

Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children

<http://tes.sagepub.com/>

The Biasing Effects of Labels on Direct Observation by Preservice Teachers

R. Allan Allday, Gary J. Duhon, Sarah Blackburn-Ellis and Jamie L. Van Dycke

Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children 2011 34: 52 originally published online 1 September 2010

DOI: 10.1177/0888406410380422

The online version of this article can be found at:

<http://tes.sagepub.com/content/34/1/52>

Published by:



<http://www.sagepublications.com>

On behalf of:

Teacher Education Division of the Council for Exceptional Children

Additional services and information for *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children* can be found at:

Email Alerts: <http://tes.sagepub.com/cgi/alerts>

Subscriptions: <http://tes.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://tes.sagepub.com/content/34/1/52.refs.html>

The Biasing Effects of Labels on Direct Observation by Preservice Teachers

Teacher Education and Special Education
34(1) 52–58
© 2011 Teacher Education Division of
the Council for Exceptional Children
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0888406410380422
<http://tese.sagepub.com>


R. Allan Allday¹, Gary J. Duhon¹,
Sarah Blackburn-Ellis¹, and Jamie L. Van Dycke¹

Abstract

Observational bias can significantly affect results attained through observation. This study focused on 122 preservice teacher educators who conducted a structured observation, using momentary time sampling procedures with 10-second intervals, to measure student on-task and off-task behaviors. The experimental variable altered was the exceptionality label of the observed student to determine if the label would have a biasing effect. Labels used were oppositional defiant disorder, attention deficit hyperactivity disorder, and gifted/talented. Results suggest that an exceptionality label (i.e., oppositional defiant disorder) significantly affects observed behavior. Practical implications of the study are discussed.

Keywords

labeling, bias, preservice teachers, observation

Approximately 13.5% of school-age children receive special education services as a result of a disability (U.S. Department of Education, 2010). Many of these disabilities result in behavioral problems that manifest in the school setting and have a significantly harmful effect on academic functioning (Barkley, DuPaul, & McMurray, 1990; Cole, 1990). To remediate behavioral problems, school-based personnel often observe the child using structured direct observations. A structured direct observation refers to measurement procedures in which behaviors are systematically measured and classified, resulting in data concerning not only the behavior but also the antecedent and consequent events occurring. These observational data have proved useful in identifying functional relationships and developing interventions to remediate behavioral concerns (Hanley, Iwata, & McCord, 2003).

Effect of Labels on Observations

School personnel often make judgments and form expectations for a student with a disability based on disability label or other information received from other school personnel, before they ever observe or interact directly with the student (Carroll & Reppucci, 1978; Fogel & Nelson, 1983). This is referred to as labeling bias, which is the expectation that people may develop toward a person who has been given a particular label (Fox & Stinnett, 1996). Labeling children has resulted in

¹Oklahoma State University, Stillwater, Oklahoma, USA

Corresponding Author:

R. Allan Allday, Oklahoma State University, 249 Willard Hall, Stillwater, OK 74078, USA
Email: allan.allday@okstate.edu

differential expectations for the one being labeled (Brophy & Good, 1970; Cooper, Findley, & Good, 1982; Rosenthal & Jacobson, 1968).

Although the use of labels was initially linked with positive intention in special education, they have been shown to hinder the success of children receiving special education services (Field, Hoffman, St. Peter, & Sawilowsky, 1992; Foster & Ysseldyke, 1976). In many cases, a child is evaluated, assigned a label, and then treated differently as a result of this label (Carroll & Reppucci, 1978; Fogel & Nelson, 1983). School personnel can interpret the label negatively and might presume that a student is incapable of being as successful as students without labels (Field et al., 1992; Foster & Ysseldyke, 1976; Rosenthal & Jacobson, 1968).

Research has shown that school personnel expect a student with a label to perform more poorly on a variety of educational and social tasks than “normal” students (Gillung & Rucker, 1977), and labels such as *emotional disturbance* elicit more negative evaluations than other labels (Levin, Arluke, & Smith, 1982; Ysseldyke & Foster, 1978). However, labels may not inevitably have a negative effect; it is possible that labels such as *gifted* and *talented* may bias expectations in the opposite direction. Ultimately, if biasing information like a label is presented prior to an observation, the observation itself may become biased and the accuracy of the observation could be reduced.

Preservice Teachers and Bias

Preservice teacher education students will be in a position, usually within 4 years, to make clinical decisions based on observed behavior; therefore, it is imperative to ensure that these individuals enter the field aware of potential biasing factors. Errors caused by biases could have a detrimental effect on future outcomes of certain students because of a bias the teacher may use in assessing student functioning.

Little research has examined the effects of labels on direct observation of student behavior

by preservice teachers. Foster, Algozzine, and Ysseldyke (1980) had practicing and preservice teachers view a video of a student completing tasks and then rate the student using a referral instrument. Participants were told that the student was either “normal” or diagnosed as “emotionally disturbed.” Findings suggested that both practicing and preservice teachers were biased by the label of emotionally disturbed. Preservice teachers actually rated the emotionally disturbed student more positive than practicing teachers. This finding illuminates the fact that bias exists prior to entering the classroom and suggests that biasing effects may become more impacting as experience is gained.

Teacher Gender and Bias

Researchers have studied the effects of bias based on teacher gender on mathematics achievement (Li, 1999) and grading policies (Klein, 2004). Sideridis, Antoniou, and Padelidu (2008) found that male teachers were more than twice as likely as female teachers to identify a student with a learning disability. Wooldridge and Richman (1985) suggested that female teachers are more likely to punish male than female students due to behavioral challenges; however, Klein (2004) found that male teachers were more susceptible to biasing factors when rating student behavior. Bias may be affected by teacher gender; hence, it is important to determine if teacher gender has a biasing effect on direct observation of behavior.

Since observations are a widely used assessment method among professionals (Hintze & Matthews, 2004), accuracy of these observations must be examined to ensure reliability and validity. Thus, it is important to understand the effect of biasing events, such as the introduction of a label, on the accuracy of direct observations. The purpose of this study was twofold. First, this study examined the effect of exceptionality labels during a structured direct observation. It was hypothesized that specific labels would bias the behavioral observation. Second, this study attempted to determine if label bias was evident in

preservice teachers and if teacher gender affected the bias.

Method

Participants

Undergraduate education majors enrolled in a general special education course were asked to participate in the study. A total of 122 students from two sections of the course participated in the training—67 and 55, respectively. The group included 58% female participants and 42% male participants. Participants' major fields of study included elementary education (21%), secondary education (39%), speech pathology (7%), agricultural education (21%), and noncertification education majors (13%). Seniors made up 28% of the sample, whereas juniors and sophomores made up 57% and 15%, respectively.

Data Collection

Data collection forms designed for this study included a first page that collected demographic data. The second page provided the operational definitions of on- and off-task behaviors, followed by a page that included three time sampling grids used for practice observations. The fourth page included the time sampling grid used for the experimental data collection with the operational definitions listed below the grid.

Data were collected using momentary time sampling procedures with 10-second intervals. Data were collected for on-task and off-task behaviors using an operational definition from Simpson (1979; see Table 1). Time sampling grids containing 18 intervals (3 minutes total) were used for coded behavior. On-task behavior was coded as +, whereas off-task behavior was coded as 0.

Procedures

Upon obtaining verbal consent, students received a brief training (approximately 30 minutes; e.g., Behets, 1993; Lakes & Hoyt, 2009) concerning the intentions and methods of

Table 1. Operational Definition of Behavioral Category

Category	Operational definition
On-task	<ul style="list-style-type: none"> • Listening to instructions • Doing what the teacher instructed • Making eye contact with the task or the teacher when either is required • Seeking help in the proper manner (e.g., raising hand) • Being in the proper geographical location (e.g., being in seat when seat work is required)
Off-task	<ul style="list-style-type: none"> • Not doing what the teacher instructed • Not making eye contact with teacher or task when either is required • Not following directions • Not in the proper geographical location

momentary time sampling procedures. Students were also taught to recognize the operationally defined behaviors and were instructed on how to complete the data collection forms. Although this level of training is considered low (Volpe, DiPerna, Hintze, & Shapiro, 2005), it is consistent with both similar studies (i.e., Lakes & Hoyt, 2009) and training requirements for commonly used tools such as the Student Observation System (SOS) within the Behavior Assessment System for Children—2nd Edition (BASC-2; Reynolds & Kamphaus, 2004). In addition, many classroom teachers receive little instruction on completing behavioral observations, yet they are, at times, required to conduct observations. The level of training provided in the current study would, in many cases, be more extensive than what most practicing teachers receive.

Following this training, students were presented with a series of three practice videos (3 minutes in length; 18 intervals) that identified the target student while providing no additional information. Following each video, students totaled on-task and off-task intervals and were given opportunities to ask questions concerning data collection. Upon completion

of practice videos, students were assigned to experimental groups using “numbered heads.” This method was chosen because students of similar majors and gender sat in proximity to each other. Using numbered heads allowed for better dispersion of majors and gender.

Participants were clustered into groups of (a) no exceptionality identified, (b) exceptionality identified as attention deficit hyperactivity disorder (ADHD), (c) exceptionality identified as oppositional defiant disorder (ODD), and (d) exceptionality identified as gifted and talented. Participants from three groups were asked to exit, whereas the other remained in the room for data collection. Videos were presented in the following order: ADHD, gifted and talented, control, and ODD. A research assistant was posted outside the observation room to ensure that participants did not communicate the information presented through the video.

Data collection included watching a 3-minute video. All groups watched the same video, which provided students with information on who was being observed (i.e., a male Caucasian student) and, in the case of the experimental groups, the exceptionality being observed. This information was presented as a voiced-over text rolling marquee.

Interobserver Agreement (IOA)

The lead authors coded each practice video independently to determine intervals of on- and off-task behaviors. The lead authors have extensive training in conducting direct observation and have conducted direct observations in practical and research settings. These authors had an IOA of 94% (17 of 18 intervals) in coding the videos used in this study. For participant data to be used in this study, the participant had to attain at least an IOA of 75%. This standard of reliability was used to ensure that participants were adequately trained in using the observation system (Volpe et al., 2005). Each training video was examined to determine the level of IOA. IOA was determined by interval measure (i.e., number of agreements divided by number of co-observations).

Table 2. Mean Number of Intervals Recorded and Standard Deviations for On-Task and Off-Task Behavior Based on Identified Exceptionality Label

Exceptionality	<i>n</i>	On-task	Off-task
Control	27	4.59 (2.33)	12.93 (3.04)
ODD	26	3.88 (2.32)	14.04 (2.52)
ADHD	27	4.67 (2.13)	13.33 (2.13)
Gifted/talented	24	5.83 (2.88)	12.17 (2.88)

Note: Standard deviations are in parentheses. ODD = oppositional defiant disorder; ADHD = attention deficit hyperactivity disorder.

Of the 122 participants, a total of 106 met the minimum IOA on the third training video. Their data were used in data analysis.

Results

Mean results of the observation are shown in Table 2. A two-way analysis of variance (ANOVA) (4×2) was conducted to determine the effects of exceptionality label and observer gender on ratings of observed off-task behavior. Results indicate that an exceptionality label significantly affected observers when rating off-task behavior, $F(3, 96) = 3.81$, $p = .01$, $\eta^2 = .11$, specifically for ODD ($M = 14.04$) and gifted/talented ($M = 12.17$). The interaction effect between exceptionality label and observer gender was approaching significance ($p = .056$). Other observer demographics did not produce statistically significant results for off-task behaviors.

Discussion

The purpose of the present study was to determine if bias exists, based on exceptionality label, in structured behavioral observations of preservice teacher educators. Results of the present study suggest that observational biases exist with preservice educators. Participants in the current study poorly rated the student identified as ODD while rating the other labels higher. Although the label of emotional disturbance was not used in this research, the use of a highly correlated term (i.e., ODD) produced significant findings that are consistent with

prior research (Levin et al., 1982; Ysseldyke & Foster, 1978). One possible rationale for this finding is that observers can be significantly influenced to produce data that are consistent with the expectations that they hold (Rosenthal & Jacobson, 1968). Understanding the differing expectations that observers have for students with exceptionality labels, especially the label of ODD, is beyond the current study's scope, although it is a worthy undertaking for future endeavors. Overall, this study provides a basic view into how preservice teachers perceive different exceptionality labels.

Limitations and Further Research

As with all research, there are limitations within the study. First, the use of a 3-minute observation video may have limited the biasing effect. Since only 18 intervals of data were collected, it is plausible that a stronger effect could be attained by the addition of more intervals. Future studies could conduct the research using a longer video (e.g., 20 minutes), which would increase the number of intervals. A second limitation could be the use of students enrolled in their only special education course. Student perceptions of exceptionality labels may become more positive while taking the class. Further research should attempt to use preservice educators prior to the special education course. In addition, comparisons between preservice and practicing teachers would show if experience affects bias. Third, it is difficult to generalize results of the study due to the use of video for observation purposes. Viewing behavior via video versus live action could produce different results because of additional extraneous variables present in live action observations. Further inquiry should provide participants an opportunity to observe live action within an appropriate educational setting.

Practical Implications

The use of behavioral observations serves a valuable purpose in that education professionals depend on behavioral observations to assist with diagnoses, functional behavioral assessments,

reevaluations, and determinations of intervention effectiveness. Although observations are both a necessary and significant aspect of the evaluation process, it is always important to remember that potential biasing variables exist, in addition to exceptionality labels. For example, perceptual bias of the observer, observer drift, and student/teacher reactivity to the presence of the observer may alter the accuracy of direct observations (Kazdin, 1978; Skinner, Dittmer, & Howell, 2000). These biasing variables are ever present; therefore, it is important for teacher educators to prepare their students to be aware of biasing variables.

There are several practical exercises that teacher educators can perform that may decrease the effects of biasing during direct observations. Preservice teachers should be provided with opportunities to conduct direct behavioral observations during their training whether in a live, simulated, or video format. These observations should include a diverse set of behaviors, settings, student ethnicities/gender, and exceptionality labels. Following the observation, preservice educators should be able to discuss and identify potential biasing variables that could have affected their results.

A second exercise that may be useful in decreasing biasing effects is providing preservice teachers with training on how to develop and use operational definitions in direct observations. When students understand the importance of using this definition, this may decrease the influence of the biasing variables. Along with this understanding, it would be useful to employ multiple observers who are unaware of any existing labels to ensure more accurate results. When multiple observers are used, potential biases will be highlighted and observers can recalibrate their operational definition to attempt to remove the biases.

In conclusion, the study of bias is not new to education. There is a wealth of information that has been disseminated to the professional community on issues associated with psychometric testing bias, racial/ethnic biases,

and gender biases. Although researchers have established that various biases exist, little research has been conducted on when these biases form in education professionals and the extent to which these biases reach into professional practice. Through the results of the present study, it is clear that there is a bias present during preservice training, specifically when conducting behavioral observations. Helping preservice and practicing education professionals to observe behaviors, not labels, could assist in providing more accurate observations and better services for students who have an exceptionality label.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Financial Disclosure/Funding

The author(s) received no financial support for the research and/or authorship of this article.

References

- Barkley, R., DuPaul, G., & McMurray, M. (1990). A comprehensive evaluation of attention deficit disorder with and without hyperactivity. *Journal of Consulting and Clinical Psychology, 58*, 775-789.
- Behets, D. (1993). Systematic observation training for pre-service physical education teachers. *Physical Educator, 50*, 87-94.
- Brophy, J., & Good, T. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology, 61*(5), 365-374.
- Carroll, C., & Reppucci, N. (1978). Meanings that professionals attach to labels for children. *Journal of Counseling and Clinical Psychology, 46*, 372-374.
- Cole, D. (1990). The relation of social and academic competence to depressive symptoms in childhood. *Journal of Abnormal Psychology, 99*, 422-429.
- Cooper, H., Findley, M., & Good, T. (1982). Relations between student achievement and various indexes of teacher expectations. *Journal of Educational Psychology, 74*(4), 577-579.
- Field, S., Hoffman, A., St. Peter, S., & Sawilowsky, S. (1992). Effects of disability labels on teachers' perceptions of students' self-determination. *Perceptual and Motor Skills, 75*(3), 931-934.
- Fogel, L., & Nelson, R. (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., Algozzine, B., & Ysseldyke, J. (1980). Classroom teacher and teacher-in-training susceptibility to stereotypical bias. *Personnel and Guidance Journal, 59*, 27-30.
- Foster, G., & Ysseldyke, J. (1976). Expectancy and halo effects as a result of artificially induced teacher bias. *Contemporary Educational Psychology, 1*, 37-45.
- Fox, J., & Stinnett, T. (1996). The effects of labeling bias on prognostic outlook for children as a function of diagnostic label and profession. *Psychology in the Schools, 33*, 143-152.
- Gillung, T., & Rucker, C. (1977). Labels and teacher expectations. *Exceptional Children, 43*(7), 464-465.
- Hanley, G., Iwata, B., & McCord, B. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147-185.
- Hintze, J., & Matthews, W. (2004). The generalizability of systematic direct observations across time and setting: A preliminary investigation of the psychometrics of behavioral observation. *School Psychology Review, 33*, 158-170.
- Kazdin, A. (1978). Artifact, bias, and complexity of assessment: The ABCs of reliability. *Journal of Applied Behavior Analysis, 10*, 141-150.
- Klein, J. (2004). Who is most responsible for gender differences in scholastic achievements: Pupils or teachers? *Educational Research, 46*, 183-193.
- Lakes, K., & Hoyt, W. (2009). Applications of generalizability theory to clinical child and adolescent research. *Journal of Clinical Child and Adolescent Psychology, 38*, 144-165.
- Levin, J., Arluke, A., & Smith, M. (1982). The effects of labeling students upon teachers' expectations and intentions. *The Journal of Social Psychology, 118*, 207-212.

- Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research, 41*, 63-76.
- Reynolds, C., & Kamphaus, R. (2004). *Behavior assessment system for children (BASC-2)* (2nd ed.). Circle Pines, MN: American Guidance Services.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York, NY: Holt, Rinehart and Winston.
- Sideridis, G., Antoniou, F., & Padelidi, S. (2008). Teacher biases in the identification of learning disabilities: An application of the logistic multi-level model. *Learning Disability Quarterly, 31*, 199-209.
- Simpson, R. (1979). Pupil-teacher classroom observation scale. *Diagnostique, 5*, 12-17.
- Skinner, C., Dittmer, K., & Howell, L. (2000). *Direct observation in school settings: Theoretical issues*. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in the schools: Theory, research and clinical foundations* (pp. 19-45). New York: Guilford.
- U.S. Department of Education, National Center for Education Statistics. (2010). *Digest of Education Statistics, 2009* (Report No. NCES 2010-013). Retrieved from Institute of Education Sciences, National Center for Education Statistics website: <http://nces.ed.gov/fastfacts/display.asp?id=64>
- Volpe, R., DiPerna, J., Hintze, J., & Shapiro, E. (2005). Observing students in classroom settings: A review of seven coding schemes. *School Psychology Review, 34*, 454-474.
- Wooldridge, P., & Richman, C. (1985). Teachers' choice of punishment as a function of a student's gender, age, race and IQ level. *Journal of School Psychology, 23*(1), 19-29.
- Ysseldyke, J., & Foster, G. (1978). Bias in teachers' observations of emotionally disturbed and learning disabled children. *Exceptional Children, 44*(8), 613-615.

Bios

R. Allan Allday is an assistant professor of special education at Oklahoma State University.

Gary J. Duhon is an associate professor of school psychology at Oklahoma State University.

Sarah Blackburn-Ellis is a doctoral student in school psychology at Oklahoma State University.

Jamie L. Van Dycke is an assistant professor of special education at Oklahoma State University.