Hammill Institute on Disabilities

Journal of Positive Behavior Interventions
2019, Vol. 21(1) 3–18

DOI: 10.1177/1098300718766657
jpbi.sagepub.com
SAGE

Evidence Review for Teacher Praise to Improve Students' Classroom Behavior

Tara C. Moore, PhD, BCBA-D[1], Daniel M. Maggin, PhD[2], Kelly M. Thompson, PhD[1], Jason R. Gordon, EdS[1], Stephanie Daniels, BA[1], and Lucinda E. Lang, PhD[3]

Abstract

In this study, a systematic review was conducted to examine the evidence base for teacher praise for students without severe disabilities in K–12 classroom settings. Specifically, reviewers followed standards presented by the Council for Exceptional Children and the What Works Clearinghouse to evaluate the methodological quality of 30 studies meeting inclusion criteria. Only 11 studies were rated as being methodologically sound. The review team then classified effects of the 11 methodologically sound studies and provided an overall evidence-based classification for the practice using both sets of standards. Results indicated there is currently insufficient evidence to identify teacher praise as an evidence-based practice for this population of students. A further evaluation of this research base provided no discernable patterns for when and for whom teacher praise is likely to be effective. Implications for practice and future research are provided.

Keywords

teacher praise, problem behavior, evidence-based practice, research-based practice

As early as the 1960s, researchers (e.g., Hall, Lund, & Jackson, 1968; Hall, Panyan, Rabon, & Broden, 1968; Madsen, Becker, & Thomas, 1968) examined the effects of teacher praise as a form of positive reinforcement to improve students' classroom behavior. Researchers have extended this line of research to examine effects of teacher praise on students' appropriate and inappropriate behaviors (e.g., Blaze, Olmi, Mercer, Dufrene, & Tingstom, 2014; Broden, Bruce, Mitchell, Carter, & Hall, 1970; Ferguson & Houghton, 1992) as well as students' academic performance (e.g., Hasazi & Hasazi, 1972; Kirby & Shields, 1972; Sutherland & Wehby, 2001) for individual students and for classes of students. More recently, researchers (e.g., Allday et al., 2012; Duchaine, Jolivette, & Fredrick, 2011; Sutherland, Wehby, & Copeland, 2000) have examined effects of contingent (occurring because of or following the desired student behavior) behavior-specific praise (i.e., teacher providing verbal praise, which explicitly specifies the desired student behavior) to improve student behavior.

Based on this body of work, providing teacher praise is often recommended as a classroom and behavior management practice, which is supported by research (e.g., Conroy, Sutherland, Snyder, & Marsh, 2008; Epstein, Atkins, Cullinan, Kutash, & Weaver, 2008; Lewis, Hudson, Richter, & Johnson, 2004; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). However, it is not enough for teachers to simply select and implement practices supported by *some* research. When possible, teachers should implement practices found to be effective across *multiple, high-quality* research studies. Kretlow and Blatz (2011) clarified for teachers, the distinction between *research-based* practice (i.e., "practices that have been studied in some way," p. 10) and *evidence-based* practice (i.e., practices with a "magnitude" of quality empirical intervention research studies and demonstrate the positive effects of the practice). For Kretlow and Blatz, education-focused government agencies, professional organizations, and researchers are responsible for conducting systematic syntheses of bodies of research to help identify for teachers' practices, which are research-based (i.e., based on *some* research) and which are evidence-based (i.e., based on results from a *body of high-quality* research).

Conducting Systematic Reviews to Identify Evidence-Based Practices

The Council for Exceptional Children (CEC), a professional organization for special education, and the U.S.

[1]The University of Tennessee, Knoxville, USA
[2]The University of Illinois at Chicago, USA
[3]Carson–Newman University, Jefferson City, TN, USA

Corresponding Author:
Tara C. Moore, Assistant Professor, Department of Theory & Practice in Teacher Education, The University of Tennessee, 409 Bailey Education Complex, Knoxville, TN 37996-3442, USA.
Email: tara.moore@utk.edu

Action Editor: Wendy Oakes

Department of Education's Institute of Education Sciences (IES) have each developed standards for evaluating the quality of and outcomes of research to determine whether a practice is evidence-based: *CEC Standards for Evidence-Based Practices in Special Education* (CEC Standards; CEC, 2014) and IES's *What Works Clearinghouse: Single-Case Design Technical Documentation* (WWC Standards; Kratochwill et al., 2010). Both standards provide guidelines for review teams examining methodology and effects of single-subject research designs; such designs are typical of published teacher praise studies.

According to both CEC and WWC Standards, review teams interested in evaluating the evidence base of a practice for a specified population, setting, and so forth, should follow a systematic, multistep review process. First, reviewers should *evaluate the methodological quality of studies* identified as meeting inclusion criteria for the review. Then, only studies that meet methodological quality standards are examined further to *classify the observed effects of the practice*. In this step, the review team considers the effectiveness of the practice based on results reported in each methodologically sound study. Finally, the review team summarizes the effects across studies to *classify the body of evidence for the practice* for the specified population. In this last step, the review team summarizes and combines effects across the body of methodologically sound studies to determine whether the practice can be deemed evidence based or not.

There is considerable overlap between CEC and WWC Standards for evaluating evidence from single-subject research. However, there are differences in terminology (e.g., studies identified as being "methodologically sound" per CEC Standards vs. "meeting design standards" or "meeting with reservations" per WWC Standards). There are also differences in specified criteria such as methodological quality indicators included and CEC Standards requiring a study to have at least three participants to consider the effects of practice (WWC Standards allow for consideration of studies with only one participant). Furthermore, CEC Standards provide specific guidelines to account for gradations to classify the overall level of evidence on a practice (e.g., "evidence-based, potentially evidence-based," "mixed evidence," "insufficient evidence," and "negative effects"), whereas review teams using WWC Standards would determine whether the overall body of evidence met or did not meet WWC Standards' recommended threshold for combining results of studies to identify evidence-based practices. These differences may result in different overall classifications of the body of evidence for a practice depending on which set of standards the review team used. Specifically, a review team could classify a practice to be evidence-based using one set of standards and not evidence-based using the other set. This may be problematic for both researchers and practitioners.

Purpose of Current Review

Teacher praise is certainly a research-based practice; it is a low-intensity strategy (see Lane, Menzies, Ennis, & Oakes, 2015) requiring low effort and minimal teacher time or resources. Researchers have presented several narrative reviews of the research base on teacher praise, including summaries of research supporting teacher praise and recommendations for practice (e.g., Gable, Hester, Rock, & Hughes, 2009; Lewis et al., 2004; Simonsen et al., 2008; Sutherland, 2000), a recent summary of descriptive research on naturally occurring rates of teacher praise (Jenkins, Floress, & Reinke, 2015), and reviews of research examining the effects of performance feedback to improve teachers' use of praise (Cavanaugh, 2013; Sweigart, Collins, Evanovich, & Cook, 2016). However, the quality and extent of the body of research on the effects of teacher praise has yet to be examined.

The purpose of this study, then, was to systematically review the body of research examining the effects of teacher praise using first CEC Standards and then using WWC Standards to determine whether teacher praise is an evidence-based practice for students without severe disabilities in K–12 classroom settings. For the purposes of this review, teacher praise was defined as "favorable verbal or nonverbal [teacher] attention" (Jenkins et al., 2015, p. 464), which indicates approval of student behavior. Although it is often recommended that teachers use behavior-specific praise (e.g., Brophy, 1981), this review was not limited to general or behavior-specific praise nor did we intend at outset to compare effects across types of praise. A secondary purpose is to provide a summary of when and for whom teacher praise was or was not effective. This review focuses on examining outcomes for students without severe disabilities to inform practice selection for problem behaviors teachers are most likely to encounter in general and special education classrooms.

Our goal is to help teachers and school leaders understand the overall quality of the body of research on teacher praise, the level of evidence supporting the effectiveness of teacher praise, limitations to the research base, and for whom teacher praise might be effective. We also hope to help teachers understand current and future directions for research on teacher praise; help them understand, interpret, and critique labels such as "research-" or "evidence-based" as applied to teacher praise; and help them make instructional decisions grounded in our best science. Last, we hope to highlight areas for future research and discussion on teacher praise.

Method

The review team consisted of two professors of special education (first and second authors) at two universities, each

with advanced training in research methodology related to group comparison research and single-subject research. Each professor teaches graduate-level courses related to research methods for special education, supervises graduate students in research, and conducts research. The third and fourth review team members were advanced PhD students in school psychology and special education (third and fourth authors, respectively); they completed advanced coursework related to research methods for special education and assisted on research projects.

Identification of Studies to Include in Review

Relevant studies were identified by the following procedures. First, electronic searches were conducted with the research databases Educational Source, ERIC, and PsycINFO, three frequently used databases for educational research, using the following descriptors: (a) *praise* and (b) *teacher** (to allow for "teacher(s)") and *praise*, using "AND" as a Boolean search operator to identify articles containing both "teacher(s)" and "praise." Limitations were set for only returning results from peer-reviewed journals, for articles written in English, and for publication dates through 2016. The electronic searches yielded 7,016 articles with duplicates included. The first author reviewed the titles and abstracts of the articles to determine potential eligibility. A resulting 100 articles were then retrieved and reviewed by the first author to determine final eligibility. For a sample (32%) of articles, a second member of the review team independently reviewed articles to evaluate reliability of search procedures (87.50% agreement on eligibility determination across sample). For some articles, the first author also consulted with the second author and came to a consensus about eligibility. A resulting 24 studies were identified as meeting eligibility criteria via the electronic search. Next, an archival search was conducted by reviewing citations included in each eligible article. Here, an additional 39 articles were reviewed. A second reviewer again reviewed a sample (25.64%) of articles in this step of the search, and the two reviewers had 100% agreement on eligibility determination. The archival search resulted in an additional six eligible articles. Finally, an archival search was conducted on each of the narrative reviews cited earlier that focused on teacher praise; however, no additional eligible articles were identified. In total, 30 articles met eligibility criteria.

Eligible studies had to meet the following eight criteria: (a) the researcher(s) used a valid experimental or quasi-experimental research design including either those drawn from group comparison or single-case methods, (b) the intervention being investigated had to be a direct, planned manipulation of teacher praise, (c) the classroom teacher had to be the implementer of the intervention, (d) the study was conducted with K–12 students (i.e., the target population), (e) students identified as having severe disabilities such as profound intellectual disabilities or lower incidence disabilities (including autism) were excluded, (f) the research report was published in a peer-reviewed scholarly journal no later than 2016, (g) the intervention was conducted in a school classroom setting (e.g., not in the cafeteria, gym), and (h) the study had to include examinations of student classroom behavior during typical instructional situations as the dependent variable (i.e., target outcomes): either appropriate behaviors (e.g., academic engagement, on-task) or inappropriate behaviors (e.g., disruptive behavior).

Studies using AB single-subject research designs or pre- and posttest designs without a control or comparison group were not included. Studies that included only academic performance as the dependent variable(s) (e.g., correct responses, percent correct, and accuracy) were not included. Studies that involved praise used in combination with other specifications for how teachers attended to student behavior (e.g., praise plus ignoring, praise plus reprimanding) were included because they represented only teacher attention as potential consequences of behavior. However, studies that paired teacher praise with other rewards (e.g., tokens, points, playing a game) or response cost system (e.g., losing break time for inappropriate behavior) were not included because additional rewards or loss of privileges may have served as reinforcers or punishers for behavior. The review did not include studies conducted in preschools, studies on teacher praise notes (i.e., written notes of encouragement), or studies where students were taught to recruit positive attention or praise from teachers.

Each of the 30 studies meeting eligibility criteria used single-subject research designs (each eligible study is noted by an asterisk in the reference list and included in Table 1). Within these studies, 76 cases (i.e., the subject or participant identified as the unit of analysis) met eligibility criteria. The student was the unit of analysis for 14 studies (40 cases), and a class of students (or group of students selected from a class) was the unit of analysis for 16 studies (36 cases). For six studies, one or more cases were excluded for not meeting all eligibility criteria and are not summarized in the "Results" section: Duncan, Dufrene, Sterling, and Tingstrom (2013; participant in a preschool setting); Hall et al. (1971; intervention included access to games); Hall, Panyan, et al. (1968; intervention included access games or loss of break as a response cost); Madsen et al. (1968; not a valid design for two participants); and Rathel, Drasgow, Brown, and Marshall (2014; participant with autism). However, the review team took into consideration reported results for noneligible cases if they were necessary in determining the classification of effects for a study (e.g., the noneligible student represented the third demonstration of effect for a study). For more information on eligibility determinations, contact the first author.

Table 1. Ratings of Methodological Quality by Study.

	CEC Standards		WWC Standards	
Study	Overall[a]	By quality indicators[b]	Overall[a]	By standard items[c]
Student-level cases				
Allday et al. (2012)	Sound	All met	Meets	All met
Broden, Bruce, Mitchell, Carter, and Hall (1970)	Sound	All met	Meets with reservations	Met with reservations: 4
Duncan, Dufrene, Sterling, and Tingstrom (2013)	Sound	All met	Meets with reservations	Met with reservations: 4
Hall et al. 1971	—	Not met: 5.1, 5.2, 6.1, 7.4	—	Not met: 1, 2, 4
Hall, Lund, and Jackson (1968)	—	Not met: 5.1, 5.2, 6.1, 7.5	—	Not met: 1, 2 Met with reservations: 4
Kirby and Shields (1972)	—	Not met: 5.1, 5.2, 6.1	—	Not met: 1
Madsen, Becker, and Thomas (1968)	—	Not met: 6.6	—	Not met: 4
Martens, Hiralall, and Bradley (1997)	—	Not met: 6.5	—	Not met: 3, 4
Rathel, Drasgow, Brown, and Marshall (2014)	Sound	All met	Meets with reservations	Met with reservations: 4
Starkweather Matheson and Shriver (2005)	Sound	All met	Meets with reservations	Met with reservations: 4
Thompson, Marchant, Anderson, Prater, and Gibb (2012)	Sound	All met	Meets	All met
Trolinder, Choi, and Proctor (2004)	—	Not met: 7.5	—	Not met: 2 Met with reservations: 4
Wasik, Senn, Welch, and Cooper (1969)	—	Not met: 7.5	—	Not met: 2
Wood, Umbreit, Liaupsin, and Gresham (2007)	—	Not met: 6.6, 7.4	—	Not met: 4
Class-level cases				
Blaze, Olmi, Mercer, Dufrene, and Tingstom (2014)	Sound	All met	Meets with reservations	Met with reservations: 4
Cossairt, Hall, and Hopkins (1973)	—	Not met: 6.5	—	Not met: 3, 4
Duchaine, Jolivette, and Fredrick (2011)	Sound	All met	Meets with reservations	Met with reservations: 4
Dufrene, Lestremau, and Zoder-Martell (2014)	—	Not met: 6.5	—	Not met: 3, 4
Ferguson and Houghton (1992)	—	Not met: 7.5	—	Not met: 2 Met with reservations: 4
Hall, Panyan, Rabon, and Broden (1968)	—	Not met: 7.5	—	Not met: 2
Haydon and Musti-Rao (2011)	—	Not met: 6.5	—	Not met: 3, 4
Hollingshead, Kroeger, Altus, and Brubaker Trytten (2016)	—	Not met: 6.6, 7.4	—	Not met: 4
Houghton, Wheldall, Jukes, and Sharpe (1990)	—	Not met: 6.5	—	Not met: 3, 4
McAllister, Stachowiak, Baer, and Conderman (1969)	—	Not met: 6.5	—	Not met: 3, 4
Mesa, Lewis-Palmer, and Reinke (2005)	—	Not met: 6.5	—	Not met: 3, 4
Myers, Simonsen, and Sugai (2011)	—	Not met: 7.5	—	Not met: 2 Met with reservations: 4
Pisacreta, Tincani, Connell, and Axelrod (2011)	Sound	All met	Meets	All met
Sutherland, Wehby, and Copeland (2000)	Sound	All met	Meets with reservations	Met with reservations: 4
Thomas, Becker, and Armstrong (1968)	Sound	All met	Meets	All met
Workman, Kindall, and Williams (1980)	—	Not met: 7.5	—	Not met: 2

Note. CEC = Council for Exceptional Children; WWC = What Works Clearinghouse.
[a]Indicates overall rating not meeting methodological standard. [b]CEC quality indicators: 5.1 (reports implementation fidelity), 5.2 (reports implementation fidelity related to dosage), 6.1 (systematically manipulates the independent variable), 6.5 (three demonstrations of effects), 6.6 (baseline phases include at least three data points), 7.4 (at least three data points per phase for demonstrations of effects), 7.5 (provides evidence of adequate interobserver reliability). [c]WWC Standards items: 1 (independent variable systematically manipulated), 2 (provides evidence of adequate interobserver agreement), 3 (includes three attempts to demonstrate effects), 4 (*met* if attempts at demonstrations of effects include at least three data points per phase, *met with reservations* if attempts at demonstrations of effects include at least three data points per phase).

Step 1: Evaluate Methodological Quality of Studies

CEC Standards. The review team conducted an initial evaluation of each of the 30 studies and their 76 cases to determine which studies and cases met *all* of the quality indicators for single-subject research design following the CEC Standards. Two members of the review team independently read each study to rate whether or not the study met each of the quality indicators and then to assign an overall

rating for each study. Only studies meeting all quality indicators for all eligible cases were given an overall rating of *methodologically sound*.

The reviewers then compared their independent ratings to evaluate their level of agreement. Interrater agreement was calculated for (a) total quality indicator items across all cases, (b) overall ratings (i.e., overall rating of being methodologically sound or not) by case, and (c) overall ratings by study, by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100 to get a percentage of interrater agreement. Interrater agreement was 99.70% for total quality indicator items across all cases, 96.05% for overall ratings by case, and 93.33% for overall agreement by study. Reviewers discussed all disagreements and assigned consensus ratings to serve as the final ratings.

WWC Standards. To further validate the evaluation of methodological quality of these studies, reviewers rated each study using the WWC Standards. Two review team members independently read each study to rate whether each study met each WWC design standard item by case and to assign an overall rating by case and an overall rating by study as *meets evidence standards, meets evidence standards with reservations*, or *does not meet evidence standards*. Interrater agreement was 100% for total design criteria items across all cases, overall ratings by case, and overall agreement by study.

Special considerations for this review. There were two special considerations agreed upon and used by the review team when evaluating studies. First, researchers had to assess and report implementation fidelity or dosage throughout the study to demonstrate control and systematic manipulation of the independent variable (CEC Quality Indicators 5.1, 5.2; WWC Item 1). Typically, this involved reporting data (either in a table or graph) on levels of praise during relevant phases of the study to demonstrate that teacher praise increased during intervention phases or conditions. Second, researchers had to report enough information to demonstrate that adequate interobserver reliability was assessed throughout all relevant phases of the study (CEC Quality Indicator 7.5, WWC Item 2). Each study eligible for this review included dependent variables assessed through direct observation procedures. Therefore, the review team agreed upon a criterion for researchers to report either data or a statement to indicate interobserver agreement was assessed for at least 20% of observation sessions across all relevant phases of the study. Although this criterion is stringent, it is consistent with WWC and other contemporary recommendations (e.g., Ayers & Gast, 2010; Kazdin, 2011). Contact the first author for more information on the review team's interpretation of standard items.

Step 2: Classify Effects for Methodologically Sound Studies

Next, the review team classified the effects presented in each methodologically sound study by CEC Standards and then by WWC Standards. Only studies with a final overall rating of *methodologically sound* (i.e., meeting all of the CEC Standards quality indicators) or being rated as *meets design standards* or *meets design standards with reservations* (per WWC Standards design criteria) were evaluated further to classify the effects of teacher praise. For each study, two members of the review team independently evaluated graphed outcomes (using visual analysis; e.g., Gast & Spriggs, 2010; Kratochwill et al., 2010) on each relevant dependent variable, focusing on only phase contrasts relevant to the purposes of this review. Reviewers rated whether or not an effect was demonstrated for *each case*. For multiple baseline experiments, this involved rating effects from baseline to intervention phase for each case in the study. For reversal designs, this involved rating effects by comprehensively considering demonstrated effects across relevant phase contrasts for each case (e.g., ABAB). Interrater agreement for classification of effects using the CEC Standards was 100%. Overall effects of each study were classified using both CEC Standards and WWC Standards as described below.

CEC Standards. Classifications of effects were based on the

> number and proportion of participants [i.e., cases] in a study for whom a functional relationship [between the independent and dependent variables] is established by reviewers' use of standard methods of visual analysis [of data and graphs presented in the research report]. (CEC, 2014, p. 7)

Each study was classified as having *positive effects, negative effects*, or *neutral or mixed effects*. For positive effects classification, the results of the study must demonstrate a functional relationship indicating a "meaningful, therapeutic change in the targeted dependent variable" (CEC, 2014, p. 7) for at least 75% of relevant cases. CEC Standards also require a minimum of three total relevant cases for positive effects classification and no results should indicate a nontherapeutic change in student behavior. For negative effects classification, the results of a study must demonstrate functional relationship indicating nontherapeutic change in student behavior for at least 75% of relevant cases, and there should be a minimum of three cases. If a study does not meet criteria for either positive or negative effects classification, then it should be classified as having neutral or mixed effects. Classifications of effects were not provided to studies with fewer than three eligible participants.

WWC Standards. Using the WWC Standards, each study that *meets design standards* or *meets design standards with reservations* was classified as providing *strong evidence, moderate*

evidence, or *no evidence*. For a *strong evidence* classification, the study must provide "three demonstrations of the intervention effects with no non-effects" (Kratochwill et al., 2010, p. 16) via visual analysis. For a *moderate evidence* classification, a study must "provide three demonstrations of effects and also include at least one demonstration of a non-effect" (Kratochwill et al., 2010, p. 16). A *no evidence* classification is assigned to studies which do not provide three demonstrations of an effect. In contrast to CEC Standards, WWC Standards provide no stipulation that a study must have a minimum of three total relevant cases to receive a positive evidence classification. To illustrate, a study including only one participant could receive a *strong evidence* rating using WWC Standards if a reversal design was used to provide three demonstrations of effects: (a) from baseline to intervention phase, (b) from intervention to reversal phase, (c) from reversal to reimplementation of the intervention phase.

As outlined in WWC Standards, effect size estimations were conducted for studies providing *strong evidence* or *moderate evidence* per WWC Standards. The review team calculated effect sizes indices, percentage of nonoverlapping data (PND; Scruggs, Mastropieri, & Casto, 1987) and Tau-*U* (Parker, Vannest, Davis, & Sauber, 2011) using the Single Subject Research web-based calculator for calculations (Vannest, Parker, & Gonen, 2011).

Step 3: Classify the Overall Evidence Base

Last, the review team classified the overall body of evidence for teacher praise as an intervention for the specified student population; classification of effects was determined using visual analysis of graphed data within each study. Per CEC Standards, teacher praise was classified as either *an evidence-based practice, a potentially evidence-based practice, mixed effects, insufficient evidence*, or *negative effects* based on the number of methodologically sound single-subject studies with positive effects, total participants across studies, and the ratio of methodologically sound studies with positive effects to methodologically sound studies with neutral or mixed effects. Per WWC Standards, teacher praise was classified as meeting or not meeting standards for being an evidence-based practice based on the number of studies meeting methodological (i.e., design) standards (and their level of evidence), diversity in research teams and geographic location of studies, and number of experiments conducted across studies. The first and second authors came to consensus about the evidence-based classification using each set of standards.

Summary of Evidence Base

The review team examined each methodologically sound study to gather information about study, student, and intervention characteristics to provide a summary of when and for whom teacher praise was and was not effective and to provide information about social validity.

Results

Evaluation of Methodological Quality of Studies

Of the 30 studies identified as being eligible for inclusion in this review, 11 studies were rated as being methodologically sound (i.e., meeting all relevant quality indicators for all eligible cases per CEC Standards and either *meets evidence standards* or *meets evidence standards with reservations* per WWC Standards for design criteria), yielding 20 student-level cases and 12 class-level cases meeting methodological standards (i.e., 32 total cases; see Table 1). There was exact agreement between the review team's methodology ratings using both sets of standards; studies rated as being *methodologically sound* using CEC Standards were also rated as *meets design standards* or *meets design standards with reservations* for WWC Standards, and studies rated as not methodologically sound by one set of standards were also rated as not being methodologically sound by the other set of standards. The distinction between some studies meeting evidence standards and some studies meeting evidence standards with reservations using WWC Standards was due to studies having at least five data points per relevant phase (*meeting standards*) or studies having at least three data points per relevant phase (*meeting standards with reservations*). CEC Standards only required at least three data points per phase.

The remaining 19 studies were rated as not methodologically sound (see Table 1). The primary methodological concerns included in the following. (a) Did not assess or report implementation fidelity or dosage throughout the study (thus compromising internal validity; CEC Quality Indicators 5.1, 5.2, 6.1; WWC Item 1). (b) Internal validity was compromised because the study's design did not provide at least three demonstrations of experimental effects at three different times (e.g., multiple baseline across two participants; CEC Quality Indicator 6.5; WWC Item 3). (c) Phase(s) had fewer than three data points in relevant phases for one or more cases (CEC Quality Indicators 6.6, 7.4; WWC Standards Item 4). (d) Researchers did not provide enough information to demonstrate that adequate interobserver reliability was assessed throughout all phases of the study (CEC Quality Indicator 7.5; WWC Standards Item 2). The stringent criterion of providing information to indicate interobserver agreement was assessed for at least 20% of observation sessions across all relevant phases resulted in six, otherwise methodologically sound studies being rated as not meeting both CEC and WWC Standards.

Classification of Effects of Methodologically Sound Studies

Visual analysis. The classifications of the effects of teacher praise based on visual analysis are presented at the study level in Table 2 and for each eligible case in Appendix A (student-level cases) and Appendix B (classroom-level cases). The review team rated positive effects for 14 of 20

Table 2. Summary of 11 Methodologically Sound Studies.

Study and geographic location	Number and type of eligible experiment(s)	Intervention and effect on students' behavior[a]	Classification of effects for study	
			CEC Standards	WWC Standards
Student-level cases				
Allday et al. (2012) Southwest and Midwest United States	One multiple baseline across four teachers' students (one to two students per teacher); seven students	Positive effect for "on-task" behavior for four out of seven students (57.14%) resulting from training teachers to provide behavior-specific praise to class	*Mixed effect*; positive effect for at least three participants, but less than 75%	*Moderate evidence*; three demonstrations of effects across tiers and at least one demonstration of a noneffect
Broden, Bruce, Mitchell, Carter, and Hall (1970) Midwest United States	Two within-case reversal; two students	Positive effect for "appropriate attending" for two out of two students (100%) when teacher attention provided for attending behavior	*No classification of effects*; less than three cases	*Strong evidence*; three demonstrations of effect in each within-case experiments
Duncan, Dufrene, Sterling, and Tingstrom (2013) Southeast United States	One multiple baseline across three teachers' students (one student not eligible); two eligible students	Positive effect for "disruptive behavior" for three out of three students (100%) resulting from training teachers to provide behavior-specific praise to target student	*Positive effect*; positive effect for at least 75% of cases; at least three cases in study	*Strong evidence*; three demonstrations of effects across tiers
Rathel, Drasgow, Brown, and Marshall (2014) Southeast United States	One multiple baseline across four teachers' students (one student not eligible); three eligible students	Positive effect for four out of four students (100%) resulting from training teachers to provide behavior-specific praise to target student	*Positive effect*; positive effect for at least 75% of cases; at least three cases in study	*Strong evidence*; three demonstrations of effects across tiers
Starkweather Matheson and Shriver (2005) Midwest United States	One multiple baseline across three teachers' students; three students	Positive effect for "academic behavior" for two out of three students (67%) when praise added to an effective commands intervention	*Neutral (or no) effect*; effect not demonstrated at least three times	*No evidence*
Thompson, Marchant, Anderson, Prater, and Gibb (2012) Western United States	One multiple baseline across three teachers' students; three students	Positive effect for "on-task behavior" for one out of three students (33%) in relevant phase contrasts where praise increased	*Neutral (or no) effect*; effect not demonstrated at least three times	*No evidence*
Class-level cases				
Blaze, Olmi, Mercer, Dufrene, and Tingstom (2014) Southeast United States	Four within-case reversals for four teachers' classes	Positive effect for "appropriately engaged" and "disruptive" behavior for four out of four *classes* (100%) when comparing phases of loud or quiet behavior-specific praise with phases of decreased praise	*Positive effect*; positive effect for at least 75% of cases; at least three cases in study	*Strong evidence*; three demonstrations of effects in each within-case experiment
Duchaine, Jolivette, and Fredrick (2011) Southeast United States	One multiple baseline across three teachers' classes	Positive effects for "on-task" behavior for zero out of three classes (0%) resulting from a coaching and feedback intervention to increase teachers' behavior-specific praise	*Neutral (or no) effect*; effect not demonstrated at least three times	*No evidence*

(continued)

Table 2. (continued)

Study and geographic location	Number and type of eligible experiment(s)	Intervention and effect on students' behavior[a]	Classification of effects for study	
			CEC Standards	WWC Standards
Pisacreta, Tincani, Connell, and Axelrod (2011) Mid-Atlantic United States	One multiple baseline across three teachers' classes	Positive effects for "disruption" for two out of three classes (67%) resulting from modeling and feedback intervention to increase praise-to-behavior correction	*Neutral (or no) effect*; effect not demonstrated at least three times	*No evidence*
Sutherland, Wehby, and Copeland (2000) Southeast United States	One within-case reversal	Positive effect for "on-task" behavior for one out of one *class* (100%) resulting from a feedback intervention to increase teacher's behavior-specific praise	*No classification of effects*; less than three cases	*Strong evidence*; three demonstrations of effects in within-case experiment
Thomas, Becker and Armstrong (1968) Midwest United States	One within-case reversal	Positive effect for "disruptive behavior" for one out of one class (100%) when teacher told to discontinue use of praise statements and to disapprove of the class's problem behaviors	*No classification of effects*; less than three cases	*Strong evidence*; three demonstrations of effects in within-case experiment

Note. CEC = Council for Exceptional Children; WWC = What Works Clearinghouse.
[a]Effects based on visual analysis at the case level.

student cases (70.00%) and positive effects for 8 of 12 classroom-level cases (66.67%) resulting in positive effects for 22 of 32 (68.75%) student- and classroom-level cases combined.

There was agreement in effects classification under both standards for 8 of the 11 studies (72.22%). However, classification of study effects did not align for three studies: Broden et al. (1970), Sutherland et al. (2000), and Thomas, Becker, and Armstrong (1968). These discrepancies occurred because CEC Standards required at least three cases for a study to be considered as having positive effects, whereas WWC Standards allowed for a single within-case reversal design (i.e., one participant) to provide the three demonstrations of effects needed for a *strong* or *moderate evidence* classification. The WWC Standards resulted in the classification of seven studies, which provided moderate or strong evidence for teacher praise whereas the CEC Standards resulted in the classification of four studies demonstrating positive or mixed effects.

Effect size estimation. Per WWC Standards, effect sizes were estimated for the seven studies with *moderate* or *strong evidence* (Allday et al., 2012; Blaze et al., 2014; Broden et al., 1970; Duncan et al., 2013; Rathel et al., 2014; Sutherland et al., 2000; Thomas et al., 1968). For appropriate behavior, average PND was 84.40% (*SD* = 18.23, range = 57.14%–100%), considered *effective* (Scruggs & Mastropieri, 1998) on average. Average Tau-*U* was 0.79 (*SD* = 0.24, range = 0.42–1) across all experiments for appropriate behavior (baseline corrected for baseline trend for one case; Parker et al., 2011). For student-level experiments, appropriate behavior average PND was 66.97% (*SD* = 11.79, range = 57.14%–82.94%), and average Tau-*U* was 0.68 (*SD* = 0.17, range = 0.48–0.90). For class-level experiments, appropriate behavior average PND was 98.33% (*SD* = 3.73, range = 91.67%–100%), and average Tau-*U* was 0.88 (*SD* = 0.26, range = 0.42–1). For disruptive behavior, average PND was 73.63% (*SD* = 21.32, range = 43.75%–100%), also considered *effective* (Scruggs & Mastropieri, 1998) on average, and average Tau-*U* was |0.90| (*SD* = 0.20) ranging from |0.68| to |1.17| (corrected baseline for two cases). Only one study (Duncan et al., 2013) examined effects for disruptive behavior at the student level (PND = 71.96%, Tau-*U* = |0.74|). For class-level experiments, disruptive behavior average PND was 73.96% (*SD* = 23.82, range = 43.75%–100%), and average Tau-*U* was |0.92| (*SD* = 0.21, range = |0.68|–|1.17|).

Overall Evidence-Based Classification

CEC Standards. Using CEC Standards to determine the overall evidence-based classification, there is *mixed evidence* for teacher praise for the population of students identified in this review. There were three methodologically sound studies with positive effects, and the ratio of methodologically sound studies with positive effects (based on visual analysis) to methodologically sound studies with neutral or mixed effects is less than 2:1 (here there was a 3:4 ratio of positive to neutral/mixed effects). Three studies received no classification under CEC standards because the studies had less than three cases (see Table 2).

WWC Standards. Per WWC Standards, teacher praise for the specified student population *does not meet the WWC Single-Case Design Standards Panel's recommended threshold* for combining results of studies to identify evidence-based practices. This threshold specifies the need for a minimum of five studies that meet evidence standards or meet evidence standards with reservations, conducted by at least three different research teams at three different locations, and the combined number of experiments across studies total at least 20. This review identified 11 studies that meet evidence standards or meet evidence standards with reservations (7 provide strong or moderate evidence). The studies represent at least three different research teams at three different locations. But, the total number of experiments providing strong or moderate evidence was 11 experiments (see Table 2).

Summary of Evidence Base

Summaries of study, student, intervention characteristics, teachers' perceptions of social validity, and effects are presented for student-level and class-level cases in Appendices A and B. Eligible studies were conducted in elementary through high school classrooms, in general and special education classrooms, and with male and female students. However, a review of student, setting, and intervention characteristics as well as teachers' perceptions of social validity provided no discernable patterns for when and for whom teacher praise was effective. Only one study (Duncan et al., 2013) reported hypothesized function of students' problem behavior (at least partially maintained by attention). When reported, teacher perceptions of social validity were generally positive. But, in two studies (Duncan et al., 2013; Starkweather Matheson & Shriver, 2005), two teachers indicated praise was not appropriate for their classrooms.

Discussion

In their introduction of CEC Standards, the CEC (2014) acknowledge that requiring studies to meet all methodological quality indicators "will necessarily limit the consideration of studies conducted before quality indicators were developed and emphasized in published studies" (p. 2). This is certainly true for the body of research on teacher praise for students without severe disabilities: Several older and often-cited studies eligible for inclusion in this review were

rated as not meeting methodological quality standards (both CEC and WWC Standards). These studies were eliminated from further consideration when classifying the evidence base for teacher praise.

Only 11 studies were rated as being methodologically sound using each set of standards. The review team classified study-level effects for the 11 methodologically sound studies and arrived at different results when using CEC Standards compared with WWC Standards because CEC Standards required at least three cases for a study to be considered as having positive effects. In contrast, WWC Standards allow for a single within-case reversal design to provide three demonstrations of effects. Despite these differences, the results of the overall evidence-based classifications using CEC and WWC Standards were effectively the same: There is currently insufficient evidence to identify teacher praise as an evidence-based practice for students without severe disabilities in K–12 classroom settings. One of the goals of this review was to provide a comprehensive summary of the evidence base for teacher praise to include information of when and for whom teacher praise was effective. However, we cannot yet draw conclusions about when and for whom teacher praise is likely to be effective.

The results of this review are not unlike results of similar systematic reviews of evidence base supporting other behavioral intervention practices for classroom settings for students without severe disabilities. For example, Maggin and colleagues conducted reviews of the research base on token economies (Maggin, Chafouleas, Goddard, & Johnson, 2011) and group contingencies (Maggin, Johnson, Chafouleas, Ruberto, & Berggren, 2012) to improve classroom behavior of students without severe disabilities. In both systematic reviews, many older studies were identified as not being methodologically sound for similar reasons, such as not demonstrating adequate levels of interobserver agreement and not including enough data points within a phase to demonstrate a clear pattern of student behavior. And, researchers conducting these two systematic reviews were unable to draw strong conclusions about when and for whom the interventions were likely to be effective.

Limitations

One limitation of this review is that the review team may have inadvertently missed one or more studies during the search process. We did not conduct reliability checks on all studies included in our search process. Second, only studies published in peer-reviewed journals were included in this review. Dissertations were not included in this review, which may have led to the exclusion of additional methodologically sound studies on teacher praise. And, this research base may be influenced by publication bias (see Cook, 2014), meaning researchers are less likely to submit for publication results of studies with negative or neutral effects and journals are less likely to publish negative findings. It is very possible that we have not included findings from unpublished or yet-to-be-published studies that might change the overall classification of teacher praise. Third, as an a priori decision, we did not include studies conducted with preschool students nor did we include student participants with severe disabilities or autism. Had we included either or both of these groups of students in our target populations, we may have arrived at different overall conclusions. Fourth, because of the small number of studies included, we did not parse out research on general praise from research on behavior-specific praise; there could be differences in effectiveness and in the overall evidence across the two types of praise. Fifth, to calculate effect sizes, the review team extracted data from images of graphed data from the research reports; we did not use the original data files. It is possible that some points were not exact, which may have influenced effect size calculations. Finally, we only calculated PND and Tau-*U* as indices of effect size because these are two that are frequently reported; other effect size indices (e.g., using multilevel analyses) could provide additional information.

Implications and Recommendations for Practice

The results of this review are important for teachers and their instructional decisions and for school leaders. Most important, the results of this review indicate, at present, there is not yet enough high-quality research to deem teacher praise as an evidence-based practice for students without severe disabilities in K–12 classroom settings. Nevertheless, teacher praise is still a *research-based practice* for this population of students. We hope the results of this review highlight this distinction for teachers and school leaders. Although teacher praise is supported by a history of more than 50 years of empirical research, many older and often-cited studies used research designs that do not meet contemporary standards for demonstrating effects (e.g., AB or pre- and posttest only designs without a control group) or other contemporary methodological standards. Furthermore, at present, we do not yet know enough to make specific recommendations about for whom and under what conditions teacher praise is more or less likely to be an effective intervention approach. We hope the results of this review enhance practitioners' understanding of the limitations of the existing research base on teacher praise.

Although disheartening, these results highlight some positive implications for teachers and school leaders. First, researchers demonstrated the effectiveness of teacher praise in lower and upper grade levels, for whole classes and individual students, for male and female students, in general and special education classrooms, and to increase appropriate behavior (e.g., on-task) as well as decrease disruptive

behavior. Results indicated teacher praise was effective for almost 70% of the 32 cases summarized in this review, and average effect sizes indicate teacher praise is an effective intervention option. Furthermore, when reported, teachers generally indicated positive perceptions about using praise in their classrooms. Finally, across methodologically sound studies, there were no cases with reported negative or non-therapeutic effects. These results suggest the broad applicability of teacher praise to address common classroom problems. For these reasons, we recommend teachers continue to use teacher praise as a preliminary approach for preventing and reducing problem behaviors, particularly because it is a low-intensity and easy-to-implement strategy that is likely an important component of a teachers' instructional repertoire.

These results also highlight some less positive implications. First, there were studies reporting mixed or neutral effects of teacher praise on student behavior at both the student and class levels. In addition, although most teachers reported positive perceptions about using praise in their classrooms, two teachers in two different studies (Duncan et al., 2013; Starkweather Matheson & Shriver, 2005) indicated teacher praise was not an appropriate intervention approach for their students. Furthermore, based on the current evidence across the body of research on teacher praise, we do not yet know when praise is most likely to be effective. These findings are relevant for teachers because they suggest that teacher praise may not be effective enough to prompt meaningful improvements in student behavior or be socially valid in all situations. For some situations, more comprehensive approaches will likely be necessary where increased teacher praise is one component of a class-wide or individual behavior management plan. These findings are also important for school leaders and behavior support teams working with teachers to prevent and reduce problem behaviors. Until more is known about conditions when teacher praise is most likely to be effective, teachers and school leaders should monitor for potential idiosyncratic differences in effectiveness for different students and classes.

Implications for Future Research and Leadership in Special Education

The results of this study can inform future methodologically sound intervention research to add to the evidence base for teacher praise for students without severe disabilities who exhibit problem behaviors in classroom settings. The most obvious implication is that more research is needed to provide additional demonstrations of the effectiveness of teacher praise to increase its evidence base for this population. We hope the results of this review provide impetus for continuing research in this area. We also hope these results prompt researchers and leaders in education and special education (e.g., agency and organization representatives as well as preparers of future teachers) to accurately label teacher praise as a research-based practice rather than an evidence-based practice for students without severe disabilities in K–12 classrooms, especially when communicating with practitioners.

We hope the results of this study prompt additional conversation in our field about the overall implications of systematic evidence-based reviews. Important questions come to light, such as the following: Should researchers allocate more time and resources in systematic replications with the sole purpose of expanding the evidence base of effective practices? Will replication studies be valued by others, even if it requires building on much older research, as in the case for teacher praise research? Will scholarly journals devote space to publish such replication studies to assist in building research bases? If we encourage and publish replication work, will we inadvertently constrain research on novel intervention approaches? Are we unnecessarily discrediting years of educational research, particularly for research on behavioral interventions, which typically involve direct observations in the natural setting? In our interpretation of the standards, we required researchers to report information demonstrating assessment of interobserver agreement for at least 20% of sessions across all experimental conditions. We feel justified in this decision based on common recommendations (e.g., Ayers & Gast, 2010; Kazdin, 2011; Kratochwill et al., 2010), and moving forward, we recommend researchers using direct observations in applied settings adhere to this standard. And, how can we improve communication with practitioners and policy makers about research-based and evidence-based practices? We encourage faculty who train future educational researchers to pose these and related questions to their students as well.

Results of our review indicate there is currently insufficient evidence to identify teacher praise as an evidence-based practice under either set of standards. However, it is possible that reviews on other practices using CEC Standards versus WWC Standards could result in different overall evidence-based classifications. This possibility should be considered by educational agencies, organizations, and researchers as they continue to work on identifying practices that are evidence-based. We hope that our application of both CEC Standards and WWC Standards prompts discussion among leaders in our field about, and potential reconciliation of, discrepancies between the two sets of standards for single-subject research methodologies.

In addition, researchers should develop systematic lines of research to broaden our understanding of when and for whom teacher praise is effective. For example, it is reasonable to predict that teacher praise is more likely to be effective for students who have attention-seeking problem behaviors. More research is needed to examine further the influence of behavioral functions when predicting

effectiveness of teacher praise. Also, researchers should examine effects of teacher praise across grade levels and student ages, across different types of classroom settings, and for quiet versus loud praise. Additional research topics might include further examination of different aspects of teacher praise (e.g., general vs. behavior-specific praise, qualitative aspects, different schedules of reinforcement), comparisons between or combinations of teacher praise and other types of intervention approaches (e.g., precorrection, rules, differential reinforcement), or additional outcomes that might be associated with increased praise (e.g., improved classroom environment, improved teacher–student relationships, improved teacher efficacy, or improved student perceptions about school).

Finally, we encourage researchers to continue to examine issues related to social validity of teachers praise. Specifically, this might include qualitative and quantitative approaches to learning more about teacher and student preferences, teachers' ability and willingness to implement teacher praise with fidelity, and teachers' sustained use of teacher praise over time. The results of this review indicate teachers generally reported positive perceptions of teacher praise across studies, with two exceptions (Duncan et al., 2013; Starkweather Matheson & Shriver, 2005). However, many studies included in this review included intensive supports, such as coaching or performance feedback, to help teachers improve their use of praise. We applaud this work and encourage others to continue research to identify barriers to and supports for teachers' sustained adoption of praise. We recognize the difficulties in conducting research in classroom settings, particularly when we are expecting teachers to implement and sustain interventions. We acknowledge teachers' and others' concerns about teacher praise (e.g., reports of teacher perceptions reported in studies in this review and Kohn, 2001). We hope this review will prompt additional attention to these concerns to help us identify when teacher praise is most likely to be adopted and effective.

Appendix A

Table A1. Study, Student, Intervention Characteristics, Social Validity, and Effects for Eligible Student-Level Cases.

Case; disability status	Problem behavior and function	Grade and gender	Setting	Support provided to teacher	Effect on student behavior[a]	Teachers' perception of social validity
Allday et al. (2012)						
Paul; EBD[b]	Inattentive, aggressive, noncompliant; function not reported	Second; male	Gen ed[c]	Training; performance feedback to increase behavior-specific praise to all students in class	*Positive effect*	Positive perception; noted it was difficult to increase behavior-specific praise
Jack; at risk for EBD	Inattentive, noncompliant, disruptive; function not reported	K[d]; male	Gen ed		*Positive effect*	
Jill; at risk for EBD	Inattentive, not following rules, somatic complaints; function not reported	K; female	Gen ed		*Positive effect*	
Tom; at risk for EBD	Inattentive, disruptive; function not reported	First; male	Gen ed		*No effect*	
Bill; at risk for EBD	Inattentive, disruptive; function not reported	First; male	Gen ed		*No effect*	
Kyle; EBD	Off-task, disruptive; function not reported	Sixth; male	Gen ed		*Positive effect*	
Chris; EBD	Off-task, disruptive; function not reported	Sixth; male	Gen ed		*No effect*	
Broden, Bruce, Mitchell, Carter, and Hall (1970)						
Edwin; none reported	Often engaged with another disruptive student; function not reported	Second; male	Gen ed	None reported; praise provided for individual student	*Positive effect*	None reported
Greg; mild intellectual disability	Disruptive and nonstudy behaviors, talking out, talking to peers, out of seat, and so forth	Second; male	Gen ed		*Positive effect*	None reported
Duncan et al. (2013)[e]						
Thomas; no disability	Off-task; inappropriate vocalizations; leaving seat; attention hypothesized as at least partial function	Fifth; male	Gen ed	Training; Goal setting and performance feedback to increase behavior-specific praise to target student	*Positive effect*	Positive perception prior to intervention
Bobby; no disability	Off-task; inappropriate vocalizations; leaving seat; attention hypothesized as at least partial function	Fifth; male	Gen ed		*Positive effect*	Positive perception prior to intervention, but after intervention, teacher stated she did not think praise was appropriate for student and that it decreased her attention to other students

(continued)

Table A1. (continued)

Case; disability status	Problem behavior and function	Grade and gender	Setting	Support provided to teacher	Effect on student behavior[a]	Teachers' perception of social validity
Rathel, Drasgow, Brown, and Marshall (2014)[f]						
Ms. Walker's student; EBD	Off-task	6-year-old; male	Spec ed[g]	Training; Performance feedback to increase behavior-specific praise and positive-to-negative communication ratio to class	*Positive effect*	None reported
Ms. Carroll's student; EBD	Off-task	10-year-old; female	Spec ed		*Positive effect*	Positive perception
Ms. Bishop's student; LD	Off-task	13-year-old; male	Sped ed		*Positive effect*	None reported
Starkweather Matheson and Shriver (2005)						
Cory; no disability	Noncompliance; function not reported	Second; male	Gen ed	Training; performance feedback to increase praise to individual student	*No effect*	Positive perception
Nate; no disability	Noncompliance; function not reported	Second; male	Gen ed		*Positive effect*	Positive perception
Andy; no disability	Noncompliance; function not reported	Fourth; male	Gen ed		*Positive effect*	Negative perception; felt uncomfortable praising fourth-grade students
Thompson, Marchant, Anderson, Prater, and Gibb (2012)						
Anna's student; disability status unclear	Noncompliance and disruptive behavior	Fourth; male	Gen ed	Tiered training (basic training, self-monitoring, coaching) to increase teachers' behavior-specific praise to class	*No effect*	Positive perception
Gail's student; disability status unclear	Noncompliance and disruptive behavior	Fourth; male	Gen ed		*Positive effect*	
Jane's student; disability status unclear	Noncompliance and disruptive behavior	Third; male	Gen ed		*No effect*	

Note. EBD = emotional and behavioral disorder; LD = learning disability.
[a]Effects based on visual analysis at the case level. [b]Emotional or behavioral disorders. [c]General education classroom. [d]Kindergarten. [e]Student John not eligible for this review (preschool setting) but positive effect demonstrated based on visual analysis. [f]Ms. Bryant's student not eligible for this review (autism) but positive effect demonstrated based on visual analysis. [g]Special education classroom.

Appendix B

Table B1. Study, Student, and Intervention Characteristics, Social Validity, and Effects for Classroom-Level Cases.

Case	Class behavior	Grade or age; gender	Setting	Support provided to teacher	Effect on student behavior[a]	Teachers' perception of social validity
Blaze, Olmi, Mercer, Dufrene, and Tingstom (2014)						
Teacher 1	Disruptive	Ninth to 11th; 44% female	Gen ed[b]	Training to increase quiet/loud praise to the class	*Positive effect*	Positive perception; teacher preferred quiet praise
Teacher 2	Disruptive	Ninth to 10th; 38% female	Gen ed		*Positive effect*	Positive perception; teacher preferred loud praise
Teacher 3	Disruptive	Ninth to 10th; 58% female	Gen ed		*Positive effect*	Positive perception; teacher preferred quiet praise
Teacher 4	Disruptive	Ninth to 10th; 48% female	Gen ed		*Positive effect*	Positive perception; teacher preferred loud praise
Duchaine, Jolivette, and Fredrick (2011)						
Kelly's class	Off-task	Ninth; 45% female	Gen ed	Training; Performance feedback to increase behavior-specific praise to class	*No effect*	Positive perception
Jamie's class	Off-task	Ninth; 50% female	Gen ed		*No effect*	
Chris's class	Off-task	Ninth; 39% female	Gen ed		*No effect*	
Pisacreta, Tincani, Connell, and Axelrod (2011)						
Teacher 1	Disruptive	Sixth; gender not reported	Gen ed	Training; Modeling and performance feedback to increase praise-to-behavior correction ratio to class	*Positive effect*	None reported
Teacher 2	Disruptive	Eighth; gender not reported	Gen ed		*Positive effect*	
Teacher 3	Disruptive	Seventh; gender not reported	Gen ed		*No effect*	

(continued)

Table B1. (continued)

Case	Class behavior	Grade or age; gender	Setting	Support provided to teacher	Effect on student behavior[a]	Teachers' perception of social validity
Sutherland, Wehby, and Copeland (2000)						
Class	Off-task	10- to 11-year-olds; 22% female	Spec ed[c] classroom for students with EBD[d]	Training; Performance feedback to increase behavior-specific praise to the class	*Positive effect*	None reported
Thomas, Becker, and Armstrong (1968)						
Class	Class well behaved	6- to 7-year-olds; gender not reported	Gen ed	None reported; praise provided to the class	*Positive effect*	None reported

Note. EBD = emotional and behavioral disorder.
[a]Effects based on visual analysis at the case level. [b]General education classroom. [c]Special education classroom. [d]Emotional or behavioral disorders.

Declaration of Conflicting Interests

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

Funding

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

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Allday, R. A., Hinkson-Lee, K., Hudson, T., Neilsen-Gatti, S., Kleinke, A., & Russel, C. S. (2012). Training general educators to increase behavior-specific praise: Effects on students with EBD. *Behavioral Disorders*, *37*, 87–98.

Ayers, K., & Gast, D. L. (2010). Dependent measures and measurement procedures. In D. L. Gast (Ed.), *Single subject research methodology in behavioral sciences* (pp. 129–165). New York, NY: Routledge.

*Blaze, J. J., Olmi, D. J., Mercer, S. H., Dufrene, B. A., & Tingstom, D. H. (2014). Loud versus quiet praise: A direct behavioral comparison in secondary classrooms. *Journal of School Psychology*, *52*, 349–360. doi:10.1016/j.jsp.2014.05.006

*Broden, M., Bruce, C., Mitchell, M. A., Carter, V., & Hall, R. V. (1970). Effects of teacher attention on attending behavior of two boys at adjacent desks. *Journal of Applied Behavior Analysis*, *3*, 205–211. doi:10.1901/jaba.1970.3-205

Brophy, J. (1981). On praising effectively. *The Elementary School Journal*, *81*, 268–278. Retrieved from http://www.jstore.org/stable/1001606

Cavanaugh, B. (2013). Performance feedback and teachers' use of praise and opportunities to respond: A review of the literature. *Education and Treatment of Children*, *36*, 111–136. doi:10.1353/etc.2013.0001

Conroy, M. A., Sutherland, K. S., Snyder, A. L., & Marsh, S. (2008). Classwide interventions: Effective instruction makes a difference. *Teaching Exceptional Children*, *40*, 24–30. doi:10.1177/004005990804000603

Cook, B. G. (2014). A call for examining replication and bias in special education research. *Remedial and Special Education*, *35*, 233–246. doi:10.1177/0741932514528995

*Cossairt, A., Hall, R. V., & Hopkins, B. L. (1973). The effects of experimenter's instructions, feedback, and praise on teacher praise and student attending behavior. *Journal of Applied Behavior Analysis*, *6*, 89–100. doi:10.1901/jaba.1973.6-89

Council for Exceptional Children. (2014). *Council for Exceptional Children standards for evidence-based practices in special education*. Retrieved from http://www.cec.sped.org/~/media/Files/Standards/Evidence%20based%20Practices%20and%20Practice/EBP%20FINAL.pdf

*Duchaine, E. L., Jolivette, K., & Fredrick, L. D. (2011). The effect of teacher coaching with performance feedback on behavior-specific praise in inclusion classrooms. *Education and Treatment of Children*, *34*, 209–227.

*Dufrene, B. A., Lestremau, L., & Zoder-Martell, K. (2014). Direct behavioral consultation: Effects on teachers' praise and student disruptive behavior. *Psychology in the Schools*, *51*, 567–580. doi:10.1002/pits

*Duncan, N. G., Dufrene, B. A., Sterling, H. E., & Tingstrom, D. H. (2013). Promoting teachers' generalization of intervention use through goal setting and performance feedback. *Journal of Behavioral Education*, *22*, 325–347. doi:10.1007/s10864-013-9173-5

Epstein, M., Atkins, M., Cullinan, D., Kutash, K., & Weaver, R. (2008). *Reducing behavior problems in the elementary school classroom. IES practice guide* (NCEE 2008-012). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/wwc/publications/practiceguides

*Ferguson, E., & Houghton, S. (1992). The effects of contingent teacher praise, as specified by Canter's assertive discipline programme, on children's on-task behaviour. *Educational Studies*, *18*, 83–93.

Gable, R. A., Hester, P. H., Rock, M. L., & Hughes, K. G. (2009). Back to basics: Rules, praise, ignoring, and reprimands revisited. *Intervention in School and Clinic*, *44*, 195–205. doi:10.1177/1053451208328831

Gast, D. L., & Spriggs, A. D. (2010). Visual analysis of graphic data. In D. L. Gast (Ed.), *Single subject research methodology in behavioral sciences* (pp. 129–165). New York, NY: Routledge.

*Hall, R. V., Fox, R., Willard, D., Goldsmith, L., Emerson, M., Owen, M., . . . Porcia, E. (1971). The teacher as observer and

experimenter in the modification of disputing and talking-out behaviors. *Journal of Applied Behavior Analysis*, *4*, 141–149. doi:10.1901/jaba.1971.4-141

*Hall, R. V., Lund, D., & Jackson, D. (1968). Effects of teacher attention on study behavior. *Journal of Applied Behavior Analysis*, *1*, 1–12. doi:10.1901/jaba.1968.1-1

*Hall, R. V., Panyan, M., Rabon, D., & Broden, M. (1968). Instructing beginning teachers in reinforcement procedures which improve classroom control. *Journal of Applied Behavior Analysis*, *1*, 315–322. doi:10.1901/jaba.1968.1-315

Hasazi, J. E., & Hasazi, S. E. (1972). Effects of teacher attention on digit-reversal behavior in an elementary school child. *Journal of Applied Behavior Analysis*, *5*, 157–162. doi:10.1901/jaba.1972.5-157

*Haydon, T., & Musti-Rao, S. (2011). Effective use of behavior-specific praise: A middle school case study. *Beyond Behavior*, *20*(2), 31–39.

*Hollingshead, A., Kroeger, S. D., Altus, J., & Brubaker Trytten, J. (2016). A case study of positive behavior supports-based interventions in a seventh-grade urban classroom. *Preventing School Failure*, *60*, 278–285. doi:10.1080/1045988X.2015.1124832

*Houghton, S., Wheldall, K., Jukes, R., & Sharpe, A. (1990). The effects of limited private reprimands and increased private praise on classroom behavior in four British secondary school classes. *British Journal of Educational Psychology*, *60*, 255–265. doi:10.1111/j.2044-8279.1990.tb00943.x

Jenkins, L. N., Floress, M. T., & Reinke, W. (2015). Rates and types of teacher praise: A review and future directions. *Psychology in the Schools*, *52*, 463–476. doi:10.1002/pits.21835

Kazdin, A. E. (2011). *Single-case research designs* (2nd ed.). New York, NY: Oxford University Press.

*Kirby, F. D., & Shields, F. (1972). Modification of arithmetic response rate and attending behavior in a seventh-grade student. *Journal of Applied Behavior Analysis*, *5*, 79–84. doi:10.1901/jaba.1972.5-79

Kohn, A. (2001). Five reasons to stop saying, "good job!" *Young Children*, *56*(5), 24–30.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from https://ies.ed.gov/ncee/wwc/Document/229

Kretlow, A. G., & Blatz, S. L. (2011). The A B Cs of evidence-based practice for teachers. *Teaching Exceptional Children*, *43*, 8–19. doi:10.1177/004005991104300501

Lane, K. L., Menzies, H. M., Ennis, R. P., & Oakes, W. P. (2015). *Supporting behavior for school success: A step-by-step guide to key strategies*. New York, NY: Guilford Press.

Lewis, T. J., Hudson, S., Richter, M., & Johnson, N. (2004). Scientifically supported practices in emotional and behavioral disorders: A proposed approach and brief review of current practices. *Behavioral Disorders*, *29*, 247–259. Retrieved from http://www.jstore.org/stable/23889473

*Madsen, C. H., Jr., Becker, W. C., & Thomas, D. R. (1968). Rules, praise, and ignoring: Elements of elementary classroom control. *Journal of Applied Behavior Analysis*, *1*, 139–150. doi:10.1901/jaba.1968.1-139

Maggin, D. M., Chafouleas, S. M., Goddard, K. M., & Johnson, A. H. (2011). A systematic evaluation of token economies as a classroom management tool for students with challenging behavior. *Journal of School Psychology*, *49*, 529–554. doi:10.1016/j.jsp.2011.05.001

Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students with challenging behavior. *Journal of School Psychology*, *50*, 625–654. doi:10.1016/j.jsp.2012.06.001

*Martens, B. K., Hiralall, A. S., & Bradley, T. A. (1997). A note to teacher: Improving student behavior through goal setting and feedback. *School Psychology Quarterly*, *12*, 33–41. doi:10.1037/h0088945

*McAllister, L. W., Stachowiak, J. G., Baer, D. M., & Conderman, L. (1969). The application of operant conditioning techniques in a secondary school classroom. *Journal of Applied Behavior Analysis*, *2*, 277–285. doi:10.1901/jaba.1969.2-277

*Mesa, J., Lewis-Palmer, T., & Reinke, W. (2005). Providing teachers with performance feedback on praise to reduce student problem behavior. *Beyond Behavior*, *15*(1), 3–7.

*Myers, D. M., Simonsen, B., & Sugai, G. (2011). Increasing teachers' use of praise with a response-to-intervention approach. *Education and Treatment of Children*, *34*, 35–59. doi:10.1353/etc.2011.0004

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*, 284–299. doi:10.1016/j.beth.2010.08.006

*Pisacreta, J., Tincani, M., Connell, J. E., & Axelrod, S. (2011). Increasing teachers' use of a 1:1 praise-to-behavior correction ratio to decrease student disruption in general education classrooms. *Behavioral Interventions*, *26*, 243–260. doi:10.1002/bin.341

*Rathel, J. M., Drasgow, E., Brown, W. H., & Marshall, K. J. (2014). Increasing induction-level teachers' positive-to-negative communication ratio and use of behavior-specific praise through e-mailed performance feedback and its effect on students' task engagement. *Journal of Positive Behavior Interventions*, *16*, 219–233. doi:10.1177/1098300713492856

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior Modification*, *22*, 221–242.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education*, *8*, 24–33.

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children*, *31*, 351–380. doi:10.1353/etc.0.0007

*Starkweather Matheson, A. S., & Shriver, M. D. (2005). Training teachers to give effective commands: Effects on student compliance and academic behaviors. *School Psychology Review*, *34*, 202–219.

Sutherland, K. S. (2000). Promoting positive interactions between teachers and students with emotional/behavioral disorders. *Preventing School Failure*, *44*, 110–115. doi:10.1080/10459880009599792

Sutherland, K. S., & Wehby, J. H. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with

emotional and behavioral disorders. *The Journal of Special Education*, *35*, 161–171. doi:10.1177/002246690103500306

*Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders*, *8*, 2–8. doi:10.1177/106342660000800101

Sweigart, C. A., Collins, L. W., Evanovich, L. L., & Cook, S. C. (2016). An evaluation of the evidence base for performance feedback to improve teacher praise using CEC's quality indicators. *Education and Treatment of Children*, *39*, 419–444. doi:10.1353/edt.2016.0019

*Thomas, D. R., Becker, W. C., & Armstrong, M. (1968). Production and elimination of disruptive classroom behavior by systematically varying teacher's behavior. *Journal of Applied Behavior Analysis*, *1*, 35–45. doi:10.1901/jaba.1968.1-35

*Thompson, M. T., Marchant, M., Anderson, D., Prater, M. A., & Gibb, G. (2012). Effects of tiered training on general educators' use of specific praise. *Education and Treatment of Children*, *35*, 521–546. Retrieved from https://muse.jhu.edu/article/487097

*Trolinder, D. M., Choi, H., & Proctor, T. B. (2004). Use of delayed praise as a directive and its effectiveness on on-task behavior. *Journal of Applied School Psychology*, *20*, 61–83. doi:10.1300/J370v20n02_05

Vannest, K. J., Parker, R. I., & Gonen, O. (2011). Single case research: Web based calculators for SCR analysis (Version 1.0) [Web-based application]. College Station: Texas A&M University. Available from singecaseresearch.org

*Wasik, B. H., Senn, K., Welch, R. H., & Cooper, B. R. (1969). Behavior modification with culturally deprived school children: Two case studies. *Journal of Applied Behavior Analysis*, *2*, 181–194. doi:10.1901/jaba.1969.2-181

*Wood, B. K., Umbreit, J., Liaupsin, C. J., & Gresham, F. M. (2007). A treatment integrity analysis of function-based intervention. *Education and Treatment of Children*, *30*, 105–120.

*Workman, E. A., Kindall, L. M., & Williams, R. L. (1980). The consultative merits of praise-reward versus praise-reprimand instruction. *The Journal of School Psychology*, *18*, 373–380. doi:10.1016/0022-4405(80)90009-6