA DUKES, Ph.D., is Assistant Professor, Dept. of Educational Psychology and Counseling, University of Tennessee.

RICHARD A. SAUDARGAS, Ph.D., is Associate Professor, Dept. of Psychology, University of Tennessee.

contact attenuates label-generated bias was tested by Yoshida and Meyers (1975). In their study, teachers watched a videotape of a child responding to a series of tasks administered by an adult. Tape footage showed child improvement on the task as well as cues of time passage. The teachers, who were informed that the child had been labeled mentally retarded, were not biased in their evaluations of the child. The researchers reasoned that the extended teacher-student contact and student learning portrayed on the videotape allowed participants to form their own evaluations of the child without relying on the child's label.

In summary, the literature on teacher bias within the classroom is inconclusive. Work by Foster and Ysseldyke (1976) and others appears to demonstrate the existence of label-generated bias; however, generalization of these findings to classrooms is limited as they do not incorporate a classroom context into the experimental procedures. In contrast, Yoshida and Meyers (1975), whose videotape included context cues of teacher-child interactions, child improvement on a task, and time passage, did not find a label-generated bias.

To determine that special education labeling biases teacher evaluation of students, the classroom context must be adequately portrayed since classroom events or cues may attenuate the bias-generating potential of special education labels. Therefore, research into this matter should show children and teachers engaged in activities and behaviors commonly found in classroom settings. The current study pursued the question of label-generated bias by including more ecologically valid tapes. The following hypothesis was tested: Classroom context cues included on videotapes will attenuate teacher bias generated by the LD label.

METHOD

Subjects

The subjects were 80 regular classroom, elementary (grades K-5) teachers from eight public schools. Sixty-three percent had Bachelor degrees, 34% held Master's degrees, and the remaining 3% did not indicate their highest degree. The teachers averaged 13 years of regular education teaching experience, and five noted special education teaching. Ninety-one percent reported experience teaching handicapped as well as nonhandicapped students in regular education classrooms.

Independent Variables

Three variables were manipulated in the experiment: (a) test sensitization, (b) expectation, and (c) instructional mode.

Test sensitization. This variable was included due to concern over potential pretest sensitization. Sensitization effects upon the posttest scores were possible because pretest and posttest were separated by only a short length of time (i.e., an 11-minute videotape). To evaluate the impact of the pretests, half of the subjects received the pretest, the other half did not.

Expectation. The teachers were randomly assigned to LD or normal expectancy groups. Expectancy was created by having each teacher read a one-page (150-word) case study which described an eight-year-old boy as being either LD or normal. Except for the key descriptors — LD or normal — the case studies were equivalent. Both groups of teachers watched the same nonhandicapped eight-year-old boy on videotape as described in the section on videotapes below.

Instructional mode. The instructional mode was either individual seatwork or group discussion. This variable was manipulated at two different points in the study. The initial manipulation was accomplished by the pretest instructions, which contained a description of the hypothetical child as participating in a large-group, teacher-led activity or an individual seatwork activity. The second manipulation consisted of presenting one of two videotapes to the teachers. One tape showed the identified child and a group of children engaged in individual seatwork activities; the other showed the same children engaged in a teacher-led group activity. As with test sensitization and expectancy, participants were randomly assigned to this variable.

Dependent Variable

The *Classroom Behavior Inventory* (CBI) (Schaefer, Edgerton, & Aronson, 1977) served as the dependent measure. The CBI is a global behavior characteristic measure (McKinney & Feagans, 1983) comprised of 42 behaviorally oriented child descriptors. Each item is scored on a 5-point numerical rating scale which ranges from *not at all like* to *very much like.* For this study, a sixth response category, *not observed/not appropriate,* was added. This modification was necessary to cover the possibility that the behavior specified in any one item did not occur on the tape.

Each item on the CBI contributes to the score of the 10 scales on the measure. Eight of the scales represent four bipolar dimensions of behavior. These are: Extroversion vs. Introversion, Distractibility vs. Task Orientation, Independence vs. Dependence, and Hostility vs. Consideration. Remaining scales are Verbal Intelligence and Creativity and Curiosity. Psychometric research has supported the CBI's structure, reliability, and validity. The scale reliably differentiates between LD and normal children on task orientation and distractibility (McKinney & Feagans, 1983).

Videotapes

The two videotape segments used in the study depicted a simulated classroom. On both 11-minute segments, six nonhandicapped elementary-school children ($M = 2$, $F = 4$) and a certified special education teacher engaged in a variety of classroom activities. The group instruction videotape showed a teacher-led science discussion. The children were seated in a semicircle while the teacher lectured and asked questions. During the discussion the children engaged in various behaviors including listening to the teacher, responding to questions, talking to each other, and so forth. The individual seatwork videotape showed the children seated in desk rows completing language arts assignments. The teacher monitored the children, gave instructions, and checked their work. Similar to the group instruction tape, the children participated in behaviors typical of a regular classroom including school work, walking about the room, talking with the teacher, and raising hands. For both tapes participants rated the same target boy. (A boy was chosen to be consistent with previous research on teacher labeling biases, which had used males on the videotapes.)

Prior to the experiment, trained observers viewed both tapes using handheld computers programmed with the State Event Classroom Observation System (SECOS) (Saudargas & Fellers, 1986). These observations generated measures of child and teacher behaviors on the tapes. This component of the study served to insure that the behaviors depicted were representative of the kind and frequency of classroom behaviors that typically take place during individual seatwork and group discussion. Observations of the target student's behavior on the videotape showed that during individual seatwork he (a) was engaged in schoolwork-related behavior 71% of the time, (b) looked around the room 16% of the time, (c) engaged in activities other than schoolwork 6% of the time, and (d) talked with other children 1% of the time. On the group instruction videotape, the boy (a) was engaged in schoolwork 45% of the time, (b) looked around 43% of the time, and (c) engaged in activities other than schoolwork 12% of the time.

Procedure

Prior to the experiment the teachers were informed that the purpose of the study was to validate a new teacher rating scale and that the videotapes depicted a summer school program at a university. Teachers in the pretest and posttest experimental condition first completed the CBI with reference to what they considered to be a typical LD or nonhandicapped boy participating in either a large group activity or individual seatwork. After completing and returning the pretest CBI, the teachers received a copy of the case study which matched their assigned expectancy (LD or normal). After having read the case studies silently, participants viewed the assigned videotape whereupon they completed the posttest CBI in reference to what occurred on the tape.

Teachers in the posttest-only condition read the appropriate case study, viewed the assigned videotape, and completed the CBI.

RESULTS

As mentioned, the CBI was modified by adding a sixth response category. The added category, *not observed/not appropriate* (*O* response), served two purposes. First, on the pretests, it gave teachers the opportunity to indicate when an item was not appropriate for describing the hypothetical child; and second, it allowed teachers completing the posttests to respond that a given behavior had not been observed on the videotape.

Prior to any statistical analyses of the CBI, the frequency of the *O* response category was evaluated. Only scales with 90% usable data were included in the analyses. All CBI scales were used for pretest analysis. For posttest data analysis, two scales — Orientation to Task and Distractibility — were used.

Three statistical analyses were conducted on the CBI. The first investigated the impact of the expectancy and instructional mode variables on the pretest CBI scores. The second examined the impact of the same variables, plus test sensitization upon the posttest scores. The last analysis assessed the observed change between pretest and posttest scores.

Table 1
Mean CBI Pretest Scores for LD-and Normal-Expectancy Teachers in the Pretest-Only Group

	Expectancy				
	LD ($N=17$)		Normal ($N=18$)		
Scale	Mean	*SD*	Mean	*SD*	*F*
Creativity and Curiosity	11.01	4.45	17.28	3.77	19.23*
Distractibility	11.94	2.52	8.33	2.53	15.11*
Verbal Intelligence	9.20	4.23	15.67	3.11	22.97*
Orientation to Task	10.39	4.66	15.73	2.96	15.24*
Introversion	9.03	2.24	5.56	1.83	24.59*

*$p<.01$.

All the CBI scales were included in the 2 (Expectancy) X 2 (Instructional Mode) MANOVA since the pretest scores were not characterized by a large number of *O* responses as were the posttest scores. Only the main effect for LD versus normal expectancy condition was significant (F $10,22=5.66$, $p<.001$). Followup univariate analyses revealed that the expectancy effect was significant for five of the 10 scales, that is, Creativity, Curiosity, Distractibility, Verbal Intelligence, Orientation to Task, and Introversion (see Table 1). For the positive oriented scales — Creativity and Curiosity, Verbal Intelligence, and Orientation to Task — the teachers receiving the LD expectancy gave the child lower scores than the normal-expectancy participants. On the Distractibility and Introversion scales, the teachers given the LD expectancy rated the child higher than did teachers given the normal expectancy.

The posttest scores were analyzed with a MANOVA in a 2 (Expectancy) X 2 (Instructional Mode) X 2 (Test Sensitization) analysis. Because of the large number of posttest *O* responses, only the Distractibility and Orientation-to-Task scales were included as dependent variables in the MANOVA. Only the main effect for instructional

Table 2
Posttest Means and Standard Deviations for the CBI Orientation-to-Task and Distractibility Scales

	Large Group ($N=33$)		Individual Seatwork ($N=34$)		
	Mean	*SD*	Mean	*SD*	*F*
Orientation to Task	7.70	3.12	11.26	4.33	14.39**
Distractibility	11.76	2.56	10.37	2.84	4.06*

*$p<.05$. **$p<.01$.

Table 3
Means, Standard Deviations, and t-Values of the CBI Distractibility Scale for Teachers in the Pretest and Posttest Condition

			Distractibility				
Instructional			**Pretest**		**Posttest**		
Mode	**Expectancy**	***N***	**Mean**	***SD***	**Mean**	***SD***	***t***
Large	LD	10	11.40	3.89	12.40	1.84	-1.15
Large	Normal	8	7.25	2.61	11.50	2.20	-5.67**
Indiv.	LD	10	13.10	1.10	9.80	3.01	2.74*
Indiv.	Normal	8	9.25	2.82	11.00	2.51	-1.13

*$p<.05$. **$p<.01$.

mode (F 2,58 = 7.31, $p<.001$) was significant. Followup univariate analyses documented significance for the Orientation-to-Task ($p<.001$) and the Distractibility scale ($p<.048$). Teachers in the large-group instructional mode rated the target child as more distractible and less on task, whereas those in the individual seatwork instructional mode rated the child as being less distractible and more on task (see Table 2).

Although the results of both prior analyses disclosed a significant impact of instructional mode on posttest ratings, the question remained as to whether or not any significant change occurred in the scores from pre- to posttesting. The results of matched-pairs t-tests for the Distractibility and Orientation-to-Task scales showed significant pretest-to-posttest changes on both scales for those teachers who saw the large group film and were given the normal expectancy. These teachers increased their Distractibility scores (t = -5.67,

Table 4
Means, Standard Deviations, and t-Values of the CBI Orientation-to-Task Scale for Teachers in the Pretest and Posttest Condition

			Orientation to Task				
Instructional			**Pretest**		**Posttest**		
Mode	**Expectancy**	***N***	**Mean**	***SD***	**Mean**	***SD***	***t***
Large	LD	6	11.67	7.01	8.83	3.87	1.22
Large	Normal	9	16.67	2.60	7.11	2.26	7.52**
Indiv.	LD	10	10.20	3.01	12.20	4.52	-1.58
Indiv.	Normal	9	15.33	3.74	10.78	4.30	1.89

**$p<.01$.

$p < .001$) (see Table 3) and decreased their Orientation-to-Task scores ($t = 7.52$, $p < .001$) (see Table 4) on the posttests. In contrast, the teachers who saw the individual seatwork film and were given the LD expectancy significantly decreased their posttest scores on the Distractibility scale ($t = 2.74$, $p < .05$) (see Table 3). These teachers also increased their posttest scores on Orientation to Task (see Table 4); however, this change did not reach statistical significance.

DISCUSSION

The current findings support the hypothesis that teachers rely on numerous classroom cues and events to form their evaluations of a child. As illustrated, classroom ecology variables (i.e., academic work, peers, teacher-child interactions, and two instructional modes) functioned to attenuate teacher evaluation bias as did Yoshida and Meyers' (1975) temporal and child improvement variables.

Teachers may be influenced more by a child's behaviors in a classroom setting than by a child's label as illustrated by the manner in which the participants in the present study rated the child after viewing the videotape. Both the LD- and normal-expectancy group teachers appropriately rated the target child's behavior more negatively on the group instruction than on the individual seatwork tape. The SECOS criteria measures documented that the teachers' ratings were justified since the child did spend less time in schoolwork on the group instruction tape than on the individual seatwork tape.

The pre- to posttest CBI comparisons provided further evidence that the LD and normal labels did not influence posttest ratings. This comparison replicated a finding of Foster and his colleagues who found that teachers' pretest ratings of a hypothetical child were more negative than their posttest ratings which suggested that the teachers were sensitive to normal behaviors (Foster & Ysseldyke, 1976; Foster et al., 1975). In the present study, the LD- and normal-expectancy group teachers did not differ in their ratings of the target child at posttest. Foster and Ysseldyke, however, found that the teachers who had received a deviancy label evaluated the child's behaviors more negatively than the teachers who were operating with a normal label. As a result, the authors proposed that the teacher ratings reflected an interaction between the child's actual behaviors and the label. Such an interaction was not found in the current study suggesting that the teachers, regardless of assigned expectancy, were influenced by the child's behavior, not the assigned label.

The addition to the CBI of the option to rate items on the videotapes as "not observed" enhanced the validity of the study. Other research using videotape segments of a child's behavior has not included such a response option. Yet, it is unreasonable to expect that all behaviors on a typical rating scale may be observed on a relatively brief videotape without making the child's behavior appear staged. The only items that were coded frequently enough to be used for posttest analyses were on the Distractibility and Attention-to-Task scales — the two scales which have been supported by research as valid discriminators of LD and nonhandicapped children (McKinney & Feagans, 1983).

The current study included the classroom setting on the videotapes, first, to show that classroom variables can attenuate bias, and, second, to improve the external validity of the findings. As in previous research, none of the children on the videotapes used in this study were handicapped. Every effort was made to insure that the behaviors of the children on the videotapes were similar to those of children in regular classrooms. Indeed, the amount of time the target child spent engaged in schoolwork was very similar to the amount of time spent on such tasks in actual classrooms (Fellers & Saudargas, 1987; Slate & Saudargas, 1986).

Although the current study did not incorporate an actual classroom nor actual LD children, the findings suggest that teachers are sensitive to the classroom ecology and will evaluate students with respect to their behavior and the context of such behavior. Future research would make further strides by using actual LD children on the videotapes or using videotapes of real classrooms (Dusek & Joseph, 1983).

REFERENCES

Algozzine, B., & Sutherland, J. (1977). The "Learning Disabilities" label: An experimental analysis. *Contemporary Educational Psychology, 2*, 292-297.

Claiborn, W. L. (1969). Expectancy effects in the classroom: A failure to replicate. *Journal of Educational Psychology, 60*, 377-383.

Dusek, J. A., & Joseph, G. (1983). The bases of teacher expectancies: A meta-analysis. *Journal of Educational Psychology, 75*, 327-346.

Dusek, J. A., & O'Connell, E. J. (1973). Teacher expectancy effects on the achievement test performance of elementary school children. *Journal of Educational Psychology, 65*, 371-377.

Fellers, G., & Saudargas, R. A. (1987). Classroom behaviors of learning disabled and nonhandicapped girls. *Learning Disability Quarterly, 10*, 231-236.

Fogel, L. S., & Nelson, R. O. (1983). The effects of special education labels on teachers' behavioral observations, checklist scores, and grading of academic work. *Journal of School Psychology, 21*, 241-251.

Foster, G., & Ysseldyke, J. (1976). Expectancy and halo effects as a result of artificially induced teacher bias. *Contemporary Educational Psychology, 1*, 37-45.

Foster, G., Ysseldyke, J. E., & Reese, J. H. (1975). "I wouldn't have seen it if I hadn't believed it." *Exceptional Children, 41,* 469-473.

Gozali, J., & Meyen, E. L. (1970). The influence of the teacher expectancy phenomenon on the academic performances of educable mentally retarded pupils in special classes. *Journal of Special Education, 4*, 417-424.

Jacobs, W. R. (1978). The effect of the learning disability label on classroom teachers' ability objectively to observe and interpret child behaviors. *Learning Disability Quarterly, 1*, 50-55.

Jose, J., & Cody, J. J. (1971). Teacher-pupil interaction as it relates to attempted changes in teacher expectancy of academic ability and achievement. *American Educational Research Journal, 8*, 39-49.

McKinney, J. D., & Feagans, L. (1983). Adaptive classroom behavior of learning disabled students. *Journal of Learning Disabilities, 16*, 360-367.

Saudargas, R. A., & Fellers, G. (1986). *State-Event Classroom Observation System.* Unpublished observation manual. University of Tennessee.

Schaefer, E. S., Edgerton, M., & Aronson, M. (1977). *Classroom Behavior Inventory.* Unpublished form.

Slate, J. R., & Saudargas, R. A. (1986). Differences in learning disabled and average children's classroom behaviors. *Learning Disability Quarterly, 9*, 61-68.

Sutherland, J., & Algozzine, B. (1979). The learning disabled label as a biasing factor in the visual motor performance of normal children. *Journal of Learning Disabilities, 12*, 17-23.

Taylor, R. L., Smiley, L. R., & Ziegler, E. W. (1983). The effects of labels and assigned attributes on teacher perceptions of academic and social behavior. *Education and Training of the Mentally Retarded, 18*, 45-51.

Yoshida, R. K., & Meyers, C. E. (1975). Effects of labeling as educable mentally retarded on teachers' expectancies for change in a student's performance. *Journal of Educational Psychology, 67*, 521-527.

Requests for reprints should be addressed to: Richard Saudargas, Department of Psychology, University of Tennessee, Knoxville, TN 37996-0900.