5-14-2023

Contingencies in Classrooms: A Systematic Review of Contingent Reinforcement for Academic Performance

Taylor Jean Hitchings
Syracuse University

Abstract

Contingency-based reinforcement interventions, in which students receive a reward for engaging in desired behaviors, have been used to target various behaviors at school. Previous research studies' findings suggest that contingency-based reinforcement interventions can be used to target a variety of behaviors, including task engagement, appropriate classroom etiquette, and academic performance across a variety of subjects. The purpose of this systematic review was to synthesize the assessment and intervention practices of 98 studies (from 1969 to 2021) that examined the effects of contingency-based reinforcement interventions when academic performance was directly targeted in the areas of math, reading, writing, spelling, and English Language Arts. Specifically, the state of the literature's use of relevant assessment practices (i.e., Can't Do/Won't Do to identify skill versus performance deficits, preference, and reinforcer assessments) was reviewed, as was the use of relevant intervention practices (i.e., contingency administration formats, contingency criteria selection, randomized and unknown intervention components, and academic subjects targeted). Overall, this review reported the percentage of studies that used Can't Do/Won't Do, preference, and reinforcer assessments, as well as the percentage that implemented the intervention in individual or group formats, used particular methods to select reinforcement criteria, included randomized or known contingency components, and targeted specific academic subjects. Descriptive information was also provided regarding the general characteristics of the studies that met inclusion criteria, as well as participant characteristics and demographics. The results of this study may be used to inform subsequent research, meta-analyses, and intervention implementation.

Keywords: contingent reinforcement, academic performance, systematic review

CONTINGENCIES IN CLASSROOMS: A SYSTEMATIC REVIEW OF CONTINGENT REINFORCEMENT FOR ACADEMIC PERFORMANCE

By

Taylor J. Hitchings

B. A., University of Colorado Boulder, 2017

Thesis

Submitted in partial fulfillment of the requirements for the degree of

Master of Science in Psychology

Syracuse University

May 2023

Acknowledgements

I want to thank everyone who has contributed to this project and supported me throughout this process.

First, thank you to my committee members—Drs. Natalie Russo, Katie Kidwell, and Meredith Martin—for your insight.

This project was also completed under two advisors: Drs. Bridget Hier and Tanya Eckert. Bridget, I want to thank you for all of your investment and time. This project is stronger because you were a part of it. Tanya, thank you for helping me to get to the finish line. Thank you for your flexibility, kindness, and support throughout my time at SU.

I also want to thank the Association of Positive Behavior Support for their assistance funding this systematic review.

To Sam and Kaytlin, who served as secondary coders: You two are the best. Thank for giving up some of your precious free time (during the summer, too!) to help.

This project is ultimately dedicated to my loved ones and those who have supported me throughout grad school and life. Dad, Mom, Kris, AJ, Jess, Matt, Eli, Emmy, Michelle, and Amanda. I love you all very, very much. Shout out to Kathy, as well, for of all of the ways that she's helped me grow over the last year. And, of course, to Dally and Dewey—the best fluffy little guys anyone could have ever asked for. Even though you were both extremely distracting whenever I tried to work on this thesis.

Funding Statement

This study was partially funded through a student research grant from The Association of Positive Behavior Support (APBS) under Dr. Bridget Hier.

TABLE OF CONTENTS

List of Tables

List of Figures

Contingencies in Classrooms: A Systematic Review of Contingent Reinforcement for Academic Performance

Current data suggest that students in the United States are academically underperforming across a range of grades and academic subjects (Hussar et al., 2020). Within the academic intervention literature, contingency-based reinforcement interventions are one strategy that has been implemented, in which researchers examine the impact of rewards on students' academic performances (e.g., Chaffee et al., 2020; Duhon et al., 2004; Eckert et al., 2002; Scott et al., 2017). Although myriad studies have examined the relationship between rewards and academic behaviors, this literature has not yet been comprehensively synthesized to provide a holistic view of this literature base and the way these interventions have been implemented. Thus, this systematic review will consolidate the current literature to identify the assessment and intervention components that are utilized within contingency-based reinforcement academic interventions targeting skills in math, reading, writing, and spelling. This review will first discuss the use of Can't Do/Won't Do assessments to inform intervention selection, alongside preference and reinforcer assessments that may be administered to determine the rewards that students will receive for their academic performances. The methods of contingency-based reinforcement interventions will also be examined, with a review of administration formats, contingency criteria, and targeted academic behaviors. Collectively, this information may allow practitioners and researchers to recognize the empirical bases for their assessment and intervention decisions, highlight current gaps in the literature, and facilitate future meta-analyses on contingency-based reinforcement interventions that have targeted academic performance.

Academic Performance in the United States

Past research indicates that students' academic achievement outcomes have important ramifications for their futures. Proficiency with academic skills predicts performance in subsequent classes during students' initial education (Duncan et al., 2007), as well as college preparedness, admissions, and graduation rates (Allensworth & Clark, 2020; Fields, 2014). Vocationally, academic achievement and educational attainment have been shown to correlate with individuals' employability and salaries (National Commission on Writing, 2003, 2005; Roth & Clarke, 1998). From a more holistic perspective, students' academic achievement has also been shown to correlate with overall life expectancy, the likelihood of engaging in risky behaviors like smoking and teenage pregnancy, and time of onset for chronic disease (Fiscella & Kitzman, 2009). Although causality cannot be assumed based on the relationship between academic achievement and these outcomes, research has shown that an individual's education and academic achievement predict health outcomes even after controlling for socioeconomic factors like occupational status, income, and wealth (Fiscella & Kitzman, 2009). As these data indicate, students' academic achievement can impact various outcomes throughout their lifetimes.

One national measure of American students' academic skills is the National Assessment of Educational Progress (NAEP). NAEP provides information about students' performances in a variety of academic subjects on the national, state, and district levels. These data are used to identify academic trends, establish and measure benchmark standards, and compare academic performance based on a range of student and school characteristics (National Center for Education Statistics [NCES], 2018). Each time a NAEP assessment is administered, a nationally representative sample of fourth-, eighth-, and twelfth-grade students are asked to demonstrate their academic skills on a series of standardized measures. Students' skills are assessed in a range of academic areas including reading, math, writing, and science (NCES, 2018). NAEP assessments are the same in every state, and depending upon the academic subject being assessed, they range from 90 to 120 minutes long. Students' skills can be classified as either basic, proficient, or advanced (NCES, 2018). When a student's skills are classified as basic, this indicates that their performance did not demonstrate competency in the subject matter as would be expected for someone at their grade level (NCES, 2020). Scores that are classified as proficient or advanced indicate that the student's performance met or exceeded competency expectations, respectively (NCES, 2020).

The most recent NAEP data indicate that many students in the United States struggle to meet academic expectations across grade levels and subject areas (Hussar et al., 2020). Less than 40% of students in fourth, eighth, and twelfth grade reach competency expectations in reading (Hussar et al., 2020). In math, only one in four seniors demonstrate competency in the subject matter (Hussar et al., 2020). Academic performance in the area of writing is similarly concerning, with only 27 to 28% of students' writing being scored as proficient or above (NCES, 2012a). In fact, students in the United States are academically underperforming in all core subjects—math, reading, writing, geography, science, and U.S. history (Hussar et al., 2020; NCES, 2011, 2015). As these data suggest, many students present significant deficits when their academic performances are assessed. These difficulties are evident on a national level, spanning the mid-elementary years through the twelfth grade. Thus, there is a clear need for evidence-based practices to better address American students' persistent academic deficits.

Skill and Performance Deficits

When students are academically underperforming, this may be caused by a skill or a performance deficit (or, perhaps, a combination of both). This conceptualization of deficits was first applied to social behavior in children (Gresham, 1981), but it has since been broadened to include other behaviors and academic challenges (e.g., Bonfiglio et al., 2004; Duhon et al., 2004; Noell et al., 1997). Broadly, if an individual has a skill deficit, they are unable to perform a particular activity—for example, staying in their seat in the classroom or accurately responding to a set of math questions—because they have insufficient skills to do so (Gansle et al., 2002). When a skill deficit is present, an individual likely requires additional instruction to develop new academic behaviors before their performance will improve (Lentz, 1988). Alternately, a performance deficit occurs when an individual has the prerequisite skills to successfully carry out a particular behavior, but this skill is only demonstrated in a particular environment or under specific circumstances (Gansle et al., 2002). For instance, if a student typically fails to complete their independent seatwork but finishes assignments accurately when told that they must complete the task before they can go to recess, this may indicate a performance deficit. A possible explanation for this variability in academic performance is that the student is motivated to complete independent seatwork when it affects their ability to go to recess, indicating that the specific situational circumstances affect the student's varying behavior across settings.

Although performance deficits generally encompass any factors that are impacting a student's performance beyond their skills, they are often conceptualized more specifically as motivational problems within the field of school psychology (Daly et al., 1997; Lentz, 1988; VanDerHeyden & Witt, 2008). Within the previous literature that has explored the relationship between motivation and academic performance, DiPerna and colleagues (2002, 2005) found a significant indirect effect of student motivation on academic performance. More specifically, the authors found that teachers' reports of primary and intermediate students' motivation indirectly affected their reading ($\beta = .20$ to .25) and math performance ($\beta = .28$ to .38) through its direct impact on study skills ($\beta = .68$ to .75) and engagement ($\beta = .46$ to .58).

While motivation can be conceptualized differently across theoretical orientations, much of the literature in the field of school psychology has examined motivation in a way that is consistent with behavioral psychology, especially within the context of academic performance deficits (e.g., VanDerHeyden & Witt, 2008). From a behavioral perspective, students' motivation can be affected through the manipulation of antecedents and consequences, such that a desired behavior (e.g., independent task completion) becomes more rewarding when it is followed by a desirable consequence (Murphy et al., 2019). Thus, to remain consistent with previous literature in this topic area, the term "motivation" will refer to the assumption that introducing a desirable reward will serve as an incentive to increase students' motivation to perform an academic skill to the best of their ability.

Reinforcement, Incentives, and Rewards

Given the importance of academic motivation for student outcomes, there has been a substantial amount of research examining ways to increase students' motivation. One common strategy involves the use of reinforcement, in which a stimulus is provided after a behavior occurs and increases the likelihood that the same behavior will occur again in the future (Miltenberger, 2016). If a student's behavior increases after a stimulus is presented, this stimulus is classified as a reinforcer for that particular student (Miltenberger, 2016). Common stimuli that may be considered potential reinforcers in schools include incentives and rewards, which may be directly manipulated in classrooms to promote positive academic behaviors or achievement (Stipek, 2004). These incentives and rewards can take a variety of forms, including teacher and peer praise, access to a preferred toy, activity, or edible object, and escape from a non-preferred activity or task (Miltenberger, 2016). Because it is impossible to know whether a reward functions as a reinforcer until subsequent behavior is examined, the term "reward" will be used to refer to stimuli that are intended to serve as reinforcers.

It is important to note, however, that the use of rewards as an intervention component has been a controversial topic in academic literature, with some researchers expressing concerns based on findings that suggest that this strategy may interfere with individuals' intrinsic motivation (Deci et al., 1999; Kohn, 1993). Despite these concerns, many studies that have explored this topic have found that external rewards do not appear to have a significant negative impact on students' intrinsic motivation (Cameron & Pierce, 1994; Cerasoli et al., 2014). Additionally, if a reward does not increase the targeted behavior, it inherently cannot be classified as a reinforcer (Stipek, 2004).

To explore the contradictory nature of results related to rewards and intrinsic motivation, Akin-Little et al. (2004) examined the literature that has found a negative relationship between extrinsic motivation and intrinsic motivation, and the authors assessed these studies' methodologies and offered alternative explanations for their findings. Akin-Little et al. (2004) argued that the occasionally observed deleterious effects of rewards on intrinsic motivation are the result of poor methodology and ineffective programming, in which researchers did not utilize rewards in empirically supported ways. For example, ineffective programming may include too much of a delay between a behavior and reward, a failure to measure baseline performance or use baseline data to inform reinforcement contingencies, or no explicit intervention to support the generalization of behaviors (Akin-Little et al., 2004). Because many intervention studies may fail to include these important components of effective reinforcement, Akin-Little et al. (2004) argued that the “logical solution is not to eliminate programmed reinforcement, but to use effective programmed reinforcement strategies” (p. 359). In this way, it is essential that reinforcement strategies—in all their various forms—are further explored to inform best practices in classrooms and to maximize students’ academic performances. When properly implemented, the use of reinforcement is a versatile strategy that can be used to identify and intervene in performance deficits (Duhon et al., 2004; Skinner et al., 2004; VanDerHeyden & Witt, 2008).

Use of Rewards in Assessment

Rewards can be integrated into assessments to examine a student’s motivation and potentially determine if a student is demonstrating a skill or performance deficit. This assessment practice plays an important preliminary role in intervention planning: It allows educators to better determine the nature of the academic problem, which then informs subsequent intervention (Graplin et al., 2018). A common assessment approach that utilizes rewards in this way is the Can’t Do/Won’t Do assessment (VanDerHeyden, 2008). These assessments are grounded in behavioral theory, as they look at the intraindividual variability in a student’s academic performance based on the presence or absence of a contingent reward. Although Can’t Do/Won’t Do assessment procedures may vary slightly across different implementations, the general method of assessment is largely consistent regardless of administrator or area of academic concern. Per VanDerHeyden and Witt's (2008) guidelines, a baseline sample of academic performance is collected first to confirm academic underperformance, establish performance goals, and demonstrate the student's performance outside of any environmental or instructional changes. Based on the student's initial score, the reinforcement contingency is established. Students are offered a reward if they are able to improve their scores in some way, and they are then administered the academic measure again to determine how their scores may have changed with the introduction of a reward.

Theoretically, a student with a performance deficit will display a significant increase in performance during Can't Do/Won't Do assessments when responding is properly incentivized (i.e., during the contingent reward condition); alternatively, students with a skill deficit will not be able to improve their scores, despite the availability of a reward. This contingent reward condition is particularly relevant to intervention planning, because it serves as a short-term pilot intervention that may later be fully implemented if a performance deficit is identified. After each of these phases (i.e., baselines and contingent reward) are implemented, the student's performance across the two conditions is then compared. Depending upon available time and resources, Can't Do/Won't Do assessments may also involve the use of brief reversals between baseline and reward conditions to demonstrate experimental control over a student's academic responding based on the presence or absence of a reward (VanDerHeyden & Witt, 2008).

Assessments that use Can't Do/Won't Do methodology (or very similar practices) have been shown to effectively identify performance deficits and effective interventions in reading (Gansle & Noell, 2002), math (Duhon et al., 2004), and writing (Duhon et al., 2004). For example, Gansle and Noell (2002) assessed a six-year-old boy's reading skills to identify the presence of a skill or performance deficit. When his performance was incentivized with coupons that could be redeemed for special privileges or rewards, the participant showed improvement across all three outcomes of interest. Despite these improvements, however, he was still not performing at a mastery level. This indicated that he also had a skill deficit and would thus benefit from additional instruction in reading in combination with a motivational intervention. Duhon et al. (2004) conducted a study with four elementary-aged participants who were underperforming in math, writing, spelling, or a combination of academic subjects. Using brief experimental analysis methodology, the authors introduced incentives to assess whether the students were demonstrating skill or performance deficits. Two students were identified as having performance deficits, while the other two did not significantly improve with the incentive and were identified as having skill deficits. Based on these assessment results, Duhon et al. (2004) conducted an extended analysis with interventions that addressed each student's identified type of deficit. During this extended analysis, students who were identified as having a skill deficit were provided additional instruction in their academic areas of need, while students who were identified as having performance deficits received interventions that focused on goal setting and reinforcement. These targeted interventions were effective for all four students, indicating the correct identification of their skill or performance deficits, as well as the effectiveness of implementing contingency-based reinforcement to improve the academic performance of students with performance deficits.

Although individual studies have demonstrated the utility of Can't Do/Won't Do assessments in identifying performance deficits and precisely matching an effective intervention to students' needs, the broader literature base in this area has yet to be systematically synthesized. For example, the percentage of academic intervention studies that use Can't Do/Won't Do assessments to inform intervention procedures is unknown. This information would provide valuable insight for the field, as it would help determine what information is currently used by researchers to identify student needs and develop interventions, particularly in cases where performance deficits may be present. If Can't Do/Won't Do assessments are not widely used in the research literature, this would suggest that academic interventions are being applied without a thorough assessment that can experimentally determine the primary source of a student's academic underperformance. On the other hand, if a systematic review of this literature base reveals that a substantial portion of studies conducts Can't Do/Won't Do assessments to inform intervention planning, this information would lay the groundwork for a future meta-analysis that may analyze the effects of using Can't Do/Won't Do assessments in problem identification and intervention planning.

Reinforcement and Intervention

Given the effects of reinforcement contingencies, this strategy is used in many school-based interventions to address performance deficits (e.g., McCurdy et al., 2020; Panahon & Martens, 2013; Skinner et al., 2004; Stage & Quiroz, 1997). Reinforcement can be implemented in two general ways: contingently and noncontingently. If reinforcement is provided regardless of the individual's behavior or performance on a specific task, it is considered noncontingent (Vollmer et al., 1993). Alternately, contingent reinforcement is a behavioral strategy in which reinforcement is only provided after a student exhibits a particular desired behavior. This ultimately serves to strengthen the reinforced behavior (Miltenberger, 2016). For interventions using contingent reinforcement to have optimal effects, they must do the following: (a) identify a specific behavior or skill to target, as is expected for any intervention (Graplin et al., 2018), (b) determine how reinforcement will be administered, and (c) identify the goal or criterion that the individual must achieve to receive reinforcement (Galbicka, 1994).

In past literature, contingent reinforcement has been effectively used to promote academic engagement, productivity, performance improvement, task completion, and skill accuracy (e.g., Chadwick & Day, 1971; Martens et al., 1992; Park et al., 2019). Contingent reinforcement has also been shown to improve student performance in reading, writing, and math (Duhon et al., 2004; Eckert et al., 2002). Regarding the acceptability and feasibility of contingent reinforcement interventions, a survey found that elementary-school teachers reported that they believed they were able to implement contingencies with no external support, though teachers also reported some concerns about the intervention's acceptability and compatibility with school culture (Briesch et al., 2015).

Although contingent reinforcement can be implemented as its own independent intervention, it often serves as one of multiple treatment elements that are integrated into a larger intervention package. Contingency-based reinforcement has been implemented in conjunction with a variety of treatment elements including peer tutoring (Piggott et al., 1986), changes in instructional elements (Axelrod et al., 1987; Codding et al., 2011), and choice (Burton, 2012). Overall, studies that examined packaged interventions that included elements of contingency-based reinforcement have demonstrated positive effects on students' academic performance across a range of academic subjects (e.g., Bendell et al., 1980; Codding et al., 2011; Eckert et al., 2002; Newstrom et al., 1999; Piggott et al., 1986).

Gilbert (1990), for instance, explored a combination of four treatments to examine the effects of contingency-based reinforcement on spelling performance, with and without an element of competition between students. In two of the treatment phases, contingency-based reinforcement was implemented in combination with a competitive group structure. In the other two phases of treatment, however, contingency-based reinforcement was implemented in isolation without the influence of competition among students. These two isolated phases of treatment allow a causal relationship to be inferred between contingency-based reinforcement and spelling outcomes without the impact of a third treatment variable (competition, in this case). Results of this study indicated that, in the cases without a competitive group structure, students scored an average of 57.14% correct on spelling tests when rewards were provided for each individual, while the average mean score was 72.71% when rewards were provided for students as a group.

In contrast, Eckert and colleagues (2002) implemented an intervention that involved the use of an antecedent intervention in combination with two potential consequences—contingent reinforcement and/or performance feedback. Results demonstrated that four of the six student participants showed increased reading performance when the antecedent intervention was implemented in conjunction with the consequence-based elements. These results provided empirical support for a treatment package for reading performance that includes both antecedent and consequence manipulations. However, because the impact of contingency-based reinforcement was never examined without the antecedent intervention in this study, it is impossible to determine the role of contingent reinforcement on reading outcomes. Thus, Eckert et al. (2002) provided valuable information about the combined effects of antecedent interventions with varying consequences, but the study cannot speak to the effectiveness of contingency-based reinforcement alone.

Another example of a treatment package that included—but did not isolate—contingency-based reinforcement was conducted by Codding and colleagues (2011). This study examined the effectiveness of Kindergarten Peer-Assisted Learning Strategies in Mathematics (KPALS) on students' early numeracy skills, with and without goal setting and contingent reinforcement. Results of this study indicated that the KPALS-only intervention had a moderate effect on number identification ($d = 0.52$), and the KPALS intervention with contingent reinforcement and goal setting had a moderate effect as well ($d = 0.56$). Although the second treatment phase did include contingency-based reinforcement, the KPALS intervention was present in both treatment conditions, so the effect of contingency-based reinforcement could not be isolated. As such, the effect size of the second treatment phase likely reflects an interaction between the KPALS intervention, goal setting, and contingency-based reinforcement, rather than contingency-based reinforcement alone.

Among the systematic reviews and meta-analyses that have explored contingency-based reinforcement interventions, inclusion criteria typically only require the use of contingency-based reinforcement in any capacity, regardless of whether it is in isolation or in combination with other intervention elements. While this provides an overarching overview of when and how contingency-based reinforcement has been included in the literature, the inclusion of packaged interventions complicates the utility of such systematic reviews and meta-analyses. It is impossible to determine the causal relationship between contingency-based reinforcement and performance outcomes if contingency-based reinforcement is never implemented in isolation. In other words, additional intervention elements essentially serve as confounding variables that complicate the way data can be analyzed and interpreted.

From an applied perspective, the inclusion of packaged treatments in systematic reviews and meta-analyses also lessens these studies' practical utility among educators who wish to implement a contingency-based reinforcement intervention in their classrooms. If an educator is solely interested in implementing a contingency-based reinforcement intervention, empirical support for interventions that also require other elements (e.g., peer tutoring, changes in instruction, manipulations of social cohesiveness) provide little value. This is primarily because a study that features a packaged intervention can only provide empirical support for that specific intervention. In other words, an educator who solely plans to implement a contingency-based intervention relies on empirical studies that are able to demonstrate the causal relationship between this isolated intervention and student outcomes.

Reward Selection

Although some researchers determine possible rewards through methods like student suggestions (e.g., Lynch et al., 2009; Theodore et al., 2004), preference and reinforcer assessments may be conducted to systematically determine what specific stimuli are preferred and reinforcing for an individual (DeLeon & Iwata, 1996; Fisher et al., 1992; Green et al., 1988; Pace et al., 1985). Although there can be some procedural variability among preference assessments, all preference assessments include three general steps. First, the educator identifies and gathers potential reinforcers. They then expose the student to each of these potential reinforcers. Finally, the educator collects data on the student's response to each item to determine which of the potential reinforcers are most likely to motivate the individual (King &

Kostewicz, 2014). Afterward, the educator may also conduct a reinforcer assessment, in which they provide the potential reinforcer contingent on a particular behavior and see if that behavior increases with the delivered reward (Northup et al., 1996).

Preference assessments have effectively identified activity-based (Daly et al., 2009), edible (Fritz et al., 2020; Resetar & Noell et al., 2008), and tangible reinforcers (Noelle et al., 2000). Based on previous empirical findings, preference assessments are an effective way to determine reinforcers when they are included in contingency-based reinforcement intervention (Daly et al., 2009; Fritz et al., 2020; Noell et al, 2000; Radley et al., 2019; Roane et al., 1998). There is also some research to suggest that preference assessments can be conducted in an individual (Radley et al., 2019; Resetar & Noell et al., 2008) or class-wide format with all students responding to the stimuli at once (Radley et al., 2019). For example, Radley and colleagues (2019) compared the results of individual and group preference assessments for 19 seventh-grade students. During the individual preference assessments, students individually met with the researchers, and they were given an opportunity to sample all of the potential edible reinforcers before the assessment began. The researchers then lined up four edible items and instructed the student to select one. Students were given 30 seconds to eat the selected item, after which the researcher lined up four new edible items and repeated the process. In contrast, the group preference assessment was conducted with the entire class participating in each trial. Before the assessment, each student received a card that allowed them to indicate their preferred item among four choices. The researchers then projected four possible reinforcers on a screen for all of the class to see at once. Students chose their most-preferred item and noted their decision on their cards. Using a phone application, the researchers compiled each student's preference.

The application used this information to create pie charts for each trial, with more preferred items receiving a larger portion of the chart than less-preferred items. For example, if 50% of the students chose chocolate, 50% of the pie chart would be devoted to chocolate. A spinner was then spun on the pie chart, and all students received the item the spinner landed on. This process was repeated with each trial. Results of this study indicated that, across trials, an average of 40% of students identified the same item as their most preferred stimulus, regardless of the preference assessment format. Moreover, the individual and group preference assessments identified similar preference hierarchies for an average of 38% of students across trials. Group preference assessments were more efficient and required 3.79 minutes per student assessed, whereas individual assessments required 10.79 minutes per student assessed. Overall, these results suggest that both individual and group preference assessments are viable options to determine rewards for a contingency-based reinforcement intervention.

Despite the use of preference and reinforcer assessments in individual studies, however, they have only been discussed in one meta-analysis about contingency-based reinforcement interventions with students (Pokorski et al., 2017), which noted that none of their included studies reported using preference assessments to inform reinforcement. Because Pokorski et al. (2017) only included studies with preschool-aged participants, the use of preference and reinforcer assessments with older students has not yet been synthesized. However, it is important for practitioners and researchers to know how frequently preference and reinforcer assessments are being conducted in contingency-based reinforcement intervention literature, as this information may impact the implications of studies that featured ineffective contingency-based reinforcement interventions. In those studies, if preference and reinforcer assessments were not conducted, it is possible that intended rewards may not have truly served as reinforcers. As such, any results indicating that contingency-based reinforcement interventions were not effective for students may have failed simply because the selected rewards were not sufficiently motivating to change students' behavior. Alternatively, if preference and reinforcer assessments are consistently being conducted to inform contingency-based reinforcement interventions in the existing research literature, future researchers may choose to meta-analyze the effect of these assessments on intervention outcomes to determine if they are a necessary component for effective contingency-based reinforcement interventions.

Target Identification

Before a contingency-based reinforcement intervention can be implemented, the intervention implementer must identify the specific skills and behaviors that the intervention will target. When implementers attempt to improve students' academic skills, they must decide whether to target academic skills directly by reinforcing academic performance (e.g., addition fluency, spelling accuracy), or indirectly through reinforcement of behaviors related to academic performance (e.g., task engagement, compliance). An intervention target may be selected based on a student's performance on academic assessments, which commonly encompass standardized state tests, curriculum-based measurements, norm- and criterion-referenced tests, and computer-adaptive tests (Graplin et al., 2018).

Within the extant research on academic intervention, some studies have examined the effect of contingency-based reinforcement interventions on students' direct academic performances (e.g., Alric et al., 2007; Daly et al., 1999; Gilbertson et al., 2008; Turco & Elliott, 1990). These studies directly targeted students' academic performances as the primary outcome of interest and included some form of reward based on academic performance. Previous research has been conducted in the areas of spelling (e.g., Popkin & Skinner, 2003; Turco & Elliott, 1990), math (e.g., Gilbertson et al., 2008; Popkin & Skinner, 2003), and reading (e.g., Alric et al., 2007; Daly et al., 1999). For example, using an alternating treatments design, Alric and colleagues (2007) examined the effectiveness of a contingent-reinforcement intervention on the reading fluency (i.e., words read correctly per minute) of eight fourth-grade students receiving remedial reading instruction. Over the course of six weeks, teachers implemented reinforcement contingencies based on students' reading performances. Each day, teachers also individually administered curriculum-based measurement in reading (CBM-R) probes and recorded the number of words that the student read correctly. Effect sizes were determined using a variation of Cohen's *d* based on calculations by Busk and Serlin (1992), and this study's results demonstrated that contingency-based reinforcement interventions had a large effect on both academic performance (ES = 3.31) and homework accuracy or completion (1.04).

There is also a great deal of research that has targeted academic performance indirectly through the reinforcement of behaviors that are related to—but separate from—academic performance. In these cases, students' academic skills were not directly targeted by the intervention; instead, researchers used the intervention to target other classroom behaviors. Frequent intervention targets include on-task behavior (Williamson et al., 2009), task engagement (Chaffee et al., 2020; Hirsch et al., 2016; Lum et al., 2019), and appropriate or non-disruptive classroom behavior (Hartman & Gresham, 2016; Ling & Barnett, 2013; McKissick et al., 2010). Oftentimes, these behaviors are selected as intervention targets because these inappropriate classroom behaviors interfere with the learning environment and, as a result, reduce opportunities to engage with class content that is necessary to develop academic skills (e.g., Hartman & Gresham, 2016; Popkin & Skinner, 2003; Williamson et al., 2009). In previous research, teachers have also reported that they believe that students' disruptive behaviors significantly impact the effectiveness of their teaching (Chaffee et al., 2020; National Center for Education Statistics, 2012b). Off-task behavior has been shown to predict later reading outcomes (Moffett & Morrison, 2019), and interventions targeting off-task behavior have demonstrated collateral improvements in math accuracy during independent work time (King et al., 2017). Problem behaviors in the classroom, including internalizing and externalizing behaviors and hyperactivity, have been shown to negatively predict students' current academic achievement (Milecki et al., 2002).

As an example, using a single-case reversal design, Chaffee and colleagues (2020) examined the effect of a contingency-based reinforcement intervention called "Tootling" on middle school students' academic engagement and disruptive classroom behavior. The Tootling intervention involves a group contingency in which students record their peers' positive classroom behaviors, and the whole class receives a reward if enough positive behaviors are observed. Academic engagement was operationalized as passive or active on-task behaviors (e.g., reading aloud or looking at the teacher during instruction), while disruptive behavior included out-of-seat behavior, unpermitted talking or other vocalizations, and unrelated motor activities (e.g., playing with objects, touching other students). In part, these behaviors were selected because of their associations with academic achievement and success. Data were analyzed at the classroom level, and a visual inspection of the data demonstrated an overall trend of increases in academic engagement as a result of the intervention, although there was a significant overlap between the intervention and withdrawal phases in particular. Other findings of this study showed that this group-contingency intervention had a moderate to large effect on increasing academic engagement (Tau-U = 0.68 to 0.79, NAP = 0.76 to 0.92). The results indicated that Tootling had a positive effect on behaviors related to students' academic performance, but the authors did not directly measure students' academic skills. Indeed, a failure to examine whether the effects of contingency-based reinforcement interventions targeting behaviors like academic engagement generalize to academic performance is common (e.g., Ling & Barnett, 2013; McKissick et al., 2010; Williamson et al., 2009).

The vast majority of systematic reviews and meta-analyses in this topic area focus on outcomes like disruptive behavior or task engagement, rather than direct academic performance (Bowman-Perrott et al., 2016; Chaffee et al., 2017; Little et al., 2015; Long et al., 2019; Maggin et al., 2012; Maggin et al., 2017; Pokorski et al., 2017; Stage & Quiorz, 1997). Collectively, the results of these studies indicated that contingency-based reinforcement interventions effectively reduce disruptive behavior and promote positive behaviors like task engagement (Bowman-Perrott et al., 2016; Chaffee et al., 2017; Little et al., 2015; Long et al., 2019; Maggin et al., 2012; Maggin et al., 2017; Stage & Quiorz, 1997). Nonetheless, previous research has shown that improvements in related behaviors do not necessarily generalize to improved academic performance for all students (Dion et al., 2011; Dolan et al., 1993; Ialongo et al., 1999). Dion and colleagues (2011) conducted a randomized controlled trial comparing the additive effects of a group contingency when it was paired with a peer-tutoring intervention on 409 first-grade students' inattention (based on teacher ratings and classroom observations) and end-of-year reading assessment scores. Results of hierarchical linear models demonstrated that the contingency-based reinforcement intervention improved students' attention levels during class (*B* = .17, $p < .01$), but the contingency had no significant effect on students' word recognition or comprehension. This study's findings suggested that improvement in classroom behaviors (i.e., attention) did not generalize to students' reading performances. Ultimately, the results of this study demonstrate that improvements in behaviors that are associated with academic performance may not always lead to improvements when academic achievement is directly measured. As such, there is a clear need for a closer examination of the literature that has specifically explored the impact of contingency-based reinforcement interventions on direct academic performance.

To date, only two meta-analyses have explored academic performance as a target behavior or outcome for contingency-based reinforcement interventions (Cerasoli et al., 2014; Little et al., 2015). Cerasoli and colleagues (2014) examined the impact of intrinsic motivation and incentives on various types of performance (e.g., job or academic performances) and found that incentives explained greater unique variance in quantity-based performances ($\beta = .33$) than intrinsic motivation ($\beta = .24$). Although these results suggest that quantity- or effort-based measures of academic performance may be more sensitive to incentives, this study did not provide information about what specific academic behaviors were targeted in studies. Moreover, this study's syntheses and analyses did not isolate child participants from adult participants or academic performance from work-related performance. Because of these limitations, it is unclear what proportion of contingency-based reinforcement intervention studies directly target academic performance, how they operationalize direct academic performance, and what participant demographics are represented in the current literature.

In contrast, a meta-analysis by Little et al. (2015) did examine the effects of various group contingencies on child participants' classroom behaviors and academic performance. The study included two performance-based dependent variables as outcomes: academic performance, which was operationalized according to quantitative measures like accuracy and fluency, as well as homework, which included participants' accuracy or completion. The effect size of contingency-based reinforcement interventions on academic performance, which was determined using a variation of Cohen's *d* based on calculations by Busk and Serlin (1992), was 3.31; however, these interventions had a smaller effect on homework accuracy or completion (ES = 1.04). While this meta-analysis provided insight into the use and effectiveness of group contingencies on some academic outcomes, only 28% (i.e., 14 out of 50) of the included studies directly targeted academic performance or homework completion/accuracy. In part, this may be due to Little and colleagues' (2015) focus on group contingencies specifically, despite previous literature that has implemented contingency-based reinforcement interventions at the individual level (e.g., Gansle et al., 2002). In addition, this meta-analysis only included studies published through 2010. Based on these limitations, further synthesis is needed to examine studies with individually implemented interventions, as well as any works published over the last decade.

Because much of the current literature about contingency-based reinforcement interventions in school settings centers around behavioral outcomes like engagement or disruptive behavior, there is a need for an updated and comprehensive systematic review that focuses explicitly on studies that have directly targeted academic performance outcomes when contingency-based reinforcement interventions are implemented. This review would synthesize how contingency-based reinforcement interventions that target academic performance have been previously implemented, which would provide guidance to researchers and practitioners who wish to develop contingency-based reinforcement academic interventions. It would also allow researchers to identify areas in need of further research. Ultimately, depending on the state of the research in this area, the review could facilitate a later meta-analysis to determine the effectiveness of various components of contingency-based reinforcement interventions on academic performance.

Intervention Format

Contingency-based reinforcement interventions can be implemented in individual or group formats. In an individual format, reinforcement contingencies are structured and provided for one specific student (Litow & Pumroy, 1975; Skinner et al., 2004). For example, Gilbertson et al. (2008) implemented individual contingencies to examine their impact on the math fluency of four elementary students. The authors first conducted a brief assessment comparing the effect of rewards versus additional instruction on students' digits correct per minute. The initial assessment and subsequent intervention utilized individual contingencies, in which each student had a different criterion for reinforcement based on their initial baseline performance. The results of the pre-intervention assessment suggested that all participants would likely benefit from an intervention that combined rewards and instruction, although contingent reinforcement independently led to improvements in all students' digits correct per minute (increases ranged from 23% to 42%). In an extended analysis, the effectiveness of this intervention was examined using a multiple baseline design. A visual inspection of the data reflects improvements in each students' math fluency with the introduction of the intervention, although the degree of change and rate of growth differed between students. Overall, individual contingencies have been utilized in prior academic literature (e.g., Gansle et al., 2002; Gilbertson et al., 2008), though they have not been explored in isolation by any systematic review or meta-analysis to date.

As their name suggests, group contingencies are administered to groups of students (Gresham & Gresham, 1982). When multiple students require or may benefit from a contingency-based reinforcement intervention, group contingencies may be preferred over individual contingencies for a variety of reasons (Gresham & Gresham, 1982; Skinner et al., 2004). One relevant benefit of group contingencies in the classroom is their efficiency; teachers can establish contingencies and provide reinforcement to groups of students at once, rather than doing so individually (Gresham & Gresham, 1982; Skinner et al., 2004).

There are three possible types of group contingencies: independent, dependent, and interdependent (Litow & Pumroy, 1975). Independent contingencies are somewhat similar to individual contingencies in that a student's access to reinforcement occurs solely as a function of their own performance or behavior, regardless of the behavior of their peers (Crouch et al., 1985). However, independent contingencies are still classified as a group contingency because all members of the group share the same criterion or contingency for reinforcement (Alric et al., 2007; Litow & Pumroy). One drawback of independent contingencies, however, is that they lack some of the efficiency of dependent and interdependent contingencies, because it is more time-efficient for teachers to monitor performances and administer reinforcement for students as a group (Skinner et al., 2004). Nonetheless, past research on independent contingencies has found them to be effective with spelling accuracy (Shapiro & Goldberg, 1986), math fluency (Gross et al., 2016), reading fluency (Alric et al., 2007), and completion of literacy worksheets (Deshais et al., 2019).

In opposition, with a dependent contingency, an individual's access to reinforcement is dependent upon their peers' performance. When a dependent contingency is implemented, the entire class receives—or does not receive—a reinforcer based on the performance of select students (Williamson et al., 2009). Compared to other types of group contingencies, dependent contingencies are sometimes considered unfair due to students' dependence on others, and this dynamic could potentially lead to high levels of peer pressure or other negative peer interactions for students whose performance impacts the entire class's access to reinforcement (Romeo, 1998; Skinner et al., 2004; Theodore et al., 2004). However, dependent contingencies have successfully targeted spelling performances (Shapiro & Goldberg, 1986), reading fluency (Alric et al., 2007), and math accuracy (Scott et al., 2017).

With interdependent group contingencies, reinforcement is provided when all members of a class or group of students meet the same goal or criterion (Coogan et al., 2007). This differs from dependent contingencies because, alongside their peers, the individual student must personally meet the goal for reinforcement. One limitation of these contingencies is that it can be difficult to select an adequate contingency goal because students' skills are at different levels (Litow & Pumroy, 1975; Sharp & Skinner, 2004). In cases like these, what may seem like a reasonable goal for the average student may be incredibly challenging for a struggling student and simple for an advanced student. This limitation notwithstanding, previous studies that have implemented interdependent group contingencies have found them to be effective when students' academic performances are directly targeted (e.g., Scott et al., 2017; Shapiro & Goldberg, 1986; Sharp & Skinner, 2004).

Although many of the aforementioned studies examined these contingencies in isolation, other research has compared the formats to each other to examine what form of contingency may be most effective under a particular set of circumstances. For example, Shapiro and Goldberg (1986) used an alternating treatments design to determine the differential effects of independent, dependent, and interdependent contingencies on the spelling accuracy of sixth-grade students ($n = 53$). During the independent contingency intervention phase, each student who scored at least 90% on their spelling test received a reward. Alternatively, the dependent contingency was structured so that the entire class still completed a spelling test, but only one student's name was drawn out of a box. If this specific student scored at least 90% on their exam, the entire class received a reward. The selected student was not identified to their classmates at any time. Finally, in the interdependent contingency phase, all students completed the spelling test, and the class's average score was calculated. If the mean score on the spelling test was at least 90%, the entire class received a reward. To analyze their results, Shapiro and Goldberg (1986) used students' baseline spelling accuracy to divide them into high- (i.e., 85 to 100%), middle- (i.e., 70 to 84%), and low-performing (i.e., below 70%) groups. A visual inspection of these data indicates that, regardless of baseline performance, students' spelling accuracy generally improved with the group contingencies. There were no clear separations in the data based on the type of contingency implemented, indicating that students' spelling accuracy did not differ as a function of the contingency format that was implemented. Results of this study are consistent with previous research that has also demonstrated similar effectiveness across contingency formats (e.g., Alric et al., 2007; Scott et al., 2017). Collectively, the results of these studies do not point to the superior effectiveness of one type of group contingency over another.

To date, both individual and group contingencies within school settings have been examined in reviews and meta-analyses on contingent reinforcement (Bowman-Perrott et al., 2016; Chaffee et al., 2017; Little et al., 2015; Long et al., 2019; Maggin et al., 2012; Maggin et al., 2017; Pokorski et al., 2017; Stage & Quiroz, Warmbold-Brann et al., 2017). Nonetheless, the information that these reviews and meta-analyses provide about contingency formats is primarily limited to contingencies based on indirect academic targets like task engagement or appropriate classroom behavior. Indeed, a meta-analysis by Little and colleagues (2015) provides the only extant synthesis of the contingency formats used in studies directly targeting academic performance. This meta-analysis was designed to examine the efficacy of group contingencies for school-aged participants, and results demonstrated that group contingencies were effective when used to target students' academic performances (ES = 3.31, based on modified Cohen's *d* calculations by Busk and Serlin [1992]). Although these findings provide valuable insight into the use of group contingencies in contingency-based reinforcement intervention literature, there are important methodological limitations to consider. Little and colleagues (2015) only included group contingencies in their analyses, which excludes any potential studies that used individual contingencies. Additionally, this meta-analysis was focused on more comprehensive classroom behaviors, and this is reflected in their generalized search terms that only included "children," "classroom," or "school" with contingency-based reinforcement terms. Given the broad nature of these search terms, it is likely that studies targeting direct academic performance may have been missed in the authors' initial search. With these limitations in mind, there is an evident need for additional exploration of contingency formats within the literature that has directly targeted academic performance.

It is still unclear what contingency formats are most often used in studies directly targeting academic performance, particularly given Little and colleagues' (2015) exclusion of individual contingencies. This information will benefit future researchers, however, as it will reveal what contingency formats have the prerequisite empirical support to be meta-analyzed. It may also have important ramifications for practitioners, as it will highlight what contingency formats have the greatest empirical support within the context of specific academic subjects or skills. As an example, this synthesis may allow a math teacher to determine if it is more appropriate to target students' addition fluency using independent or dependent contingencies based on the breadth of research that has examined each contingency format when directly targeting math fluency. Thus, an examination of contingency formats within the literature directly targeting academic performance will provide valuable insight for the field.

Criterion Identification and Selection

By definition, an intervention that uses contingent reinforcement requires established criteria for reinforcement (Miltenberger, 2016). That is, once a target behavior is identified, researchers must then determine the level that behavior must reach for students to access reinforcers (Galbicka, 1994). There are multiple ways to determine what criteria will be used for contingent reinforcement. One method to determine the reinforcement criterion is percentile shaping, in which the criterion is calculated using a specific equation that integrates the student's previous performances on the task (Galbicka, 1994). For example, Athens and colleagues (2007) used a reversal design to examine the effect of percentile shaping on four students' task engagement across a variety of academic assignments (e.g., independent writing, copying sentences). During the intervention, the researchers systematically increased the amount of task engagement required for a reward until the reinforcement criterion matched the duration of engagement recommended by the students' teachers. Results of this study showed that percentile shaping was most effective on students' task engagement when the reinforcement criterion was based on multiple previous trials, rather than only using data from the most recent trials to determine the next trial's reinforcement criterion.

Alternatively, reinforcement may be delivered if a student's performance improves by a certain amount compared to their previous attempts (Gilbertson et al., 2008), or researchers may simply tell the student that they will receive reinforcement as long as their next performance improves upon the last (VanDerHeyden & Witt, 2008). For instance, Gilbertson and colleagues (2008) used a multiple-baseline design to examine the effects of a contingency-based reinforcement intervention targeting the math computational fluency (i.e., digits correct per minute) of four elementary-aged participants. During baseline, each student completed as many math probes as they could in six minutes, and reinforcement criteria were subsequently set at 10% above the students' median score across the baseline probes. When contingent reinforcement was introduced using these reinforcement criteria, all four students' math computation fluency improved, with increases ranging from 23% to 42%.

At times, reinforcement may be provided on a continuous, item-by-item basis, rather than based upon an overarching criterion for cumulative performance. For example, under this type of reinforcement, a student may receive a piece of candy after each correct response. This differs from the aforementioned types of criteria because the reward is provided after a correct or incorrect response to an individual question, item, or task. In essence, some contingency-based reinforcement interventions require that a student's cumulative performance reaches a certain criterion (e.g., 90% of math problems solved correctly), while others are structured dichotomously (e.g., a student receives a sticker when they answer one math question correctly, and they do not receive a sticker when the answer one math question incorrectly). Hofstadter-Duke and Daly (2015), for example, used a dichotomous method of reinforcement in all three of their treatment conditions. For the student and peer attention conditions, each participant received positive social attention upon completion of each math problem. Similarly, during the escape condition, the student was granted a 15-second break from work after each completed problem.

Unknown and Randomized Contingency Components

Although successful reinforcement contingencies require established criteria to determine if reinforcement should be provided, these criteria are not always communicated to students at the outset of the intervention. Instead, some contingency-based reinforcement interventions utilize unknown and randomly selected contingency components (e.g., Kelshaw-Levering et al., 2000; Scott et al., 2017). With unknown or randomized contingency criteria, students do not know the specific goal or target behavior for reinforcement until the behavior has already occurred, or they may be unsure whose performance will be assessed to determine if a reward has been earned (Kelshaw-Levering et al., 2000; Skinner et al., 2004).

Unknown or randomized contingency components are particularly valuable in the context of group contingencies, in part because they help negate potential limitations that are associated with group contingencies (Kelshaw-Levering et al., 2000; Scott et al., 2017; Skinner et al., 2004). For one, unknown and randomized contingency components allow educators to provide reinforcement based on the performance of randomly selected students, and this decreases the amount of time required for contingency-based reinforcement interventions because the teacher does not need to assess each student's performance each time reinforcement is provided (Heering & Wilder, 2006; Scott et al., 2017). Additionally, when reinforcement criteria are unknown, this increases the likelihood that all students will perform to the best of their abilities because they do not know what the criteria will be until after their work has already been submitted (Popkin & Skinner, 2003; Scott et al., 2017). When target behaviors are unknown and randomly selected, this may also encourage students to focus on improvement across a variety of behaviors (e.g., both math and spelling accuracy), even if only one of these behaviors is ultimately selected for reinforcement (Kelshaw-Levering et al., 2000).

For example, Scott and colleagues (2017) used unknown contingency criteria as part of a contingency-based reinforcement intervention targeting the math accuracy of 16 first-grade students. Using an alternating treatments design, the authors implemented two contingency interventions. In the first intervention, the authors implemented an interdependent group contingency, in which reinforcement criteria were based on a class-wide average; in the second intervention, the authors implemented a dependent contingency, where the entire class's rewards were provided based on the performance of four randomly selected students. Results of this study showed that both the class-wide and random selection interventions effectively increased students' math accuracy above baseline performance (percent non-overlapping data [PND] = 100% for both intervention conditions). When the two treatments were compared, Hedges' g analyses demonstrated a large effect of the randomized treatment on math accuracy compared to baseline performance ($g = 4.09$) and a moderate effect of the class-wide treatment compared to baseline performance ($g = 2.57$). There were negligible differences in performance between the two interventions ($g = 0.08$). Overall, this study demonstrated that students' math accuracy improved with the interventions, even when students did not know if their personal performance would be randomly selected and scored to determine the class's access to reinforcement.

In another study, Kelshaw-Levering and colleagues (2000) conducted an intervention targeting the disruptive classroom behaviors of 12 second-grade students using a multi-phase time-series design (i.e., A-B-A-C-B-C). Target behaviors and target students were randomized as part of the intervention, as the teacher would first randomly select a target behavior from a jar (e.g., noncompliance, inappropriate vocalizations), and they would then randomly select a sheet indicating if the entire class' behavior would be monitored according to the reinforcement criterion versus an individual student's behavior. As a result of the intervention, students exhibited lower and more stable levels of disruptive behavior during intervention phases compared to baseline. The authors also compared the effectiveness of randomizing multiple contingency components (i.e., target behavior and target student) compared to randomizing reinforcers in isolation, and they found that randomization of multiple contingency components led to at least equal decreases in students' disruptive behaviors across intervention phases.

Overall, interventions using unknown and randomly selected contingency components have been shown to successfully target on-task behavior (Williamson et al., 2009), appropriate classroom behaviors (Coogan et al., 2007; Kowalewicz & Coffee, 2014), and homework accuracy (Ferneza et al., 2013). They have been successfully implemented with students who are typically developing (Ferneza et al., 2013; Kowalewicz & Coffee, 2014) and non-typically developing (Coogan et al., 2007; Williamson et al., 2009). Evidently, randomized or unknown contingencies have been utilized with positive results in many contingency-based reinforcement intervention studies targeting academic behaviors or performance. Despite the substantial presence of randomized or unknown contingency criteria in past academic literature, however, this strategy has not yet been synthesized in any existing review or meta-analysis of the contingency-based reinforcement intervention literature. Previous reviews and meta-analyses (Chaffee et al., 2017; Little et al., 2015; Long et al., 2019; Maggin et al., 2012; Pokorski et al., 2017) have acknowledged that contingency-based reinforcement interventions often utilize randomized or unknown contingency components, but no specific information has been provided regarding the use and effectiveness of unknown or randomized contingency criteria specifically (vs. unknown rewards, for example).

Despite the variety of techniques that can be used to identify the criteria necessary for reinforcement delivery (i.e., percentile shaping, alternate calculations, dichotomies, and/or unknown or randomized criteria), the prevalence of those techniques in the contingency-based reinforcement intervention literature has yet to be synthesized. However, a review of the contingency criteria that have been implemented in intervention studies targeting academic performance will provide valuable information for practitioners and researchers alike. This information will provide practitioners with guidance on which reinforcement criteria have been most well-researched if they choose to select a contingency-based reinforcement intervention to address their students' academic underperformances. For academic intervention researchers, a summary of the reinforcement criteria evident in contingency-based reinforcement interventions will facilitate a later meta-analysis of the effects of various reinforcement criteria. Additionally, a systematic review of randomized or unknown contingency criteria will reveal how frequently this specific criteria strategy is used, alongside what student populations it has been used with, to reveal the empirical precedence for designing a contingency-based reinforcement intervention for academics that includes randomized or unknown contingency criteria. With this information, future researchers may decide to include (or not include) randomized or unknown contingency components with a concrete sense of how this decision aligns with—or perhaps addresses gaps in—the extant literature on contingency-based reinforcement interventions.

Purpose of the Current Study

The most recent NAEP data (Hussar et al., 2020; NCES, 2011, 2015) indicate that students in the United States are demonstrating widespread academic deficits. As such, there is a need for further research to determine the best ways to intervene upon these deficits. One method to intervene upon academic deficits is through the use of contingent reinforcement, particularly in cases where students are demonstrating performance deficits (Stipek, 2004). Within school settings, contingency-based reinforcement interventions have been shown to effectively target both academic and behavioral outcomes for a broad range of student participants (e.g., Chadwick & Day, 1971; Duhon et al., 2004; Eckert et al., 2002; Martens et al., 1992; Park et al., 2019).

Although there is no dearth of individual studies about classroom contingencies, the literature in this area has not yet been comprehensively synthesized using the rigorous methodology of a systematic review. This is perhaps unsurprising, as behavioral research often utilizes single-case experimental design methodology, which has historically been excluded from meta-analyses (Allison & Gorman, 1993). Nonetheless, reviews are crucial in the interpretation and application of research. They synthesize and descriptively quantify the findings from individual studies to provide a more complete view of the relevant literature, such that “review articles have a power and value that no single study can match” (Baumeister & Leary, 1997;

Cumming, 2004; Siddaway et al., 2019). Although multiple types of reviews exist (e.g., narrative or scoping reviews), systematic reviews are considered a "gold standard" within the field of psychology due to their systematic, transparent, and replicable search procedures (Boland et al., 2017; Higgins et al., 2019). When correctly done, the comprehensive and methodological nature of systematic reviews minimizes potential bias and allows researchers to critically appraise the state of literature in a specific area of study (Higgins et al., 2017; Siddaway et al., 2019). Through the rigorous methodology of a systematic review, additional synthesis of the contingency-based reinforcement intervention literature targeting academics will effectively consolidate disparate pieces of the field's current knowledge. This will aid future researchers and practitioners to address the widespread academic underperformance in the United States as they design and implement contingency-based reinforcement interventions that are informed by a holistic, updated synthesis of the literature in this area.

Ten previous meta-analyses and systematic reviews have examined the effectiveness of contingency-based reinforcement interventions in classrooms (Bowman-Perrott et al., 2016; Cerasoli et al., 2014; Chaffee et al., 2017; Little et al., 2015; Long et al., 2019; Maggin et al., 2012; Maggin et al., 2017; Pokorski et al., 2017; Stage & Quiroz, 1997; Warmbold-Brann et al., 2017). Nonetheless, large gaps remain in the field's understanding of the broad literature surrounding the use of school-based contingencies (see Table 1). To date, only two meta-analyses (Cerasoli et al., 2014; Little et al., 2015) have focused on the effects of contingency-based reinforcement interventions on direct academic targets (e.g., accuracy, fluency, task completion), rather than behavioral proxies that are believed to generalize to academic improvements. However, previous research indicates that behavioral improvements do not consistently lead to improvements in students' direct academic performances (Dion et al., 2011; Dolan et al., 1993; Ialongo et al., 1999). Thus, there is an evident need for additional review to determine the effectiveness of contingency-based reinforcement interventions that directly target students' academic performances. Moreover, among the extant systematic reviews and meta-analyses on contingency-based reinforcement interventions for academics, little information is provided regarding the use of specific intervention components or practices. The field would benefit from additional information regarding the inclusion of assessments to inform intervention selection and reinforcement (i.e., Can't Do/Won't Do assessments, preference and reinforcer assessments), as well as the structure and administration of contingency-based reinforcement interventions that have targeted direct academic outcomes (i.e., contingency format and criterion identification).

The purpose of the current systematic review was to address these existing limitations. This information can be used to facilitate future empirical studies, meta-analyses, and intervention implementation by providing an updated and comprehensive view of the literature surrounding contingency-based reinforcement interventions with direct academic targets. The first aim of this systematic review was to summarize the basic characteristics of the relevant studies that have examined contingency-based reinforcement interventions. Specifically, the goal was to report when the relevant literature was published, what methodology was used, where data were collected, and how large the sample size was. As such, the following research question was explored:

Research Question 1: What are the basic publication, setting, and methodological characteristics of the studies that have explored contingency-based reinforcement interventions for academic performance?

The second purpose of this study was to identify the demographic characteristics of participants who have been (and have not been) included in studies that examine the efficacy of contingency-based reinforcement interventions for academics. This purpose was addressed through the following question:

Research Question 2: What are the demographic characteristics of participants who have been included in studies examining contingency-based reinforcement interventions?

The third goal of this systematic review was to examine the assessment practices used by studies that have directly targeted academic performance with contingency-based reinforcement interventions. To address this aim, the literature around these studies was evaluated to answer the following research questions:

Research Question 3: What percentage of studies included Can't Do/Won't Do assessments to inform intervention selection?

Research Question 4: What percentage of studies used preference and reinforcer assessments to identify effective rewards?

The fourth aim of this study was to examine the specific contingency intervention components that were utilized in previous studies that directly targeted academic performance through contingency-based reinforcement interventions. Specifically, the following research questions were explored:

Research Question 5: What percentage of studies used individual contingencies versus group contingencies? Of the interventions that were administered in a group-wide format, what percentage were implemented using independent, dependent, and interdependent formats?

Research Question 6: How frequently has each specific academic subject (i.e., math, reading, writing, spelling, or Language Arts) been directly targeted with contingency-based reinforcement interventions?

Research Question 7: What percentage of studies used percentile shaping, alternative calculations, general improvement, or a dichotomous variable to establish reinforcement criteria?

Research Question 8: What percentage of studies used randomized or unknown work, criteria, and/or rewards as part of their contingency-based reinforcement intervention?

Method

This review's methodology was developed using current standards within the field (Higgins et al., 2019; Moher et al., 2009), exemplars of systematic reviews published in the peer-reviewed literature (Ardoin et al., 2013; January & Klingbeil, 2020; Klingbeil et al., 2019; Wolfe et al., 2016), and one informational text (Boland et al., 2017). In line with those guidelines, this review included systematic procedures to search the literature, identify studies, and code and synthesize variables of interest. Each of these procedures is described below. Given that the focus of this review was on the assessment and intervention practices of contingency-based reinforcement interventions, rather than on the effect of these interventions on students' academic performances, neither methodological quality nor bias risk were assessed. If the effectiveness of contingency-based reinforcement interventions is explored in subsequent meta-analyses, however, methodological rigor and bias risk should be examined.

Study Identification and Search Procedures

To ensure a comprehensive review of the current literature, studies were identified in two stages (see Figure 1). During the preliminary stage, a series of key terms were searched in the American Psychological Association (APA) database PsycInfo (PsycINFO) and Education Resources Information Center (ERIC) databases. These databases were selected because they are consistent with search procedures in previous related literature (Bowman-Perrott et al., 2016; Long et al., 2019; Pokorski et al., 2017). Search results in these databases also include "gray" or unpublished literature (Higgins et al., 2019). Gray literature was incorporated into the review to reduce the risk of publication bias, as is consistent with recommendations in the *Cochrane Handbook for Systematic Reviews of Intervention* (Higgins et al., 2019). Boolean search procedures were used to augment the relevancy of results. The searches related to contingencies included the keywords: *contingen*, reinforcement,* and *reward*. To return results related to academic behavior and targets, the term "*academic performance*" was searched in combination with the contingency-related keywords using the operator "AND." To ensure that all relevant studies were discovered through these search procedures, two additional searches were conducted with the following assessment-related terms: *Can't Do Won't Do* and *performance deficit*. The aforementioned assessment-related terms were also combined with "academic performance" using Boolean search procedures. These terms were identified based on a previous meta-analysis (Warmbold-Brann et al., 2017), related terms found in the APA Thesaurus of Psychological Index Terms (American Psychological Association [APA], 2022), and the keywords and descriptors of relevant studies that were identified in a scoping search (e.g., Breaux et al., 2019; Gilbertson et al., 2008; Popkin & Skinner, 2003).

After the first and second stages of screening were complete, the first author conducted an ancestral review to increase the likelihood that all relevant studies were identified for inclusion. This step occurred upon the completion of initial screening and coding. During this stage, the first author manually reviewed the full reference list of each study that met the inclusion criteria. To identify additional studies that may have been relevant to the systematic review, each study's title was reviewed to determine if it met the inclusion criteria identified for relevancy in the first stage of screening; that is, the cited study's title must have included at least one prespecified academic term, as well as at least one prespecified intervention term. A full list of terms can be found in the coding manual in Appendix A. An example of academic terms includes: "academic achievement," "reading," and "academic performance." Examples of relevant intervention terms include: "contingency," "reinforcement," and "behavioral modification."

As part of this stage, the first author also used the "Cited By" function in Google Scholar to identify additional potential studies for inclusion. The "Cited By" function identifies sources that have cited a particular study. Each study that met inclusion criteria during the screening was entered into Google Scholar for the "Cited By" function to determine if there were any other relevant sources to be screened for inclusion.

The first author determined if the relevant studies found in the ancestral review were duplicates of any of the previously identified studies from first- and second-stage screening. For studies that were not duplicates and met inclusion requirements for the first stage of screening, the study's full text was located and screened according to second-stage screening criteria. Based on these two stages of screening, the studies found through the ancestral review were included in the systematic review.

Manual Content and Development

All coding during screening and full article review were completed in accordance with a coding manual (see Appendix A). This manual included step-by-step instructions for the initial study search, text storage, screening, ancestral review, and data extraction. Before coding began, however, the manual was piloted by the first author through an iterative development process. Using a series of studies that were found in the initial study search, the first author engaged in practice coding to determine how the manual needed to be updated before coding began. All significant changes to the manual (i.e., beyond language clarification) were documented in the manual.

Inclusion Criteria

All of the studies that were retrieved from this search were screened according to a set of inclusion criteria for further analyses. These criteria were:

1. The study's full text can be accessed.
2. The study was written in English.
3. The study used experimental or quasi-experimental methods to directly examine variables. This includes secondary data analyses and replication studies.
4. The study only included participants in kindergarten through twelfth grade. If participants' grades were not reported, participants must have been between the ages of 4 and 19.

5. The study was conducted in a school setting.
6. The study included a contingency-based reinforcement intervention that directly targeted academic skills in math, reading, writing, spelling, or English Language Arts. At least one phase of treatment must have examined contingency-based reinforcement as an isolated treatment component.
7. The study's dependent variables were sensitive to incremental changes in skill and directly applicable to the academic skill targeted.

Screening Procedures

Detailed screening procedures can be found in the complete coding manual in Appendix A. In essence, the screening process included two stages. The first stage was designed to eliminate studies that were non-applicable based solely on their titles and abstracts. The second stage of screening involved an in-depth review of the studies that were not excluded during the first stage of screening based on the title and abstract alone. During the second stage of screening, studies' full texts were examined to gather all necessary information to determine whether they met the inclusion criteria.

Screening, Stage 1

During this stage, the first author read the title and abstract of each study that was found during the initial study search to determine the study's relevancy to this systematic review. If any of the information presented in the title or abstract indicated that the study did not meet inclusion criteria (e.g., the sample included adult participants or was a meta-analysis), the study was excluded. Relevancy was determined based on the inclusion criteria outlined in the coding manual regarding the first stage of screening. Broadly, a study's title and abstract must have included specific terms related to academic performance and contingency-based reinforcement to be included during the first stage of screening. Duplicates were identified and removed if a study's title, authors, and publication year were an exact match.

Screening, Stage 2

After all the studies were screened for relevancy and duplication, each study's full text was meticulously reviewed to complete the screening process and identify all eligible studies for coding. 33% of these studies were also randomly selected for coding by an additional research assistant to calculate inter-scorer agreement. Characteristics examined for eligibility are outlined in the coding manual under the second-stage screening procedures. Additional duplicates were removed during this stage based on information provided in the studies' full texts. In cases where a thesis or dissertation was duplicated in a published study, the published study was included, as it is the most recent product and has undergone the peer review process.

Coding Procedures

Once all studies were screened for inclusion in the systematic review, those that met the inclusion criteria were reviewed in-depth by the first author and a research assistant. For all included studies, the following variables were coded: basic characteristics, participant demographics, assessments practices, and intervention components. Each variable is described below, and more detailed coding procedures can be found in the complete coding manual (Appendix A).

Basic Characteristics and Participant Demographics

Basic descriptive information was reported for each included study. This included the year that the study was published, the study's methodology, and its participant demographics.

Coders quantitatively noted the study's sample size, participants' ages and grades, and the percentage of participants that were female versus male and belonged to specific races and ethnicities. Participants' diagnoses or access to special education services were also noted. Coders also determined if the study used a group or single-case design methodology.

Assessment Practices

Three assessment variables were coded according to each study's inclusion of Can't Do/Won't Do assessments and preference and reinforcer assessments. All variables were scored dichotomously to reflect the inclusion or exclusion of the given assessment within the study. Regarding the "Can't Do/Won't Do" variable, studies did not need to name a specific assessment procedure for researchers to code that the authors included a Can't Do/Won't Do assessment. Broadly, the authors must have conducted an assessment prior to intervention implementation that involved offering participants rewards to improve their academic performances. The results of this assessment must have also been used to inform subsequent intervention selection. A study was coded as including a preference assessment if it involved the systematic examination of students' reward preferences before the intervention is implemented. A study was coded as including a reinforcer assessment if it involved a systematic examination of the effect of identified rewards on students' behaviors prior to intervention implementation.

Intervention Components

Researchers also coded variables related to contingency-based reinforcement intervention implementation. These variables included contingency format, criterion selection, academic subject, and use of randomized or unknown components. To record each intervention's contingency format, coders indicated whether the intervention was implemented in an individual or group setting. If the coder determined that a group format was used, in which a contingency-based reinforcement intervention was implemented with multiple students in a class, the coder then identified what type of group contingency was implemented. The contingency criteria variable reflected the way contingency criteria were selected (e.g., through percentile shaping) in each study. A separate criteria variable was scored to reflect whether tasks, criteria, or rewards were randomized or unknown to participants at any point in the study. Finally, studies were coded based on the academic subject of the behaviors being targeted by the intervention. Possible coding options included math, reading, writing, spelling, and English or Language Arts.

Coder Training and Interscorer Agreement

For both full-text screening and data extraction, ratings were completed by the first author (primary coder) and a research assistant (secondary coder). Secondary coders were trained using training manuals for both screening (see Appendix B) and data extraction (see Appendix C). Each coder was required to achieve at least 90% proficiency with the primary coder on a sample of at least 5 practice studies prior to coding independently.

After initial relevancy screening was completed, approximately one-third (*n* = 146) of all relevant studies were randomly selected for ratings by both the primary and secondary coders during full-text screening. This selection occurred using a random number generator (Haahr, 2022). Interscorer agreement (ISA) was calculated on a trial-by-trial basis for each variable coded. Specifically, the total number of agreements was divided by the total number of agreements plus disagreements and then multiplied by 100. Overall ISA between the primary and secondary coder was 87.9%. When disagreements arose in coding, the primary and secondary coders discussed the discrepancy and reached a consensus to determine the appropriate coding.

Similarly, during the data extraction phase, 33% of included studies (*n* = 32) were randomly selected using a random number generator (Haahr, 2022) for coding by the primary and second coders. ISA was again calculated on a trial-by-trial basis for each variable coded. As with screening, the primary and seconder coders discussed discrepancies and reached a consensus for the appropriate coding. Across all variables, overall ISA was 93.8% for data extraction. On average, interscorer agreement was highest on variables related to assessment practices (100%), followed by participant demographics (95.5%), basic characteristics (94.7%), and intervention practices (89.1%). Within-variable ISA ranged from 80% (contingency format) to 100% on 10 separate variables, including some of those associated with basic study characteristics and participant demographics, as well as all assessment-related variables.

Synthesis Procedures

Basic Study Characteristics (Research Question 1)

As part of this review's results, descriptive information about general characteristics was provided for all studies that meet inclusion criteria. Specifically, information was reported regarding each study's publication year, methodological design, and settings.

Participant Demographics (Research Question 2)

To evaluate the types of participants that have been included in the literature in this area, descriptive data were collected about a variety of participant demographics. This included each study's number of participants, as well as information about participants' ages, grade levels, races, ethnicities, genders, and disability statuses.

Assessment Practices (Research Questions 3 and 4)

To determine what percentage of included studies utilized Can't Do/Won't Do assessments, the total number of studies that included an assessment of performance deficits prior to intervention was calculated and divided by the total number of studies that met inclusion criteria. Similar synthesis procedures were applied to indicate the percentage of studies in the relevant literature that conducted preference and reinforcer assessments, in which the total number of studies that included either preference or reinforcer assessments was divided by the total number of studies included.

Contingency-Based Reinforcement Intervention Practices (Research Questions 5, 6, 7, and 8)

The percentage of studies that used individual versus group contingencies (Research Question 5) was calculated by determining the number of studies that implemented the intervention in group formats divided by the total number of included studies. Of the studies that used a group format, the percentage that included independent, dependent, or interdependent contingencies was also calculated. The number of studies that targeted skills in math, reading, writing, spelling, and Language Arts were individually divided by the total number of studies to demonstrate how frequently each academic subject was targeted using contingency-based reinforcement interventions (Research Question 6). To identify how studies determined reinforcement criteria (Research Question 7), the percentage of studies that used percentile shaping, general improvement, dichotomous variables, or other methods to identify criteria were calculated individually. Finally, the percentage of studies that used randomized or unknown work, criteria, and/or rewards was calculated by identifying and counting which studies included these randomized or unknown components and dividing this value by the overall number of included studies (Research Question 8).

Results

Excluding duplicates (*n* = 1,145), initial search and ancestral review procedures returned a total of 2,961 studies for screening (see Figure 1). A total of 2,628 studies were excluded for irrelevance based on title and abstract screening. An additional 96 studies were identified for full-text screening through an ancestral review. Altogether, 429 studies were screened for inclusion in the full-text review. Another 331 studies were excluded during the full-text, final stage of screening; almost half of these studies (i.e., 44%) were excluded because their intervention did not meet inclusion criteria, oftentimes because contingent-based reinforcement for academic performance was not implemented in isolation from other intervention components. Upon conclusion of both stages of screening, a total of 98 studies met the inclusion criteria for this study.

Basic Study Characteristics

Basic study characteristics are outlined in Table 2. Included studies that examined the use of contingency-based reinforcement interventions for direct academic performance were published from 1969 through 2021. Approximately 14% of these studies were published in the last decade. The majority of these studies used a single-case experimental design methodology, with only 30.6% of studies utilizing group designs, including the review conducted by Speltz (1982), which included elements of both methodologies. Each study's setting was also examined, including the type of school and classroom where the intervention was implemented. Studies were located in the following types of schools: public (48%), private or alternative (4.1%), and laboratory or experimental educational environments (4.1%). However, because a substantial number of studies did not explicitly state the type of school where the intervention was implemented, these data do not reflect the school settings for approximately 43.8% of studies. In contrast, 69.4% of studies explicitly reported the type of classroom where the contingency-based reinforcement intervention was implemented. Interestingly, the majority of the studies that specified classroom type were conducted outside of students' typical classrooms. A total of 23.5% of the contingency-based reinforcement interventions were implemented in locations like empty classrooms or offices. 21.4% of studies occurred in general education classrooms, 18.4% occurred in special education classrooms, and 6.1% of studies occurred in inclusion or integrated classrooms.

Participant Demographics

Participant demographics are reflected in Table 3. There were 3,047 participants across all studies combined, with a range of 1 to 330 participants in a single study (M = 31.1, SD = 47.6). A total of 21 studies did not report on participants' genders; among studies that did report these data, however, there were similar percentages of female (46.4%) and male (53.6%) participants. Participants' ages ranged from 5 to 18, with an average age of 10.13 years old. Although only 61% of included studies reported specific information about participants' ages, 99% of studies provided data about participants' grade levels. Included grades ranged from Kindergarten through twelfth grade. The vast majority of studies were conducted with elementary students (72%), with the remaining studies conducted with middle school students (23%), high school students (3%), or participants across multiple levels of schooling (6%). One study (Dalton et al., 1973) did not report participants' grade levels.

Racial or ethnic data were only reported for 55.7% of participants. Among this group, the vast majority of participants were White (71%). Black or African American participants accounted for 16.6% of the sample. Less than 2% of participants (1.7%) were Hispanic, though this value excludes Mexican-American participants, as indicated below. Individuals with Asian (0.12%) or multiracial (0.24%) backgrounds accounted for less than 1% of participants. Whenever possible, more specific racial or ethnic data were noted, in lieu of the broader categories listed above. As such, participants also reflected the following racial or ethnic backgrounds: Mexican American (4.7%), Cambodian (2.4%), Navajo (1.5%), Hopi (0.4%), Asian Pacific Islander (0.2%), Middle Eastern (0.2%), Puerto Rican (0.1%), and Apache (0.05%).

Approximately 12% of participants had reported disabilities or received special education services. Of these participants, 42% were diagnosed with a learning disability, 16% were diagnosed with an emotional behavioral disorder, and 2% were diagnosed with other health impairments, most frequently in the form of ADHD. The remaining 40% of these participants received special education services for an unspecified disability or a disability outside of the aforementioned categories.

Assessment Practices

Overall, there were relatively few studies that included experimental assessments of participants' skill versus performance deficits (i.e., Can't Do/Won't Do), preferences, or reinforcing items (see Table 4). Only 6.1% of studies assessed for skill versus performance deficits prior to the contingency-based reinforcement intervention. Formal preference assessments, which involved the experimental manipulation of items to identify students' most preferred items, only occurred in approximately 5% of included studies. The use of reinforcer assessments was virtually non-existent in the included literature, as it only appeared in a single study (Hoff, 2020). Hoff (2020) was also the only study to include multiple assessment practices.

Intervention Practices

Contingency Format

For their interventions, 79.6% of studies implemented a single contingency format, while 20.4% of studies examined two or more contingency formats (see Table 5). The substantial majority (i.e., 70%) of studies included at least one condition that examined contingency-based reinforcement interventions in group formats. Among these group-based formats, 79.9% of studies included a treatment condition with an independent contingency, 39.1% included an interdependent contingency, and 17.4% included a dependent contingency. A comparison of all three group contingency formats occurred in six studies (Alric et al., 2007; Hargis, 2012; Lynch et al., 2009; Shapiro & Goldberg, 1986, Speltz et al., 1982).

Academic Subject

All included studies examined math, reading, spelling, writing, or English Language Arts (see Table 5). Math was the most common academic subject studied, as math performance served as an intervention target in 64.3% of studies. Reading was targeted in 28.6% of studies, followed by spelling (13.3% of studies). The least common academic target in contingency-based reinforcement interventions was writing, which was only directly targeted in 7.1% of studies. However, an additional 6.1% of studies targeted English Language Arts, which typically encompassed combined elements of reading, writing, and spelling.

Criterion Selection

Many of the included studies featured more than one criterion selection procedure, typically because disparate contingency formats required different reinforcement criteria (see Table 6). Overall, only 1% of studies used percentile shaping to determine the criterion for reinforcement. In contrast, 29.6% of studies used alternative calculations to determine the criterion. Oftentimes, these alternative calculations involved the integration of participants' previous scores, including the mean or median score of a series of past items. In 7% of studies, students simply needed to improve upon their previous scores to receive a reward. Alternatively, in 29.6% of studies, reinforcement was provided based on a dichotomous variable, in which students received a reward after each desired response on an item. Reinforcement criteria were determined using other methods (e.g., randomized selection, student and teacher choices, competition between students) in 25.5% of studies. In 17.3% of included studies, the method of criterion identification was not specified, or the criterion was not clearly stated.

Randomized and Unknown Components

Randomized or unknown intervention components were used in 36.7% of the contingency-based reinforcement interventions (see Table 6). Rewards were the most common randomized or unknown contingency component used, appearing in 61.1% of the interventions that featured randomized or unknown intervention components. Meanwhile, criteria were randomized or unknown in 58.3% of the studies. Randomized or unknown work—which included cases in which a student did not know which task or items would be scored, or which student's work would be scored—appeared in 50% of the studies with randomized or unknown contingency components. Interestingly, 44.4% of studies with randomized or unknown components used interventions that implemented a combination of randomized or unknown rewards, criteria, and work; in fact, Alosio (2007), Hargis (2012), Landy (2013), Lynch and colleagues (2009), Ralston (2012), Reinhardt and colleagues (2009), Scott and colleagues (2017), and Velazquez (2014) each implemented interventions using all three unknown or randomized elements.

Discussion

In the United States, students are academically underperforming across a variety of academic subjects, including math, reading, and writing (Hussar et al., 2020; NCES, 2011, 2015). Interventions using contingency-based reinforcement have been implemented to address academic underperformance, and past research demonstrates that these interventions have successfully targeted both direct academic performance and related classroom behaviors (e.g., Chadwick & Day, 1971; Park et al., 2019). Although several systematic reviews and meta-analyses have examined the use of contingency-based reinforcement interventions in classrooms (e.g., Bowman-Perrott et al., 2016), the vast majority of these studies examined the impact of contingency-based reinforcement interventions on general classroom behavior (e.g., compliance), rather than academic performance on items and tasks. The literature has also not yet been systematically analyzed with respect to the specific assessment and intervention practices associated with contingency-based reinforcement interventions that have been used in classrooms. Thus, the purpose of this systematic review was to provide a current and comprehensive review of the general characteristics, participant demographics, and assessment and intervention practices present in the studies that have targeted direct academic performance using contingency-based reinforcement interventions.

Key Findings

General Characteristics

The findings of this review demonstrate that the majority of studies that have examined the effects of contingency-based reinforcement interventions on direct academic performance utilized single-case experimental designs. Most commonly, these studies were conducted outside of primary classroom settings in public schools. These results indicate that contingency-based reinforcement interventions were likely being implemented with select students in external settings, rather than integrated into general classroom procedures.

Altogether, however, these data are incomplete; a noticeable trend in the literature is that studies often failed to explicitly report information about the types of schools (e.g., public versus private) where contingency-based reinforcement interventions were implemented. However, these data are valuable because they contextualize the intervention in important ways. For instance, the rewards offered as part of a contingency-based reinforcement intervention are limited by a school's financial resources—resources that often differ as a function of school type. Non-public schools may also have increased flexibility and variability regarding classroom structure, curricular requirements, or typical classroom procedures. Classroom size also tends to differ based on school type; larger class sizes may serve as a barrier to treatment, particularly with independent group contingencies. Because of these potential differences between types of schools, these missing data may have provided essential information about the likelihood that a contingency-based reinforcement intervention would be successful or acceptable in a given location.

Participant Demographics

The included studies' average participant was elementary-aged and White, with no reported disability or use of special education services. The lack of participants with disabilities is particularly noteworthy, given that most interventions were provided to select students outside of their primary classroom. This may indicate that select students participated in contingency-based reinforcement interventions because they required additional academic support, but that contingency-based reinforcement interventions served as an intermediate supportive measure prior to or outside of more intensive programming that might be provided for students with disabilities.

Many studies also failed to report data on participants' racial or ethnic demographics. In part, this is perhaps because some researchers may not view participants' racial or ethnic characteristics as central to their research questions about the effects of contingency-based reinforcement interventions on direct academic performance. Nonetheless, because the majority of the included studies utilized single-case experimental designs, these data would assist with generalization. Moreover, it is impossible to entirely determine the representativeness of the studies' samples because of the missing demographic data about participants' races or ethnicities. However, based on the most recent national demographic data for children ages 5 to 17 years old (de Brey et al., 2019), the results of this review revealed that White participants were overrepresented in the included studies, while Hispanic, Asian, and multiracial individuals were significantly underrepresented. Black students had similar levels of representation among study participants and national demographics.

Notably, similar racial or ethnic categories were used between federal reporting categories (de Brey et al., 2019), most of the included studies, and this systematic review itself. In the included literature as well as this review, racial or ethnic categories generally reflected similar language to that used in the federal reporting categories: "White," "Black," "Hispanic," "Asian," "Pacific Islander," "American Indian/Alaska Native," and "Two or more races." However, in some cases, the federal racial or ethnic categories, as well as those used by many included studies, do not directly coincide with the categories used and reported in this systematic review. Specifically, this review differs from typical federal categorizations because it features more specific racial or ethnic categories whenever possible (e.g., reporting "Navajo" versus "American Indian/Alaska Native"). The decision to divert from federal reporting categories reflects the most recent APA guidelines for bias-free language (APA, 2020).

Assessment Practices

The results of this review suggest that Can't Do/Won't Do, preference, and reinforcer assessments were rarely used as part of contingency-based reinforcement interventions. These assessment practices, however, provide essential data that inform the selection and structure of contingency-based reinforcement interventions. This represents a significant limitation within the field of studies that examined the impact of contingency-based reinforcement interventions, and it affects the way the studies' results can be interpreted.

The assessment of skill versus performance deficits—typically in the form of Can't Do/Won't Do assessments—helps to identify the primary source of academic underperformance in a given area. Previous research has shown that Can't Do/Won't Do assessments can accurately identify whether a student presents with a performance deficit; this information can then be used to determine the most appropriate subsequent intervention (e.g., Duhon et al., 2004; Gansle & Noell, 2002). Because these assessment data were not gathered in the vast majority of included studies, it is likely that contingency-based reinforcement interventions were applied in cases where students needed additional instruction, rather than increased motivation. This mismatch would have led to lost resources and opportunity costs, and it raises alternative interpretations of the studies' results. For example, if a study found that a contingency-based reinforcement intervention did not have a significant impact on a student's academic performance, this may be due to the intervention being generally ineffective or the intervention being applied to an inappropriate case.

Contingency-based reinforcement interventions also rely upon the assumption that a selected "reward" is properly motivating for students. Preference and reinforcer assessments systematically test this assumption. Previous research demonstrates that preference and reinforcer assessments can successfully determine which items are most preferred by students, which can then be used as rewards for contingency-based reinforcement interventions (e.g., Daly et al., 2009; Fritz et al., 2020). Although some of the included studies conducted "informal" preference assessments, in which students suggested rewards items or filled out surveys, this was not done systematically or experimentally. Thus, without the inclusion of more formal assessment practices like preference and reinforcer assessments, the included studies' results have additional potential interpretations. This is particularly true in cases where a contingency-based reinforcement intervention had no significant effect. Again, this may suggest that the intervention is generally ineffective. Alternatively, students' academic performances may have failed to improve because the offered "reward" was not desirable enough to increase motivation and effort.

Intervention Practices

Contingency Format. The findings of this review indicated that most of the studies that have examined the use of contingency-based reinforcement interventions on direct academic performance implemented independent group contingencies. Previous research has demonstrated that independent group contingencies can effectively improve students' academic performances (e.g., Gross et al., 2016; Shapiro & Goldberg, 1986). Given the substantial number of studies that examined this contingency format, there is likely sufficient empirical support to warrant a meta-analysis of the effects of independent group contingencies. This would provide a more robust measure of the contingency format's effectiveness, which may then be used to inform intervention selection in practice. However, because independent group contingencies require greater levels of individualized attention compared to other group contingency formats, the literature base has largely focused on what is likely the least efficient group contingency format (Skinner et al., 2004). As a result, independent group contingencies may have promising empirical support, but the field must also consider and evaluate the feasibility of this contingency format in practice.

It is also noted that, although studies typically featured a singular contingency format, there was a sizeable literature base comparing multiple contingency formats within a singular study. Given the many contingency formats are available for potential intervention selection, the results of these studies may be particularly beneficial, especially when single-case designs were used. The studies that examined the use of a single contingency format may assist in generally answering whether a specific contingency format should be implemented, but the studies comparing multiple formats directly inform *which* contingency format should be implemented.

Academic Subject. A sizeable majority of the included studies focused on math as an academic target. As such, there is likely sufficient evidence to reasonably determine the effects of contingency-based reinforcement interventions on math performance, specifically. Because national data (Hussar et al., 2020) demonstrate that most students in the United States are academically underperforming in math, a meta-analysis of the results of these studies is likely warranted to inform potential intervention selection for math underperformance on a national level. Among the remaining academic subjects, reading served as the most common academic target. Although this subject was targeted less frequently than math, there are likely still enough studies to warrant a meta-analysis of the effect of contingency-based reinforcement interventions on reading performance. This meta-analysis would also have ramifications for intervention selection for the large portion of students in the United States who struggle with reading.

In contrast, the effect of contingency-based reinforcement interventions on writing and spelling skills is still largely unexplored. This is apparent even when the studies that targeted general English Language Arts—which includes specific skills in reading, writing, and spelling—are accounted for. This represents a significant gap in the literature, particularly because rates of math underperformance closely mirror the rates of writing underperformance, which includes the measurement of spelling skills as part of NAEP assessments (Hussar et al., 2020). Evidently, there is a need for evidence-based interventions for writing and spelling, and further research is warranted to determine whether contingency-based reinforcement interventions might meet this demand.

Criterion Selection. Based on the findings of this review, it was apparent that there was considerable variation in the ways that reinforcement criteria were selected and structured in the studies that have examined the impact of contingency-based reinforcement interventions on direct academic performance. Interestingly, many of the criteria used were not based on cumulative goals; rather, reinforcement was frequently provided on an item-by-item basis. This finding is important to note, as this style of reinforcement would typically result in more immediate rewards following each instance of a desired behavior. As such, it is possible that studies using dichotomous reinforcement criteria should be evaluated and analyzed in isolation, given the unique structure of their reinforcement criteria when compared to other cumulative methods.

Many studies also used alternative calculations to determine appropriate reinforcement criteria. These calculations were individual to each study, making it difficult to clearly synthesize what specific calculations were used or how they were selected. The reported use of alternative calculations may assist with the replication of specific studies, but the field would likely benefit from further syntheses of these data. Additional review and potential subsequent meta-analyses would meaningfully determine what calculations are most commonly used and whether specific calculations for criterion identification impact study outcomes.

Beyond item-by-item reinforcement or the use of alternative calculations, a sizeable minority of included studies used other methods of criterion identification, including elements of competition and choice. Given the diverse range of methods that comprised this “other” category, it may be challenging to examine and meta-analyze any one specific method in future research. However, it is possible that these miscellaneous methods may potentially be divided into smaller, more cohesive categories (e.g., choice of criterion) upon further study.

Ultimately, the results of this review highlighted that methods of criterion identification in this literature base were highly variable and lacked standardization compared to other elements of contingency-based reinforcement assessments and interventions. Two reliable and consistent methods of criterion identification (i.e., percentile shaping and general improvement) only appeared in a very small percentage of the literature. There are likely not sufficient empirical bases for either of these methods to be meta-analyzed to determine their effects on a more generalizable scale. In fact, it was more common for studies to entirely omit their criterion selection procedure than to employ either of the aforementioned methods. The lack of clarity, standardization, and generalizability of criterion identification methods represents a significant limitation in the field’s understanding of contingency-based reinforcement interventions. The extent of this limitation should perhaps be explored in future research to determine how much variance in study outcomes is due to differing methods of criterion identification.

Randomized and Unknown Components. The results of this review demonstrated that a substantial portion of included studies used randomized or unknown intervention components.

When an intervention included the use of randomized or unknown intervention components, it often featured a combination of randomized or unknown criteria, work, and rewards. As such, there is likely sufficient empirical evidence to meta-analyze each of these randomized or unknown intervention components to gauge their overall impact.

Reinforcement criteria were often randomized or unknown, which is intriguing given the finding that criterion identification methods were generally imprecise among the included studies. In part, this ambiguity may be explained by the fact that studies often featured multiple potential criteria. A singular calculation or selection method would inherently provide a singular reinforcement criterion, necessitating the use of more flexible methods of criterion identification for interventions that require multiple potential reinforcement criteria. Nonetheless, because the selection methods for a reinforcement criterion (or multiple potential criteria) remain largely undefined, the implications of the recurrent use of randomized or unknown reinforcement criteria remain unclear.

Interestingly, although the majority of included studies used independent group contingencies, elements of students' work were frequently unknown or randomized. Although this review conceptualized unknown or randomized work in two possible ways—students were unsure (a) which student's work would be evaluated or (b) what task would be scored—the prevalence of independent contingencies suggested that task type is the more typical form of unknown or randomized work. This makes sense, given that unknown or randomized student selection is regularly used with dependent contingencies, which were the least-used contingency format in the included studies.

The high prevalence of randomized or unknown rewards is also noteworthy in the context of this review's finding that formal preference and reinforcer assessments were rarely implemented. It is possible that the use of randomized or unknown rewards may somewhat ameliorate concerns that rewards were not sufficiently motivating for students, as contingency-based reinforcement interventions would not rely upon students' preferences for a singular motivator if the reward was unknown or randomized. In future research, it may be beneficial to examine the necessity of formal preference or reinforcer assessments in cases where a contingency-based reinforcement intervention involves randomized or unknown rewards.

Considerations and Methodological Limitations

The results of this systematic review should be viewed in the context of this study's limitations. Although every effort was made to synthesize all of the relevant articles for this review, no search is exhaustive. It is possible that additional search terms would have returned other works for screening and inclusion, particularly as "academic performance" was the only achievement-related search term used. This excluded subject-specific search terms (e.g., "math" or "reading") or other related terms (e.g., "test scores"). Searches that used terms related to preference and reinforcer assessments were also not conducted as part of the study identification procedures. Similarly, searches were limited to the PsycInfo and ERIC databases. While these databases did return some "gray" literature, namely in the form of unpublished theses and dissertations, these results only reflect a portion of the gray literature available. Additional search methods and databases (e.g., ProQuest Dissertations and Theses Global) likely better assist with the retrieval of unpublished literature. As such, the results of this review may be skewed by a publication bias. This risk for potential publication bias is further exacerbated by the predominant inclusion of single-case experimental design studies, given that authors must demonstrate experimental control over participants' academic performances in order to be published.

There is also one noteworthy limitation related to the screening and coding process. Namely, ISA was calculated during the second-stage screening, ancestral review, and data extraction process; it was not, however, calculated during the initial screening for relevancy. This stage of screening was completed solely by the first author. To combat this limitation, the coding manual aimed to provide particularly strict and objective rules for this stage of screening to improve reliability. Nonetheless, it is possible that the initial screening process was less reliable than the subsequent steps.

The findings of this review are limited by the fact that each study's methodological quality was not assessed. However, the exclusion of methodological quality reflects the overall purpose of this systematic review, which was to identify and synthesize the available relevant literature without significant consideration of actual study effects. Indeed, this review likely includes some studies that lack the methodological rigor necessary for confident interpretations of results, replication, or meta-analyses. With this in mind, the applied use of this review to inform intervention or conduct meta-analyses should only occur with adequate consideration of each study's methodological quality and the validity and reliability of its findings.

From a statistical perspective, this review's results are also limited by the integration of both single-case experimental and group design studies, given that each type of study was given equal weight, regardless of its number of participants. That is, this review's results indicated the number of studies that examined each variable, but the extent to which each variable was examined within each study was not reflected. For example, spelling was only examined in a limited number of interventions; however, two studies that did examine spelling (Benowitz & Busse, 1976; Benowitz & Rosenfeld, 1973) featured very large sample sizes relative to the single-case experimental design studies that were included. In this case, spelling may not have been targeted by a large quantity of studies, but it may have been targeted with many *students* within these group-design studies.

Additionally, although this review attempted to isolate contingency-based reinforcement as an intervention, its common elements are sometimes viewed as separate interventions in their own right. In particular, goal setting and performance feedback appear frequently in the included studies as part of contingency-based reinforcement interventions. However, both of these components have been also been conceptualized as independent interventions for study (e.g., Koenig et al., 2016; Moore et al., 2001). Within the included studies themselves, there were contradictory views regarding these intervention components. For example, Carson and Eckert (2003) treated feedback and goal setting as independent interventions and implemented separate intervention conditions for feedback, goal setting, and contingent reinforcement. Duhon and colleagues (2004) did similar work with performance feedback, as they examined baseline, performance feedback, and reward conditions separately. In contrast, one of Dickerson's (1981) interventions began with a conversation between the experimenter and student, in which they established task goals and reinforcement criteria together. This form of contingency-based intervention was integrated with elements of goal setting (as well as another potential confound: choice). Because of the prevalence of goal setting and performance feedback in the literature, however, this review would be overly limited and nonrepresentative in its findings if all studies with these elements were excluded. Nonetheless, because of the disparate ways that these elements are conceptualized, there is some ambiguity regarding whether a given study includes contingency-based reinforcement as an isolated intervention in the purest sense.

Future Directions

This review's findings and limitations provide direction for future research in this area. Namely, future researchers may wish to:

1. **Examine methodological rigor.** While this systematic review synthesized the available literature in this area, the methodological strength of its included studies would provide an important context for its clinical and theoretical applications. While the review's results demonstrated the prevalence of studies that examined the impact of contingency-based reinforcement interventions, it is still unknown whether these studies featured the necessary characteristics to provide reliable and valid data.
2. **Consider evaluating goal setting and/or feedback in isolation as their own treatment conditions during intervention studies.** To best determine whether contingency-based reinforcement effectively targets academic performance, as well as which intervention components are most impactful, researchers may wish to distinguish between goal setting, performance feedback, and contingent reinforcement in their studies. When possible, studies evaluating contingency-based reinforcement interventions may benefit from having separate intervention phases for goal setting and feedback outside of contingent reinforcement.
3. **Explore methods of criterion identification and the impact of these selected methods.** As noted in this review's results, criterion identification procedures were

arguably the most complex and varied aspect of the contingency-based reinforcement intervention literature. In further review, researchers may examine whether there are clear patterns in the calculations and methods used to determine reinforcement criteria. Through intervention studies, future researchers may also consider examining whether certain criterion identification methods have a significant impact on intervention outcomes, and under which contexts they are most beneficial.

4. **Determine the necessity of formal assessments (i.e., Can't Do/Won't Do, preference, and reinforcer) on study outcomes compared to informal methods (e.g., preference surveys).** As this review's findings demonstrated, there is little use of standardized assessment practices to identify performance deficits, preferences, and reinforcers. As such, it may be informative for researchers to examine whether Can't Do/Won't Do assessments should be included as an integral part of contingency-based reinforcement interventions, prior to full implementation of the intervention. Regarding preferences and reinforcers, there seems to be a range of "unofficial methods," (e.g., surveys) that are used in practice. A potential avenue for investigation is whether formal or informal assessments are more effective in identifying rewards that are most motivating for students. Given the prevalence of contingency-based reinforcement interventions in group formats, it may be useful to know whether the potential benefits of individualized formal assessments outweigh the potential costs in time and resources.
5. **Meta-analyze whether the overall intervention, as well as specific assessment or intervention components, significantly affects students' academic performances.**

This systematic review may assist with the early steps of a meta-analysis that will help determine the overall empirical basis for the use of contingency-based reinforcement interventions for direct academic performance. Future researchers may wish to analyze the outcomes of these relevant studies. These analyses may be based upon contingency-based reinforcement interventions in their entirety, or researchers may choose to examine the effects and importance of specific assessment and intervention practices.

Implications

For the last few decades, contingency-based reinforcement interventions have been implemented in classrooms to target and potentially improve students' academic performances. The results of this review demonstrated that, although contingency-based reinforcement interventions are theoretically based upon supported behavioral concepts, many of the applied interventions do not always feature key behavioral concepts in practice. Of course, this partly reflects that behavioral theory and its established practices have evolved since 1969, when the first study was published. Today, however, the field recognizes that there are behavioral practices that assist with intervention selection, structure, and outcomes. Some of these practices—which were often absent in the included literature—include theoretically grounded criterion selection procedures, systematic determination of motivating rewards, and the selection of an appropriate intervention using baseline performance data and identification of specific behavioral needs. The absence of these practices has two potential implications: (1) The gap between the intervention's theoretical grounding and its applied use may be impacting the intervention's effectiveness, and (2) It may be difficult to adhere to strict systematic behavioral principles in practice when these contingency-based reinforcement interventions are used in actual classrooms. Additionally, the use of randomized and unknown elements, as well as the prevalence of group contingency formats, indicates that contingency-based reinforcement interventions may be particularly appealing to researchers and educators when they are designed to serve multiple students at once. Moreover, the studies disproportionately included participants with particular demographics and targeted certain academic subjects. Many races and ethnicities were underrepresented in the study samples compared to true national demographics. Study results also predominantly reflect the intervention's impact on the academic performances of students without disabilities or those who are not receiving special education services. From an academic perspective, the findings of extant literature in this area may apply when math or reading are targeted, but without further study, it is difficult to ascertain whether the intervention would similarly impact academic performance in the areas of writing, spelling, or English Language Arts.

Conclusion

Contingency-based reinforcement interventions have been used to target students' classroom skills through the use of positive reinforcement and rewards for desired behaviors. This systematic review gathered and synthesized the literature surrounding contingency-based reinforcement interventions that examined the impact of contingency-based reinforcement interventions on direct academic performance. The results indicated that the majority of studies occurred in public school settings with White, elementary-aged participants who were not diagnosed with any disabilities. Can't Do/Won't Do assessments were occasionally used to inform intervention selection, but formal preference and reinforcer assessments were rarely implemented. Most interventions occurred in independent group formats, and very few studies featured dependent group contingencies. Math and reading skills were commonly targeted, while spelling, writing, and English Language Arts appeared less frequently. A substantial set of studies included randomized or unknown components, most often in the form of unknown or randomized rewards. Criterion identification practices varied widely and may be a particularly informative area of future research. Meta-analyses and studies focusing on assessment practices or understudied academic areas are also potential avenues of further research.

Table 1

Systematic Reviews and Meta-Analyses on Contingency-Based Reinforcement Interventions Targeting Academic Performance and Behaviors

Study	Target Behaviors	Administration Format	Reported Use of Can't Do/Won't Do Assessments	Reported use of Preference or Reinforcer Assessments	Reported Reinforcement Criteria	Discussed Randomized or Unknown Intervention Components
Bowman-Perrott et al. (2016)	Behavioral	Group	No	No	No	No
Cerasoli et al. (2014)	Academic and Behavioral	Not Reported	No	No	Yes	No
Chaffee et al. (2017)	Behavioral	Group	No	No	No	No
Little et al. (2015)	Academic and Behavioral	Group	No	No	No	No
Long et al. (2019)	Behavioral	Individual and Group	No	No	No	No
Maggin et al. (2012)	Behavioral	Group	No	No	No	No
Maggin et al. (2017)	Behavioral	Group	No	No	No	No

Study	Target Behaviors	Administration Format	Reported Use of Can't Do/Won't Do Assessments	Reported use of Preference or Reinforcer Assessments	Reported Reinforcement Criteria	Discussed Randomized or Unknown Intervention Components
Pokorski et al. (2017)	Behavioral	Group	No	Yes	No	Yes
Stage & Quiroz (1997)	Behavioral	Individual and Group	No	No	No	No
Warmbold-Brann et al. (2017)	Behavioral	Individual and Group	No	No	No	No

Table 2

Settings and Design of Included Studies (*n* = 98)

Study	School Type	Classroom Setting	Design
Aloisio (2007)	Public	General Education	Single-Case
Alric et al. (2007)	Public	Not Specified	Single-Case
Amado (1982)	Public	Special Education	Single-Case
Andersen et al. (2013)	Not Specified	Outside Main Classroom	Single-Case
Ardi (1989)	Public	Special Education	Group
Auge (2021)	Not Specified	Not Specified	Single-Case
Ayllon et al. (1972)	Public	Special Education	Single-Case
Ayllon et al. (1975)	Alternative/Private	Special Education	Single-Case
Ayllon & Roberts (1974)	Public	Not Specified	Single-Case
Baker (2013)	Public	Special Education	Single-Case
Bear & Richards (1980)	Alternative/Private	Not Specified	Single-Case
Bennett (2006)	Public	Inclusion/Integrated	Single-Case
Benowitz & Busse (1976)	Not Specified	Not Specified	Group
Benowitz & Rosenfeld (1973)	Not Specified	Not Specified	Group
Billingsley (1977)	Laboratory or EEE	Outside Main Classroom	Single-Case
Bit (1981)	Public	General Education	Group
Brigham et al. (1972)	Public	Not Specified	Single-Case
Brooks & Snow (1972)	Not Specified	General Education	Single-Case

Study	School Type	Classroom Setting	Design
Broughton (1983)	Not Specified	General Education	Single-Case
Broughton & Lahey (1978)	Not Specified	Not Specified	Group
Brownell et al. (1977)	Public	Outside Main Classroom	Group
Carlson et al. (1970)	Public	General Education	Group
Carson & Eckert (2003)	Public	General Education & Inclusion/Integrated	Single-Case
Christ, & Schanding (2007)	Not Specified	Not Specified	Group
Coleman (1970)	Public	Special Education	Single-Case
Cracco (2006)	Not Specified	Special Education	Single-Case
Dalton et al. (1973)	Not Specified	Special Education	Group
Daly et al. (2005)	Not Specified	General Education & Outside Main Classroom	Single-Case
Denison (2013)	Public	Outside Main Classroom	Single-Case
Dickerson & Creedon (1981)	Public	Outside Main Classroom	Group
Diedrick (2009)	Not Specified	Special Education	Single-Case
Doherty (1981)	Public	Outside Main Classroom	Group
Duhon et al. (2004)	Not Specified	General Education & Outside Main Classroom	Single-Case
Eckert et al. (2000)	Not Specified	Outside Main Classroom	Single-Case
Felixbrod & O'Leary (1973)	Public	Outside Main Classroom	Group
Felixbrod & O'Leary (1974)	Public	Outside Main Classroom	Group

Study	School Type	Classroom Setting	Design
Ferritor et al. (1972)	Not Specified	Outside Main Classroom	Single-Case
Fowler et al. (1977)	Not Specified	Not Specified	Single-Case
Freeland & Noell (1999)	Not Specified	Outside Main Classroom	Single-Case
Gilbert (1990)	Public	Not Specified	Group
Hardy (2014)	Not Specified	General Education	Single-Case
Hargis (2012)	Public	Inclusion/Integrated	Single-Case
Harris & Sherman (1974)	Not Specified	Not Specified	Single-Case
Hauserman & McIntire (1969)	Not Specified	Not Specified	Group
Hay et al. (1977)	Public	Not Specified	Group
Hoff (2020)	Public	General Education & Outside Main Classroom	Single-Case
Hofstadter-Duke & Daly (2015)	Alternative/Private	Outside Main Classroom	Single-Case
Holt (1971)	Laboratory or EEE	Not Specified	Single-Case
Hopkins et al. (1971)	Not Specified	Not Specified	Single-Case
Kirby & Shields (1972)	Laboratory or EEE	General Education	Single-Case
Lahey et al. (1973)	Public	Outside Main Classroom	Single-Case
Landy (2013)	Public	Inclusion/Integrated	Single-Case
Little et al. (2010)	Not Specified	General Education	Single-Case
Lynch et al. (2009)	Public	Special Education	Single-Case
Madaus et al. (2003)	Not Specified	General Education	Single-Case
Madsen & Forsythe (1973)	Public	Not Specified	Group

Study	School Type	Classroom Setting	Design
Marholin & Steinman (1977)	Public	Special Education	Single-Case
McEvoy & Brady (1988)	Public	Special Education	Single-Case
McLaughlin et al. (1980)	Not Specified	Special Education	Single-Case
Melton (1970)	Not Specified	Not Specified	Group
Metallo (2015)	Public	General Education	Single-Case
Noell et al. (2001)	Public	Outside Main Classroom	Single-Case
O'Connor & Daly (2018)	Public	Outside Main Classroom	Single-Case
Panahon & Martens (2013)	Public	Outside Main Classroom	Single-Case
Pipkin et al. (2007)	Not Specified	Outside Main Classroom	Single-Case
Ralston (2012)	Public	General Education	Single-Case
Reinhardt et al. (2009)	Public	General Education	Single-Case
Reisener (2009)	Public	Outside Main Classroom	Single-Case
Reiss et al. (1974)	Public	General Education	Single-Case
Rieth et al. (1977)	Not Specified	Special Education	Single-Case
Robin (2014)	Not Specified	General Education	Single-Case
Rosenberg et al. (1985)	Not Specified	Not Specified	Group
Rosenfeld (1972)	Alternative/Private	Not Specified	Single-Case
Ross & Braden (1991)	Public	Special Education	Group
Rothberg (1973)	Public	Not Specified	Group
Schellenberg et al. (1991)	Public	Not Specified	Single-Case
Schunk (1983)	Not Specified	Outside Main Classroom	Group

Study	School Type	Classroom Setting	Design
Schunk (1984)	Not Specified	Outside Main Classroom	Group
Scott et al. (2017)	Public	General Education	Single-Case
Shapiro & Goldberg (1986)	Public	General Education	Single-Case
Sharp & Skinner (2004)	Not Specified	Not Specified	Single-Case
Simon et al. (1982)	Public	Not Specified	Single-Case
Slavin (1978)	Not Specified	Not Specified	Group
Speltz et al. (1982)	Laboratory or EEE	Special Education	Group & Single-Case
Strandy et al. (1979)	Not Specified	Special Education	Single-Case
Suter (1993)	Public	Outside Main Classroom	Group
Swain & McLaughlin (1998)	Not Specified	Special Education	Single-Case
Taffel & O'Leary (1976)	Not Specified	Outside Main Classroom	Group
Terry et al. (1978)	Not Specified	General Education	Single-Case
Theodore et al. (2009)	Public	Inclusion/Integrated	Single-Case
Thomson & Galloway (1970)	Not Specified	Not Specified	Group
Turco & Elliott (1990)	Public	Outside Main Classroom	Single-Case
Velazquez (2014)	Not Specified	Inclusion/Integrated	Single-Case
Weekley (1980)	Not Specified	Not Specified	Single-Case
Wilder (2011)	Public	Inclusion/Integrated	Single-Case
Winn (2005)	Not Specified	Not Specified	Single-Case
Wodarski et al. (1973)	Not Specified	Not Specified	Group
Yarbrough et al. (1977)	Not Specified	Not Specified	Group

Note. EEE = Experimental Education Environment.

Table 3

Participant Gender, Age, Grade, and Racial or Ethnic Demographics

Study	*n*	*n* Female	Age	Grade	*n* Disability	*n* Race or Ethnicity
Aloisio (2007)	6	3	9-11	4	0	H: 5, B: 1
Alric et al. (2007)	8	NR	NR	4	0	NR
Amado (1982)	4	2	10-11	5-6	Alt: 4	B: 4
Andersen et al. (2013)	6	1	NR	3	0	W: 6
Ardi (1989)	15	3	9-12	4-6	EBD: 15	H: 2, W: 9, B: 2*
Auge (2021)	39	21	NR	1-2	0	NR
Ayllon et al. (1972)	4	0	12-13	E	Alt: 4	W: 2, B: 2
Ayllon et al. (1975)	3	1	8-10	M	SLD: 3	NR
Ayllon & Roberts (1974)	5	0	NR	5	0	NR
Baker (2013)	12	1	11-15	6-8	EBD, SLD: 12 Alt: 5	W: 2, B: 7
Bear & Richards (1980)	10	NR	11-15	5-8	0	NR
Bennett (2006)	6	3	7	2	Alt: 3	NR
Benowitz & Busse (1976)	330	170	NR	4	0	W: 264, B: 66
Benowitz & Rosenfeld (1973)	174	96	NR	4	0	W: 165, B: 9
Billingsley (1977)	8	0	9-12	E	0	NR
Bit (1981)	60	26	6-16	E - HS	0	Cambodian: 60
Brigham et al. (1972)	6	3	5	K	0	NR
Brooks & Snow (1972)	1	0	9	E	0	NR
Broughton (1983)	6	3	NR	4	0	NR

Study	*n*	*n* Female	Age	Grade	*n* Disability	*n* Race or Ethnicity
Broughton & Lahey (1978)	33	NR	NR	4-5	0	NR
Brownell et al. (1977)	40	24	NR	3-4	0	NR
Carlson et al. (1970)	22	9	8-9	3	0	NR
Carson & Eckert (2003)	3	2	9-10	4	0	W: 1, B:2
Christ, & Schanding (2007)	90	49	NR	2-5	Alt: 13	B: 55*
Coleman (1970)	1	0	8	E	Alt :1	NR
Cracco (2006)	11	3	11-12	6	SLD: 5, OHI: 2, Alt: 4	H: 1, W: 10
Dalton et al. (1973)	13	7	6-13	NR	Alt: 13	NR
Daly et al. (2005)	2	1	9-10	4-5	0	W: 2
Denison (2013)	4	1	NR	3	0	W: 4
Dickerson & Creedon (1981)	30	11	NR	2-3	0	NR
Diedrick (2009)	6	2	12-14	6-8	EBD: 1, SLD: 4, OHI: 1	W: 6
Doherty (1981)	36	16	7-9	3	0	W: 29, B: 6, A: 1
Duhon et al. (2004)	4	0	8-10	3-5	0	B: 4
Eckert et al. (2000)	4	0	7-8	E	0	NR
Felixbrod & O'Leary (1973)	24	12	NR	2	0	NR
Felixbrod & O'Leary (1974)	24	12	8-9	3	0	W: 24*
Ferritor et al. (1972)	23	NR	NR	3	0	NR
Fowler et al. (1977)	1	0	11	6	0	NR
Freeland & Noell (1999)	4	1	9-10	4	0	NR
Gilbert (1990)	74	39	8-10	3	0	NR

Study	*n*	*n* Female	Age	Grade	*n* Disability	*n* Race or Ethnicity
Hardy (2014)	4	0	8-10	3-4	EBD: 2	H: 1, W: 2, B: 1
Hargis (2012)	17	7	13-14	8	SLD: 10	NR
Harris & Sherman (1974)	25	NR	NR	6	0	NR
Hauserman & McIntire (1969)	12	NR	NR	1-2	0	NR
Hay et al. (1977)	10	0	NR	2-4	0	NR
Hoff (2020)	2	0	9-10	3-4	Alt: 1	NR
Hofstadter-Duke & Daly (2015)	3	3	8-11	E	Alt: 1	W: 3
Holt (1971)	21	10	NR	1	0	NR
Hopkins et al. (1971)	24	NR	NR	1-2	0	NR
Kirby & Shields (1972)	1	0	13	7	0	NR
Lahey et al. (1973)	4	1	11	6	0	W: 3, B: 1
Landy (2013)	20	10	12-12	7	Alt: 12	NR
Little et al. (2010)	65	27	NR	4	SLD: 6	H: 4, W: 34, B: 26, A:1
Lynch et al. (2009)	6	4	10-11	5	SLD: 4, Alt: 2	NR
Madaus et al. (2003)	5	1	NR	5	0	NR
Madsen & Forsythe (1973)	88	NR	NR	6	0	NR
Marholin & Steinman (1977)	8	4	10-12	5-6	Alt: 8	NR
McEvoy & Brady (1988)	5	2	6-9	E	EBD: 3, Alt: 2	NR
McLaughlin et al. (1980)	10	NR	10-12	E	EBD: 10	NR
Melton (1970)	37	23	NR	2	0	NR
Metallo (2015)	6	4	NR	6	0	NR
Noell et al. (2001)	4	2	7-8	2-3	0	NR
O'Connor & Daly (2018)	4	3	NR	1, 3	0	B: 1, Middle Eastern: 3

Study	*n*	*n* Female	Age	Grade	*n* Disability	*n* Race or Ethnicity
Panahon & Martens (2013)	3	0	9-11	4	0	H: 1, B: 2
Pipkin et al. (2007)	2	1	5	K	0	NR
Ralston (2012)	50	NR	NR	M	0	W: 48*
Reinhardt et al. (2009)	6	3	9-11	4	0	H: 5, B: 1
Reisener (2009)	4	2	10-12	4, 6	SLD: 1	W: 2, B: 2
Reiss et al. (1974)	22	NR	NR	3	0	NR
Rieth et al. (1977)	3	0	7-11	E, 1, 4	EBD: 2, Alt: 1	NR
Robin (2014)	20	12	14-16	10	OHI: 5	H: 3, W: 17
Rosenberg et al. (1985)	44	NR	8-12	E	Alt: 44	NR
Rosenfeld (1972)	60	NR	NR	6	0	W: 60
Ross & Braden (1991)	94	32	NR	1-5	SLD: 94	W: 68, B: 23*
Rothberg (1973)	140	NR	NR	5	0	W: 61, Mexican-American: 79
Schellenberg et al. (1991)	100	NR	17-18	12	0	NR
Schunk (1983)	33	20	9-11	E	0	NR
Schunk (1984)	36	25	8-10	E	0	NR
Scott et al. (2017)	16	10	6-7	1	0	H: 1, W: 14, B: 1
Shapiro & Goldberg (1986)	53	28	NR	6	0	NR
Sharp & Skinner (2004)	14	5	7-9	2	0	B: 14
Simon et al. (1982)	7	2	NR	M	Alt: 7	NR
Slavin (1978)	205	NR	NR	7	0	W: 202*
Speltz et al. (1982)	12	4	7-10	E	SLD: 12	NR
Strandy et al. (1979)	6	NR	15-17	HS	Alt: 6	NR

Study	*n*	*n* Female	Age	Grade	*n* Disability	*n* Race or Ethnicity
Suter (1993)	40	4	8-11	E	EBD: 13	H: 1, W: 38, B: 1
Swain & McLaughlin (1998)	4	NR	13-14	M	Alt: 4	NR
Taffel & O'Leary (1976)	48	24	NR	4	0	W: 31, B: 15, Puerto Rican: 2
Terry et al. (1978)	1	0	NR	4	0	NR
Theodore et al. (2009)	21	8	9-10	4	Alt: 6	H: 3, W: 14, Asian-Pacific Islander: 4
Thomson & Galloway (1970)	91	NR	8-14	E	0	NR
Turco & Elliott (1990)	74	35	NR	5	0	W: 58, B: 16
Velazquez (2014)	17	7	10-12	5	SLD: 2	W: 17
Weekley (1980)	37	11	12-14	7	0	W: 3, M:1, Navajo: 25, Hopi: 7, Apache: 1
Wilder (2011)	44	21	6-8	1	0	NR
Winn (2005)	23	8	9-10	4	0	H: 1, W: 3, B: 19
Wodarski et al. (1973)	94	NR	NR	5	0	NR
Yarbrough et al. (1977)	90	NR	NR	5	0	NR

Note. NR = Not Reported. K = Kindergarten; E = General Elementary; M = General Middle or Junior High; HS = General High School. EBD = Emotional or Behavioral Disorder; SLD = Specific Learning Disability; OHI = Other Health Impairment; Alt = Alternative disability that does not fall under EBD, SLD, or OHI, or is otherwise not specified. H = Hispanic or Latino/Latina/Latine; W = White; B = Black or African American; A = Asian; M = Multiracial. * = Not all participants are accounted for.

Table 4

Assessment Practices in Contingency-Based Reinforcement Academic Interventions

Study	Can't Do/Won't Do	Preference Assessment	Reinforcer Assessment
Aloisio (2007)	No	No	No
Alric et al. (2007)	No	No	No
Amado (1982)	No	No	No
Andersen et al. (2013)	Yes	No	No
Ardi (1989)	No	No	No
Auge (2021)	No	No	No
Ayllon et al. (1972)	No	No	No
Ayllon et al. (1975)	No	No	No
Ayllon & Roberts (1974)	No	No	No
Baker (2013)	No	No	No
Bear & Richards (1980)	No	No	No
Bennett (2006)	No	No	No
Benowitz & Busse (1976)	No	No	No
Benowitz & Rosenfeld (1973)	No	No	No
Billingsley (1977)	No	No	No
Bit (1981)	No	No	No
Brigham et al. (1972)	No	No	No
Brooks & Snow (1972)	No	No	No

Study	Can't Do/Won't Do	Preference Assessment	Reinforcer Assessment
Broughton (1983)	No	Yes	No
Broughton & Lahey (1978)	No	No	No
Brownell et al. (1977)	No	No	No
Carlson et al. (1970)	No	No	No
Carson & Eckert (2003)	Yes	No	No
Christ, & Schanding (2007)	No	No	No
Coleman (1970)	No	No	No
Cracco (2006)	No	No	No
Dalton et al. (1973)	No	No	No
Daly et al. (2005)	No	No	No
Denison (2013)	Yes	No	No
Dickerson & Creedon (1981)	No	No	No
Diedrick (2009)	No	No	No
Doherty (1981)	No	No	No
Duhon et al. (2004)	Yes	No	No
Eckert et al. (2000)	No	No	No
Felixbrod & O'Leary (1973)	No	No	No
Felixbrod & O'Leary (1974)	No	No	No
Ferritor et al. (1972)	No	No	No
Fowler et al. (1977)	No	No	No
Freeland & Noell (1999)	No	No	No

Study	Can't Do/Won't Do	Preference Assessment	Reinforcer Assessment
Gilbert (1990)	No	No	No
Hardy (2014)	No	No	No
Hargis (2012)	No	No	No
Harris & Sherman (1974)	No	No	No
Hauserman & McIntire (1969)	No	No	No
Hay et al. (1977)	No	No	No
Hoff (2020)	Yes	Yes	Yes
Hofstadter-Duke & Daly (2015)	No	No	No
Holt (1971)	No	No	No
Hopkins et al. (1971)	No	No	No
Kirby & Shields (1972)	No	No	No
Lahey et al. (1973)	No	No	No
Landy (2013)	No	No	No
Little et al. (2010)	No	No	No
Lynch et al. (2009)	No	No	No
Madaus et al. (2003)	No	No	No
Madsen & Forsythe (1973)	No	No	No
Marholin & Steinman (1977)	No	No	No
McEvoy & Brady (1988)	No	No	No
McLaughlin et al. (1980)	No	No	No
Melton (1970)	No	No	No

Study	Can't Do/Won't Do	Preference Assessment	Reinforcer Assessment
Metallo (2015)	No	No	No
Noell et al. (2001)	No	No	No
O'Connor & Daly (2018)	No	Yes	No
Panahon & Martens (2013)	No	Yes	No
Pipkin et al. (2007)	No	No	No
Ralston (2012)	No	No	No
Reinhardt et al. (2009)	No	No	No
Reisener (2009)	Yes	No	No
Reiss et al. (1974)	No	No	No
Rieth et al. (1977)	No	No	No
Robin (2014)	No	No	No
Rosenberg et al. (1985)	No	No	No
Rosenfeld (1972)	No	No	No
Ross & Braden (1991)	No	No	No
Rothberg (1973)	No	No	No
Schellenberg et al. (1991)	No	No	No
Schunk (1983)	No	No	No
Schunk (1984)	No	No	No
Scott et al. (2017)	No	No	No
Shapiro & Goldberg (1986)	No	No	No
Sharp & Skinner (2004)	No	No	No

Study	Can't Do/Won't Do	Preference Assessment	Reinforcer Assessment
Simon et al. (1982)	No	No	No
Slavin (1978)	No	No	No
Speltz et al. (1982)	No	No	No
Strandy et al. (1979)	No	No	No
Suter (1993)	No	No	No
Swain & McLaughlin (1998)	No	No	No
Taffel & O'Leary (1976)	No	Yes	No
Terry et al. (1978)	No	No	No
Theodore et al. (2009)	No	No	No
Thomson & Galloway (1970)	No	No	No
Turco & Elliott (1990)	No	No	No
Velazquez (2014)	No	No	No
Weekley (1980)	No	No	No
Wilder (2011)	No	No	No
Winn (2005)	No	No	No
Wodarski et al. (1973)	No	No	No
Yarbrough et al. (1977)	No	No	No

Table 5

Contingency Formats and Academic Subjects of Contingency-Based Reinforcement Interventions

Study	Contingency Format	Academic Subject
Aloisio (2007)	Interdependent	Math; Reading; Spelling
Alric et al. (2007)	Independent; Dependent; Interdependent	Reading
Amado (1982)	Independent; Interdependent	Reading
Andersen et al. (2013)	Individual	Reading
Ardi (1989)	Independent	Reading
Auge (2021)	Interdependent	Writing
Ayllon et al. (1972)	Independent	Math; Reading
Ayllon et al. (1975)	Independent	Math; Reading
Ayllon & Roberts (1974)	Independent	Reading
Baker (2013)	Independent; Interdependent	Reading
Bear & Richards (1980)	Interdependent	Math; English or Language Arts
Bennett (2006)	Independent	Spelling
Benowitz & Busse (1976)	Independent	Spelling
Benowitz & Rosenfeld (1973)	Independent	Spelling
Billingsley (1977)	Individual	Reading
Bit (1981)	Independent; Interdependent	Math
Brigham et al. (1972)	Independent	Writing
Brooks & Snow (1972)	Individual	Math

Study	Contingency Format	Academic Subject
Broughton (1983)	Independent	Math
Broughton & Lahey (1978)	Independent	Math
Brownell et al. (1977)	Individual	Math
Carlson et al. (1970)	Individual	Reading
Carson & Eckert (2003)	Individual	Math
Christ, & Schanding (2007)	Independent	Math
Coleman (1970)	Individual	Math
Cracco (2006)	Interdependent	Math
Dalton et al. (1973)	Independent	Math; English or Language Arts
Daly et al. (2005)	Individual	Reading
Denison (2013)	Individual	Math
Dickerson & Creedon (1981)	Individual	Math; Writing
Diedrick (2009)	Independent	Math
Doherty (1981)	Individual	Math
Duhon et al. (2004)	Individual	Math; Reading; Writing
Eckert et al. (2000)	Individual	Reading
Felixbrod & O'Leary (1973)	Individual	Math
Felixbrod & O'Leary (1974)	Individual	Math
Ferritor et al. (1972)	Independent	Math
Fowler et al. (1977)	Individual	Math; English or Language Arts
Freeland & Noell (1999)	Individual	Math

Study	Contingency Format	Academic Subject
Gilbert (1990)	Independent; Interdependent	Spelling
Hardy (2014)	Interdependent	Math
Hargis (2012)	Independent; Dependent; Interdependent	Math
Harris & Sherman (1974)	Independent	Math
Hauserman & McIntire (1969)	Individual	Reading
Hay et al. (1977)	Independent	Math; Reading
Hoff (2020)	Individual	Reading
Hofstadter-Duke & Daly (2015)	Individual	Math
Holt (1971)	Independent	Math; Reading
Hopkins et al. (1971)	Independent	Writing
Kirby & Shields (1972)	Individual	Math
Lahey et al. (1973)	Independent	Reading
Landy (2013)	Dependent; Interdependent	Math
Little et al. (2010)	Interdependent	Math
Lynch et al. (2009)	Independent; Dependent; Interdependent	Math; Reading; Writing
Madaus et al. (2003)	Independent	Math
Madsen & Forsythe (1973)	Independent	Math
Marholin & Steinman (1977)	Independent	Math
McEvoy & Brady (1988)	Independent	Math
McLaughlin et al. (1980)	Independent; Interdependent	Spelling
Melton (1970)	Independent	Spelling

Study	Contingency Format	Academic Subject
Metallo (2015)	Independent; Dependent	Math
Noell et al. (2001)	Individual	Reading
O'Connor & Daly (2018)	Individual	Math
Panahon & Martens (2013)	Individual	Math
Pipkin et al. (2007)	Individual	Reading
Ralston (2012)	Dependent	Math
Reinhardt et al. (2009)	Interdependent	Math; Reading; Spelling
Reisener (2009)	Individual	Math
Reiss et al. (1974)	Independent	Math
Rieth et al. (1977)	Individual	Reading; English or Language Arts
Robin (2014)	Dependent; Interdependent	Math
Rosenberg et al. (1985)	Independent	Math
Rosenfeld (1972)	Independent	Math
Ross & Braden (1991)	Independent	Math
Rothberg (1973)	Independent; Interdependent	Reading; Spelling
Schellenberg et al. (1991)	Independent	English or Language Arts
Schunk (1983)	Independent	Math
Schunk (1984)	Independent	Math
Scott et al. (2017)	Dependent; Interdependent	Math
Shapiro & Goldberg (1986)	Independent; Dependent; Interdependent	Spelling
Sharp & Skinner (2004)	Interdependent	Reading

Study	Contingency Format	Academic Subject
Simon et al. (1982)	Independent	Math
Slavin (1978)	Independent; Interdependent	English or Language Arts
Speltz et al. (1982)	Independent; Dependent; Interdependent	Math
Strandy et al. (1979)	Independent	Math; Reading
Suter (1993)	Independent	Math
Swain & McLaughlin (1998)	Independent	Math
Taffel & O'Leary (1976)	Individual	Math
Terry et al. (1978)	Individual	Math
Theodore et al. (2009)	Independent	Spelling
Thomson & Galloway (1970)	Independent	Spelling
Turco & Elliott (1990)	Independent; Dependent	Spelling
Velazquez (2014)	Dependent; Interdependent	Math
Weekley (1980)	Independent; Interdependent	Math
Wilder (2011)	Interdependent	Reading
Winn (2005)	Independent	Writing
Wodarski et al. (1973)	Independent; Interdependent	Math
Yarbrough et al. (1977)	Independent	Math

Table 6

Criteria Selection and the Use of Unknown or Randomized Components in Contingency-Based Reinforcement Interventions

Study	Criteria Selection	Unknown or Randomized Components
Aloisio (2007)	Other	Work; Criteria; Reward
Alric et al. (2007)	Other	Work; Reward
Amado (1982)	Alternate Calculation; General Improvement	None
Andersen et al. (2013)	Alternate Calculation	Work
Ardi (1989)	Not Specified	None
Auge (2021)	Other	Work; Criteria
Ayllon et al. (1972)	Dichotomy	None
Ayllon et al. (1975)	Dichotomy	None
Ayllon & Roberts (1974)	Not Specified	None
Baker (2013)	Not Specified	Reward
Bear & Richards (1980)	General Improvement	None
Bennett (2006)	Not Specified	Reward
Benowitz & Busse (1976)	General Improvement	None
Benowitz & Rosenfeld (1973)	General Improvement	None
Billingsley (1977)	Other	None
Bit (1981)	Alternate Calculation; Other	None
Brigham et al. (1972)	Dichotomy	None
Brooks & Snow (1972)	Not Specified	None

Study	Criterion Selection	Unknown or Randomized Components
Broughton (1983)	Dichotomy	None
Broughton & Lahey (1978)	Dichotomy	None
Brownell et al. (1977)	Other	None
Carlson et al. (1970)	Alternate Calculation	None
Carson & Eckert (2003)	Alternate Calculation	None
Christ, & Schanding (2007)	General Improvement	None
Coleman (1970)	Dichotomy	Criteria
Cracco (2006)	Other	Criteria; Reward
Dalton et al. (1973)	Dichotomy	Criteria
Daly et al. (2005)	Alternate Calculation	None
Denison (2013)	Alternate Calculation	None
Dickerson & Creedon (1981)	Other	None
Diedrick (2009)	Dichotomy	None
Doherty (1981)	Other	None
Duhon et al. (2004)	Percentile Shaping	None
Eckert et al. (2000)	Not Specified	None
Felixbrod & O'Leary (1973)	Alternate Calculation	None
Felixbrod & O'Leary (1974)	Alternate Calculation; Other	None
Ferritor et al. (1972)	Dichotomy	None
Fowler et al. (1977)	Not Specified	None
Freeland & Noell (1999)	Alternate Calculation	Work; Criteria

Study	Criterion Selection	Unknown or Randomized Components
Gilbert (1990)	Alternate Calculation; Other	None
Hardy (2014)	Alternate Calculation	Reward
Hargis (2012)	Other	Work; Criteria; Reward
Harris & Sherman (1974)	Other	None
Hauserman & McIntire (1969)	Alternate Calculation; Dichotomy	None
Hay et al. (1977)	Dichotomy	None
Hoff (2020)	Alternate Calculation	Work; Criteria
Hofstadter-Duke & Daly (2015)	Dichotomy	None
Holt (1971)	Not Specified	None
Hopkins et al. (1971)	Dichotomy	None
Kirby & Shields (1972)	Alternate Calculation	None
Lahey et al. (1973)	Dichotomy	None
Landy (2013)	Other	Work; Criteria; Reward
Little et al. (2010)	Alternate Calculation	Reward
Lynch et al. (2009)	Alternate Calculation	Work; Criteria; Reward
Madaus et al. (2003)	Not Specified	Reward
Madsen & Forsythe (1973)	Not Specified	Criteria
Marholin & Steinman (1977)	Alternate Calculation	None
McEvoy & Brady (1988)	Other	None
McLaughlin et al. (1980)	Dichotomy	None
Melton (1970)	Dichotomy	None

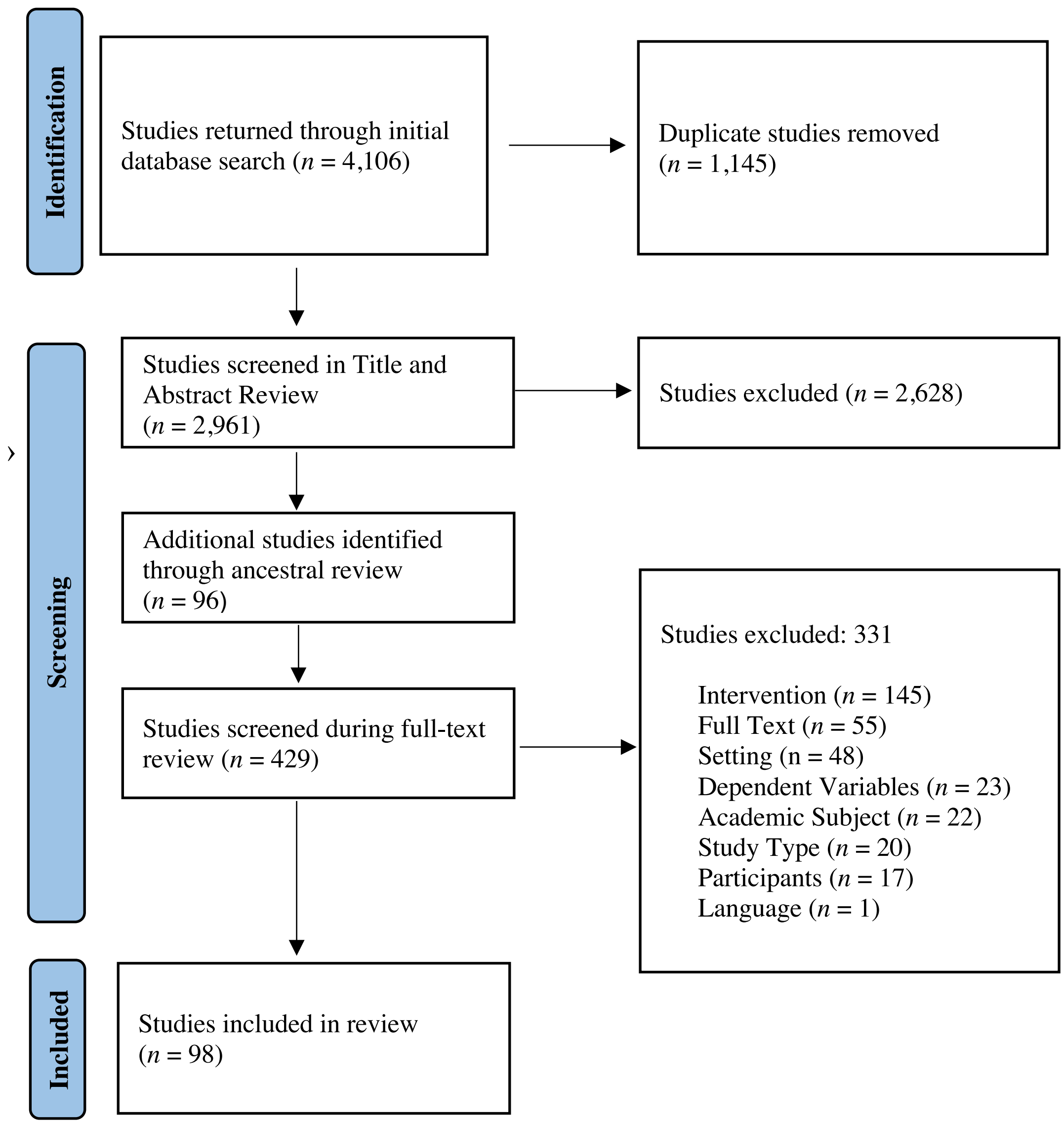
Study	Criterion Selection	Unknown or Randomized Components
Metallo (2015)	Not Specified	Work
Noell et al. (2001)	Alternate Calculation	None
O'Connor & Daly (2018)	Alternate Calculation	Criteria; Reward
Panahon & Martens (2013)	Alternate Calculation	None
Pipkin et al. (2007)	General Improvement	Criteria
Ralston (2012)	Other	Work; Criteria; Reward
Reinhardt et al. (2009)	Other	Work; Criteria; Reward
Reisener (2009)	Alternate Calculation	None
Reiss et al. (1974)	Alternate Calculation	None
Rieth et al. (1977)	Dichotomy	None
Robin (2014)	Other	Work; Reward
Rosenberg et al. (1985)	Dichotomy	None
Rosenfeld (1972)	Dichotomy	None
Ross & Braden (1991)	Dichotomy	None
Rothberg (1973)	Alternate Calculation; Dichotomy	None
Schellenberg et al. (1991)	Dichotomy	None
Schunk (1983)	Alternate Calculation	Reward
Schunk (1984)	Dichotomy	None
Scott et al. (2017)	Other	Work; Criteria; Reward
Shapiro & Goldberg (1986)	Not Specified	Work
Sharp & Skinner (2004)	Dichotomy; Other	Criteria

Study	Criterion Selection	Unknown or Randomized Components
Simon et al. (1982)	Dichotomy	None
Slavin (1978)	Other	None
Speltz et al. (1982)	Dichotomy	Work
Strandy et al. (1979)	Not Specified	None
Suter (1993)	Dichotomy	None
Swain & McLaughlin (1998)	Not Specified	None
Taffel & O'Leary (1976)	Alternate Calculation	None
Terry et al. (1978)	Not Specified	None
Theodore et al. (2009)	Other	Criteria; Reward
Thomson & Galloway (1970)	General Improvement	Reward
Turco & Elliott (1990)	Not Specified	Work
Velazquez (2014)	Other	Work; Criteria; Reward
Weekley (1980)	Not Specified	None
Wilder (2011)	Alternate Calculation; Other	Reward
Winn (2005)	Dichotomy; Other	Criteria
Wodarski et al. (1973)	Alternate Calculation; Dichotomy	None
Yarbrough et al. (1977)	Alternate Calculation	Criteria

Note. Alternative Calculation = a broad category which indicates that criterion was determined based on a calculation outside of percentile shaping (e.g., criterion based on previous mean or median performance)

Figure 1

Study Identification and Screening



Appendix A

Procedural and Coding Manual

Contingencies in Classrooms

Procedural and Coding Manual

Updated February 2, 2023

Investigators:
Taylor Hitchings, Syracuse University
Dr. Bridget O. Hier, Syracuse University
Dr. Tanya L. Eckert, Syracuse University

Coding Team:
Dr. Tanya Eckert, Syracuse University
Taylor Hitchings, Syracuse University
Samantha Maguire, Syracuse University
Kaytlin Nelson, Syracuse University

STUDY SEARCH AND STORAGE

Search Terms

- Target-Related
    - *Academic Performance*
- Contingency-Related
    - *Contingen**
    - *Reward*
    - *Reinforcement*
- Assessment-Related
    - *Can't Do Won't Do*
    - *Performance Deficit*

Search Procedure

1. Go to PsycInfo
2. Above the search field, select "choose databases"
3. Select both ERIC and PsycInfo
4. Search all target-related terms in combination with all contingency- and assessment-related terms) with the following searches:
    - Academic Performance AND Contingen*
    - Academic Performance AND Reward
    - Academic Performance AND Reinforcement
    - Academic Performance AND Can't Do Won't Do
    - Academic Performance AND Performance Deficit

Storage Procedure

- After each search is completed, download search results on the PsycInfo and ERIC databases in a .ris format
- In the Covidence software, import each .ris file
- All gathered search results are now stored in Covidence and ready for screening

SCREENING, STAGE 1: TITLE AND ABSTRACT

SCREENING, STAGE 1 PROCEDURES

1. Go to www.covidence.org and log in
2. Select "Contingent Reinforcement for Academic Performance" under "Your Reviews" heading
3. Under "Title and abstract screening," click "Continue" to reach the page with all of the studies that need to be screened
4. Read the title and abstract for each study
5. Click "No" (i.e., exclude) or "Yes" (i.e., include) to the right of each study based on screening procedures listed below.

Screening, Stage 1: Title and Abstract		
Variable	**Coding**	**Criteria**
Inclusion/Exclusion	"No" (Exclude) = Study can be excluded based on title and abstract alone "Yes" (Include) = Study can NOT be excluded based on title and abstract alone	**EXCLUSION CRITERIA** -The study was not written in English (Note: abstracts are sometimes written in English even when the study itself is not, so if you are unsure, do NOT exclude the study until we get a full PDF in Screening, Stage 2) -The study included participants outside the K-12 grade range -The study is an informational text, review, meta-analysis, opinion paper, book chapter, or curriculum guide -The study used non-experimental methods and/or only solely includes qualitative data -The study was not conducted in a school setting **INCLUSION CRITERIA** On the Covidence program page, click "Yes" (i.e., include the study) ONLY if both of the following criteria are met:

		A: The title and/or abstract references academics in some way. This includes any of the following terms: • Academics • Academic Achievement • Academic Skill • Academic Performance • Reading • Writing • Spelling • Math This does not include references to general measures of overall skill, including GPA or non-specific standardized tests (e.g., ACT, SAT) **AND** B: The title and/or abstract references academic interventions and uses any of the following terms: • Contingency • Contingent • Reinforcer • Reinforcement • Token Economy • Reward • Incentive • Behavioral Modification • Behavioral Intervention

SCREENING, STAGE 2: FULL TEXT

SCREENING, STAGE 2 PROCEDURES

1. Go to www.covidence.org and log in
2. Select "Contingent Reinforcement for Academic Performance" under "Your Reviews" heading
3. Under "Full text review"," click "Continue" to reach the page with all of the studies that need to be screened
4. Click the blue link to each study's full text, which can be found directly below the "Full text" and "Abstract" dropdown menus
5. Screen according to "Screening, Stage 2: Full Text" procedures below.

Screening, Stage 2: Full Text - Complete		
Variable	**Coding and Criteria**	**Notes**
Full Text	**0: Exclude:** Full text cannot be accessed (PDF or hyperlink) **1: Include:** Full study is accessible (PDF or hyperlink)	**Do we have access to the full text through PDF or online?**
Language	**0: Exclude**: Study is not written in English **1: Include**: Study is written in English	**What language was the study written in?**
Study Type	**0: Exclude**: -Book Review -Book -Opinion Article -Curriculum Guide -Qualitative -Meta-Analysis -Literature Review -Survey -Observational -Correlational	**Did the paper use experimental or quasi-experimental methods?**

	1: Include: -Experimental or quasi-experimental design -Variable (typically intervention) is directly manipulated -Replication Studies -Secondary Data Analysis	
Participants	**0: Exclude**: -Participants are outside the ages of K-12 (or ages 4-19*) -Preschool -College/University/Post-Secondary -Teachers, schools **1: Include**: - Participants are students in grades K -12 - Participants are students in the age range of 4-19* - Classrooms (of participants with appropriate demographics)	**What were participants' grade/age demographics?** *Exclusion criteria should first be applied based on participants' grade levels. If grade level is not reported, then exclusion is based on participant ages.
Setting	**0: Exclude**: -Online -Summer Camps -After School Programs -Military-Based Schools/Programs	**What TYPE of setting was the study conducted in?** **Exclude** if… ● Any part of the intervention or reinforcement occurs outside included settings

	-Vocational Training Programs -Homes (including homeschool) -Residential, inpatient, or outpatient settings/clinics -Juvenile Corrections Locations -Special Education, segregated, or schools set up for students with behavioral, emotional, or academic special needs (NOT physical needs, e.g., a school for the deaf) **1: Include**: ***UNLESS the following school types conflict with exclusion criteria above*** -Public School -Private School (unless it fits one of the above cases) -Parochial/Religious School -Independent School (unless it fits one of the above cases) -Charter School -Summer School (in typical school setting) -Laboratory Schools, Experimental Educational Environments (EEE)	FAQs: 1.) What is a special education or segregated school in this context? These refer to schools that only serve students with behavioral or academic special needs. These are different than specific classrooms in traditional school setting that serve students with special needs, which should be included. 2.) What is a laboratory school or experimental educational environment (EEE)? These are school settings that are affiliated with a university or lab, often for the partial purpose of research and data collection 3.) What if the study doesn't specify what type of school served as the study's setting? Studies should only be excluded for the setting variable if it is explicitly stated that the setting falls under one of the identified exclusion settings.
Intervention	**0: Exclude**: -Non-Contingent Reinforcement **OR**	**Was this an academic intervention in which reinforcement was contingent upon direct academic performance? During at least one treatment phase, was the reinforcement the ONLY treatment variable being manipulated?**

	Contingent reinforcement is only provided for: -Engagement -Time on-task -Time in seat -Disruptive behavior -Compliance -Hand-Raising -Attendance **1: Include**: Direct academic behaviors -Accuracy -Task Grade -Completion -Problems attempted -Fluency -Error rates **AND** -Contingent reinforcement is treated as an isolated intervention during at least one treatment phase	**Reinforcement:** The consequence of a behavior increases the likelihood that that behavior will occur again. Often take the forms of rewards or incentives in this literature. **Contingent:** Reinforcement is only provided if a student performs to a certain level, criterion, or goal -**Exclude** if… ● Researcher does not manipulate the contingency-based reinforcement intervention in isolation in any phases of treatment ○ Are there confounding treatment variables/components/elements that were also implemented, so that it is unclear whether treatment effects are result of the reinforcement intervention versus another manipulated intervention component? ▪ Example: A treatment is manipulated that includes both peer tutoring and contingent reinforcement, but never CR alone ▪ Example: Group cohesion is manipulated, as well as presence or absence of reinforcement ▪ Example: A specific intervention is being examined as a treatment,

		with and without a reinforcement component.

		FAQ's:
		1.) What is considered a reinforcer?
		
		Reinforcement includes any of the following:
		● Access to toys, foods, or objects
		● Increased free time, access to special activities
		● Access to a resource room or special room with reinforcers inside
		● Escape from an undesired task
		    ○ Example: a student only needs to continue solving problems until they get three correct in a row
		● Praise or other positive attention from teachers, aides, peers, or other figures in classroom
		● Performance social recognition
		    ○ Example: Winning students and/or teams are mentioned in a newsletter, written on the board, etc.)
		● Special privileges
		● Tokens, coupons, points, that can be exchanged later for prizes in the classroom
		● Access to a more preferred (or less aversive) activity, task, or assignment
		
		2.) What if the contingency, target, or reinforcement is not explicitly stated?

		If the study's procedures are general or unclear, to the extent that the study could not be replicated and this variable cannot confidently be coded, the study should be excluded. 3.) What if a behavioral target is not clearly an academic task? These studies should be excluded. Some examples of these targets could include tasks with a made-up language, circling images, finding differences in pictures, or other measures of performance that aren't clearly part of a traditional school curriculum for academic skills. 4.) What if a study only provides contingency-based reinforcement for a non-direct academic behavior, but direct academic outcomes are still measured and reported? Exclude these. Studies must directly target academic performance as part of the intervention to be included. 5.) What if the intervention we're interested in was already part of typical classroom procedures, prior to the start of the study? Exclude these. If our specific intervention of interest (i.e., contingency-based reinforcement for direct academic performance) cannot be

		experimentally manipulated, it doesn't fit inclusion criteria. 6.) What if contingency-based reinforcement is only implemented as part of a larger treatment package? Unless the treatment package only includes different elements related to contingency-based reinforcement (e.g., mystery motivators), these studies should be excluded. For example, if one phase of an intervention introduces both corrective feedback and reinforcement (but never reinforcement on its own), this study should be excluded. Common packaged intervention components that should be excluded include: self- and peer-monitoring and management; peer tutoring; response-cost procedures; repeated practice or instruction after incorrect responses; task choice 7.) What if reinforcement is provided for both direct and non-direct academic performance at once? Studies must isolate direct academic performance in isolation during at least one treatment phase. 8.) How are **brief experimental analyses (BEA)** or **functional assessments** coded?

		If a study only includes assessment for the purposes of identifying potential interventions for implementation, the study should be excluded. However, if a contingency-based reinforcement intervention is implemented during extended analyses, the study should be included. Example: A study uses BEA methods to identify an intervention, and one of the assessment conditions is contingency-based reinforcement. The study should be **excluded** if contingency-based reinforcement was not identified as an indicated treatment (or selected by students), so it was not implemented as an intervention during extended analysis. It would be **included** if contingency-based reinforcement was implemented as an isolated treatment phase during extended analyses.
Academic Subject	**0: Exclude:** Studies that only include data outside of targeted subjects -Science -Social Studies -PE -Finance -Computer/Tech -Geography -ESL (English as a second language)	**What academic skills were targeted and analyzed?** FAQ's 1.) What if the intervention is implemented in classes that feature both included and excluded subjects? This depends on how data were analyzed. If an intervention is implemented in both a non-included subject (e.g., science) and an included one (e.g., arithmetic), the study must provide data about arithmetic

	1: Include: -Math -Reading -Writing (handwriting, words written, accuracy, etc.) *Note: This does NOT include the motor skill of typing -Spelling	in isolation. If data are only analyzed as overall academic performance (which includes both subjects), the study should be excluded.
Dependent Variables	**0: Exclude**: -GPA -SAT, ACT, MAT, etc. -District or state-wide exams - WIAT, WISC, Peabody, WRAT, etc. -NAEP **1: Include**: -Measures created by teachers or experimenters -Homework accuracy/completion -Meeting goals related to improvement or achievement -Curriculum-Based Measurement (CBM) -AIMSweb -Workbooks	**What were the dependent variables of the study? What outcome data were analyzed in the results?** 1.) What if intervention results are reported with both included and excluded measures? As long as at least one outcome is reported that is consistent with the included dependent variables in the manual, the study should be included.
Decision	**0: Exclude**: Any previous variables were scored a “0” **1: Include**: All previous variables were scored a “1”	**Does the study meet all inclusion criteria?**
Studies marked “Include” will move forward to the next stage of the systematic review process, the ancestral review.		

ANCESTRAL REVIEW

ANCESTRAL REVIEW PROCEDURES

1. For all studies that meet inclusion criteria at the end of the second stage of screening, the first author will review all citations included in the References section of each study to determine if any relevant articles were missed
2. The citation's title will be examined to determine if it meets the inclusion criteria identified in the first stage of screening:

EXCLUSION CRITERIA

-The study was not written in English

-The study included participants outside the K-12 grade range

-The study is an informational text, review, meta-analysis, opinion paper, book chapter, or curriculum guide

-The study used non-experimental methods and/or only solely includes qualitative data

-The study was not conducted in a school setting

INCLUSION CRITERIA

A: The title and/or abstract references academics in some way. This includes any of the following terms:

- Academics
- Academic Achievement
- Academic Skill
- Academic Performance
- Reading
- Writing
- Spelling
- Math

This does not include references to general measures of overall skill, including GPA or non-specific standardized tests (e.g., ACT, SAT)

AND

B: The title and/or abstract references academic interventions and uses any of the following terms:

- Contingency
- Contingent
- Reinforcer
- Reinforcement
- Token Economy
- Reward
- Incentive
- Behavioral Modification
- Behavioral Intervention

3. Identify citations that fit the above criteria, then compare these to the master list of studies exported into Excel from Covidence
4. If there are any citations not in the master list, note them in the "Ancestral Review Identification" spreadsheet

Once the references of all included studies are examined, the ancestral review will have a second step using Google Scholar:

5. Go to www.scholar.google.com
6. Enter the included study's title in quotes
7. Find the included study's citation on the Google Scholar page
8. Underneath the study's citation, click on the blue hyperlink that states "Cited by ____" and a particular number
9. Review the titles of all results using the first-stage screening process
10. Identify relevant citations
11. Consult the master spreadsheet to determine if relevant citations have already been screened
12. Note non-duplicated results in the "Ancestral Review Identification" spreadsheet

From here, all studies documented in the "Ancestral Review Identification" spreadsheet will undergo the "Screening, Stage 2: Full Text" procedures outlined above.

DATA EXTRACTION

Full Article Review Preparation and Extraction Procedures:

- Once all screening has been completed, all studies that met inclusion criteria will be compiled in Covidence for data extraction.
- Coders will extract relevant data from each included article using the data extraction template created in Covidence:
    1.) Login to Covidence and select "Contingent Reinforcement for Academic Performance" under "My Reviews"
    2.) Open the dropdown menu under "Extraction"
    3.) Click "Continue."
    - For ISA coders only: Click the dropdown menu next to "Filter by tags" and select the tag "Extraction ISA."
    4.) Click "Begin extraction" on the first article at the top of the screen
    5.) Fill in relevant information in the data extraction template on the far right of the screen, using the manual below for guidance
    6.) Once all information is entered, click the blue "Complete" button in the top right corner of the screen.

Coding Manual for Data Extraction

General Notes

1. If a text includes multiple studies or experiments, data for all experiments that met inclusion criteria should be combined and reported as one.
    a. Example: If experiments A, B, and C all meet inclusion criteria, the number of participants recorded will be equal to the total number of participants studied across all experiments—unless the text states that the same students were used in multiple experiments.
2. For studies that include multiple experiments—but only one or some experiments meets inclusion criteria—only data from the included experiments should be entered. In these cases, the experiments to be included will be noted in the "Notes" section beneath the study's entry on Covidence.
    a. Example: Experiments A and C met inclusion criteria, but not B. The total number of participants should only include the number of participants from A and C.
3. On the data extraction form, items should be selected if they occur at any point in any of the included experiments.
    a. Example: One experiment used an interdependent group contingency, while another used an individual group contingency. Both "Group, Interdependent" and "Individual" should be selected for the Contingency Format variable.

b. Example: One experiment reported that they used percentile shaping to select the criterion, but another experiment did not discuss how the criterion was selected. Both "Percentile Shaping" and "Value or Method Not Specified" should be selected.

Variable	Coding	Notes
Basic Characteristics		
Title	Enter the study's title in the text box	
Authorship	Copy and paste the full list of authors from the Covidence study entry	
Year Published	Enter year published in text box	This is the year that is listed at the top of each study entry next to the study ID number and author's last name.
School Type	Select one of the following options: - Public - Laboratory or EEE - Alternative/Private - Not Specified	Public: Study specifically states that school(s) are public or part of a school district. Laboratory of EEE: These are school settings that are affiliated with a university or lab, often for the partial purpose of research and data collection Alternative: Schools that do not fit the "Public" or "Laboratory or EEE" categories. Includes parochial schools, charter schools, independent schools, private schools, etc. Not Specified: If the school type is not clearly identified as any of the above options

Classroom Setting	Select one or more of the following options: - General - Special Education - Inclusion - Study not conducted in S's typical classroom - Not Specified	Where did the study procedures occur? This does NOT necessarily refer to the students' typical classroom or the type of classroom they are usually enrolled in. Example: A student is enrolled in a general education classroom, but the study procedures occurred in an empty classroom nearby. This should be coded as "Study not conducted in S's typical classroom" only. Special education includes: Resource room, self-contained classroom.
Number of Participants	Enter total number of participants in the study in text box	This should be the final value of individuals who participated—not the starting value that may include students who did not consent to participate or students whose data were removed later on.
Research Methodology	Select one of the following options: -Group Design -Single-Case Design	The study's design is often stated at the end of the method section. Otherwise, examine the procedures used to determine the design. Group design: Participants are assigned to alternative conditions (e.g., one treatment vs. one control) and data are analyzed at a group-wide level (i.e., analysis is looking at group statistics like averages, rather than data for individual participants in isolation). May include within-subjects or between-subjects methodologies. Single-case design (SCED): Types of SCED may include "multiple baseline," "alternating treatments," "reversal," or "changing criterion." Data are analyzed at the individual level (i.e., changes in a single subject)

		This could refer to one individual student, one individual classroom, one individual school, etc., depending on the level of analysis. In short, changes in individuals' behaviors are examined, rather than using info related to group differences or score averages across different groups.
Participant Demographics		
Number Females	Enter the study's number of female participants in the text box, using full numbers	Type "NR" if not reported If clearly gendered pseudonyms are used for students, participants' genders can be inferred.
Age (Youngest)	Enter age of youngest participant (in years) in text box, using whole numbers	Type "NR" if not reported
Age (Oldest)	Enter age of oldest participant (in years) in text box, using full numbers	Type "NR" if not reported.
Grade	Select one or more of the following: -Kindergarten - 1st - 2nd - 3rd - 4th - 5th - 6th - 7th - 8th - 9th - 10th - 11th - 12th - International	The "general" variables should be selected if no specific grade levels are reported. The "general" options may be selected in cases where participants' ages or other information lead to reasonable assumptions about participants' general grade levels.

	- General Elementary - General Middle - General High School - Not Reported	
Number Hispanic or Latinx	Enter number of participants who were Hispanic or Latinx in text box	Should always be a value, which may simply be zero. If no racial or ethnic data are reported, all values would be "0" except for "Number Unspecified." In this case, "Number Unspecified" would be equal to the total number of participants if no data were provided about ethnicity/race. In short, all students should be accounted for in one of these categories. Values should not be duplicated, so the total across these boxes should add up to the number of participants exactly. If a percent is reported rather than a number, calculate the number by hand. If this calculation results in an answer like "4.8," the answer should be rounded to the nearest whole number. Values less than .5 should be
Number White	Enter number of participants who were White of Caucasian in text box	
Number Black and/or African American	Enter number of participants who were Black and/or African American in text box	
Number American Indian or Alaska Native	Enter number of participants who were American Indian or Alaska Native in text box	
Number Asian	Enter number of participants who were Asian (Far East, Southeast Asia, or the Indian subcontinent) in text box	

Number Native Hawaiian or Other Pacific Islander	Enter number of participants who were Native Hawaiian or OPI (Hawaii, Guam, Samoa, or other Pacific Islands) in text box	rounded down, and values .5 and up should be rounded up. **Whenever possible, the "Other Race/Ethnicity" option should be used to provide specific information that is more consistent with current APA guidelines related to bias-free language. For example, if "Chinese-American" is specified, that should be noted here, rather than in the broad "Asian" category.**
Number Other Race/Ethnicity	Enter the specific race or ethnicity in question, followed by the number of participants with that race or ethnicity. If more than one race or ethnicity needs to be listed, separate entries with a comma. Example: Irish: 8, Spanish: 10	
Number Multiracial	Enter number of participants of two or more races or ethnicities in text box	
Number Unspecified Race/Ethnicity	Enter the number of participants who had no racial or ethnic data reported.	
Number EBD	Enter the number of participants receiving special education services for EBDs	A value should always be entered for these variables (rather than "NR"). If no students are identified as having a disability, the value for all categories should be zero. Because students may have multiple disabilities, the total value of disabilities may be greater than the number of participants. If students are receiving special education services for an unspecified disability, those students would fall under the "Number Other Disability" variable. If a student is enrolled in an inclusion classroom, it must be
Number SLD	Enter the number of participants receiving special education services for SLDs	
Number Autism	Enter the number of participants receiving special education services for autism	
Number Other Health Impairment	Enter the number of participants receiving special education	

	services for "Other Health Impairment"	specified that the target student is personally receiving special education services in the classroom. These classifications are based on IDEA, not the DSM-5. Further information on IDEA classifications: https://sites.ed.gov/idea/regs/b/a/300.8 *If a listed disability does not fall under IDEA, this should not be included in this count.
Number Other Disability*	Enter the number of participants that fall under any of the following categories: 1. Receiving special education services for an unspecified disability. This does NOT include remedial services or programs unless they are explicitly labeled as "special education." 2. Disability falls under a disability category besides EBD, SLD, autism, or Other Health Impairment. 3. A student is described as having a disability (e.g., ADHD), but it is not clear whether this label is based on IDEA or DSM standards. If a diagnosis is listed but the student is not receiving special education services, do not count this diagnosis.	
Assessment		
Can't Do/Won't Do (CDWD)	Select one of the following options:	**To inform intervention selection before the full treatment,** did the researchers/educators offer students a reward (with no other treatment components) if they improved their initial baseline score?

	Yes = Can't Do/Won't Do assessments were conducted prior to implementing intervention No = Can't Do/Won't Do assessments were not conducted prior to implementing intervention	By themselves, functional analyses should be coded as "No" for this variable. The study does not have to name the assessment as "Can't Do/Won't Do," but (A) the authors must explicitly state that they are trying to identify whether a student has a performance deficit and/or performs better with a reward only; and (B) a reward must be offered as part of the initial assessment to determine what intervention will be implemented. NOTE: If a reward is only offered as part of an intervention that has already been selected, code this as a zero. To qualify as an assessment, rather than an intervention component, rewards must have been offered before intervention selection, with assessment results being used to inform subsequent intervention selection.
Preference	Select one of the following options: Yes = Formal preference assessments were conducted prior to intervention No = Formal preference assessments were not conducted prior to intervention	A preference assessment is defined as: "A process of identifying reinforcers for an individual that involves presenting potential reinforcers and measuring whether the individual approaches, manipulates, or consumes the item. Preference assessments can be conducted in at least three different ways; single stimulus assessment, paired stimulus assessment, and multiple stimulus assessment" (Miltenberger, 2016, p. 534). This does not include general preference inventories or student suggestions for preferred items. Students' preferences must be systematically explored to produce a ranked list of the student's preferred rewards. NOTE: If changes in the participant's behaviors are measured when rewards are being identified, this is

		NOT a preference assessment and should be coded as a reinforcer assessment (see below)
Reinforcer	Select one of the following options: Yes = Reinforcer assessments were conducted prior to intervention No = Reinforcer assessments were not conducted prior to intervention	A reinforcer assessment is defined as: "A process in which an item from a preference assessment is delivered contingent on a behavior to see if the behavior increases. If the behavior increases, the item functions as a reinforcer" (Miltenberger, 2016, p. 535)
Intervention		
Contingency Format	Select one or more of the following: - Individual - Group, Independent - Group, Dependent - Group, Interdependent - Ambiguous, Individual/Independent	Individual: Reinforcement contingencies are structured and provided for one specific student Group, Independent: Behaviors and reinforcement are occurring in group (i.e. 2+ students) settings, but an individual's own performance is the sole determinant of whether they receive reinforcement Group, Dependent: Behaviors and reinforcement are occurring in group settings, and the individual's access to reinforcement is entirely dependent upon the performance of their classmates (e.g., the student only receives access to a reinforcer if their tablemate scores 80% or higher) Group, Interdependent: A student's access to reinforcement is determined based on their performance, as well as that of their classmates (e.g., reinforcement based on class averages, all students

		must score 90% or above for the class to get a pizza party) Ambiguous, Individual/Independent: A contingency is formatted as either independent or individual, but it is impossible to differentiate due to unclear procedures, setting, or group structure.
Criterion Selection	Select one of the following: - Percentile Shaping - Alternate Calculation - General Improvement - Dichotomy - Other - Value or Method Not Specified	**How was the specific criterion value determined? Why/How did the researchers, teachers, and/or students select the criterion of interest?** **NOTE: This refers to the criterion for immediate reinforcement and NOT the criterion needed to access backup reinforcers.** ● For example, if a student receives a sticker for each correct answer, "each correct answer" is the criterion of interest that should be scored. ● Even if ten stickers are then needed to "purchase" a homework pass or other backup reinforcer, the "Criterion Selection" variable focuses on the immediate relationship between the target behavior criterion and subsequent reinforcement. **Percentile Shaping:** The criterion is calculated using a specific equation that integrates the student's previous performances on the task. The term "percentile shaping" must be used. **Alternate Calculation:** The criterion was determined based on a specific calculation outside of percentile shaping.

		Example: The mean of all students' baseline performances was used as the criterion value for the intervention. Example: The student's lowest score at baseline was used as the initial criterion during the intervention. Example: The new criterion was the student's previous quiz score plus 10% (i.e., criterion = 1.1x the student's past score). **General Improvement:** The student's performance must simply be equal to or higher than their previous performance, regardless of amount of improvement Example: The student must beat their previous score(s) to access reinforcement. **Dichotomy:** Instead of an overarching goal that the student must achieve, reinforcement is provided based on a dichotomous variable on an individual question, item, or task. That is, reinforcement is not provided based on cumulative performance. This can also be thought of as "Per Item" or "Per Task" criterion. This is most often seen in cases where a student receives a reward based on correct/incorrect responses to single items, or in cases where students receive a reward based on complete/incomplete items or assignments. Rarely, this occurs when the target

		behavior is only discussed as pass/fail, with no information about what percentage is needed to pass. Please keep in mind: ● Any criteria that discuss percentages, averages, total numbers correct, etc., should NOT be classified as a dichotomy. Example: A student is allowed to go play as soon as they complete an assignment in class. Example: A child receives a token or point for each correct response to a question. (Even if a certain number of tokens or points must be used to "purchase" a back-up reinforcer.) *Non-Example: A child receives a token after they answer seven problems correctly*. This would be coded as "Value or Method Not Specified" unless an explanation is provided for the reason seven problems need to be correct for a token. Example: A child receives a homework pass if they complete the previous day's homework (with no requirements for accuracy or anything beyond individual task completion). *Non-Example: A child receives a homework pass if they complete their previous three homework assignments*. This would be coded as "Value or Method Not Specified" unless an explanation is

		provided for the reason that three completed homework assignments is the criterion. **Other:** The study states the criterion selection method, but the method does not fit the established categories. **Value or Method Not Specified:** The study reports the criterion for reinforcement but does not specify the method or process used to determine specific criterion values. The following are potential situations that fall under this category: Example: Students must have five or fewer errors in oral reading, but it is not clear who selected this criterion or how it was determined. Example: An experimenter selects criterion based on students' baseline performances without specifying how these data were used to select the criterion. Example: A study states the criterion with no reference to how it was selected. Example: Students were told that they would receive a box of crayons if they did well on their quizzes. (With no other information provided.)
Unknown or Randomized Components	Select one or more of the following if present: - Randomized/Unknown Work	Randomized/Unknown Work: Participants do not know which students' work will be selected for potential reinforcement OR they don't know what specific task will be used to determine reinforcement.

	- Randomized/Unknown Criteria - Randomized/Unknown Reward - None	Example: All students turn in homework, but three students' papers will be randomly selected and scored to see if the students receive candy Example: A student turns in a packet of five worksheets, but the teacher will only score one to determine whether the student gets extra free time Example: A student must turn in homework for math and reading, and the teacher randomly selects either reading or math as the basis for reinforcement. Randomized/Unknown Criteria: Participants do not know how well they have to perform for reinforcement. Includes variable reinforcement schedules. Example: A box contains slips of paper with 75%, 90%, and 100% on them. The teacher randomly draws one of these slips after work is collected, and the slip drawn specifies how many students needed to pass the quiz to receive a pizza party Example: A student sometimes receives reinforcement for getting a single problem right and sometimes receives reinforcement after every 3 responses.

		Randomized/Unknown Reward: Participants do not know how they will be reinforced (or what reward they will receive) Example: A box contains slips of paper for extra recess, candy, free time, or a homework pass. A slip is then drawn to determine how academic performance will be reinforced.
Academic Subject	Select one or more of the following: - Math - Reading - Writing - Spelling - English Language Arts	English or Language Arts refers to cases in which specific skills (e.g., reading, writing, spelling) aren't named as the target, but rather broader subjects like English Language Arts. This does NOT refer to learning English as a new language. If it is ambiguous, lean on the side of assuming the study is referring to English in a language arts context.

CHANGE TABLE

Pilot Screening	
Questions/Concerns	**Changes Made for Final Manual**
The utility of coding the year published during initial screening stage	"Year Published" variable was moved to full article-coding under "Basic Characteristics"
The original manual reflected processes based on Zotero PDF storage and Excel for screening. Zotero was excluded, and Covidence was used for both PDF storage and screening.	Background information and screening procedures were updated to reflect the use of Covidence and Excel during screening
Proquest Dissertations & Theses Global (PQDT) database retrieved a very large amount of irrelevant and duplicate studies	PQDT was removed from search procedures, while Google Scholar was added to search procedures for the ancestral review

Significant overlap between the "original study" and "study type" variables, resulting in difficulty differentiating between the two	"Original study" and "study type" were condensed into a single variable: "Study Type"
Difficulty separating "Intervention" and "Target" variables for individual coding of each variable	"Intervention" and "target" were condensed into a single variable: "Intervention"
Lack of clarity for screening somewhat unique cases, especially those related to the "setting" and "intervention" variables	Specific examples of included and excluded information were added to the manual. Example studies for inclusion or exclusion were added to the notes section of each variable. A "frequently asked questions (FAQ's) section was also added below a variable's notes to specify screening procedures for common issues that occur when coding for a respective variable.
The dependent variables of some studies were inconsistent with or not sensitive to incremental changes in the targeted academic behavior	A variable called "dependent variables" was added before a final decision was made for a study's inclusion. Specific outcomes were then identified in the manual as reasons for inclusion or exclusion.
Many studies that appeared in electronic searches only included contingency-based reinforcement as part of a packaged intervention, making it impossible to isolate the potential impact of contingency-based reinforcement	Clarification of the intervention variable was added to the manual, in which contingency-based reinforcement must have been implemented in isolation during at least one treatment phase for a study to be included.
Studies using brief experimental analysis (BEA) or functional analyses were found in searches but were not addressed in the manual	An FAQ section was added beneath the "intervention" variable to address the situation. This stipulated that contingency-based reinforcement needed to be implemented during extended analyses for inclusion (rather than during brief experimental analysis procedures alone)
Ancestral review procedures were too vague for potential replication, and the scope of the initial ancestral review appeared to be missing relevant studies.	Ancestral review procedures were outlined with more procedural specificity and a brief ancestral review coding table was added.

	A second phase of ancestral review was also added, in which all included studies were entered into Google Scholar and the “Cited By” function was used to review all titles and abstracts of studies that had cited the included study in question.
Pilot Data Extraction	
Questions/Concerns	**Changes Made for Final Manual**
The manual wording and coding options did not reflect the Covidence software’s template.	The general wording around procedures was updated. Additionally, the option to select more than one option for relevant variables was introduced.
Research Methodology Variable: Based on updated screening criteria, all non-experimental studies were already excluded. Many group design studies were also quasi-experimental, rather than experimental with random assignment to conditions.	Coding options are limited to group or single-case design. Studies that used experimental or quasi-experimental design were both coded as “Group Design.”
The variables beneath the “Basic Characteristics” sub-header did not address each study’s setting in any way, though this information is relevant to additional research as well as the applied uses of this review.	Two variables were introduced: School Type and Classroom Setting.
Studies often reported the range of participants’ ages, but all participants’ ages were very rarely reported.	The “Age (Mean)” variable was removed.
Data related to participants’ racial/ethnic characteristics was not fully inclusive, did not reflect the most recent APA guidelines relate to bias-free language, and did not account for cases in which some (but not all) participants’ racial/ethnic data were not reported.	Wording of pre-existing variables was updated to be more inclusive and reflective of the various ways that participants’ racial/ethnic data were reported. Two additional variables were introduced: “Number Other Race/Ethnicity” and “Number Unspecified Race/Ethnicity.” A note was added to the manual instructions that the most specific racial/ethnic data possible should be reported using the “Number Other Race/Ethnicity” variable. For example, “Chinese-American” and

	"Japanese-American" would be reported using the "Number Other Race/Ethnicity," rather than the broad "Asian" category.
Participants' disability or special education statuses were not coded, though this information is relevant to future research and applied work.	Five new variables were created: "Number EBD," "Number SLD," "Number Autism," "Number Other Health Impairment," and "Number Other Disability."
The coding for the "Preference" assessment variable was unclear regarding common methods of discerning participants' reward preferences.	It was clarified that students' preferences needed to be systematically explored in a way that results in a ranked list of preferences. It was also specifically noted that general preference inventories or student suggestions were not preference assessments.
In rare cases, studies did not adequately report setting information, making it difficult to discern whether the intervention occurred in a group format.	An "ambiguous" option was added for contingency format.
There criterion selection coding options were overly limited and not applicable to the criterion selection procedures in many included studies.	Additional criterion identification coding options were added, including "dichotomy," "alternate calculation," and "Value or Method Not Specified." "Stable percent improvement" was removed as a possible coding option.
Many studies utilized additional alternative randomized/unknown components, beyond randomized/unknown criteria alone.	Additional options to code for randomized/unknown rewards and work were added to this variable.
Some studies reported the use of contingency-based reinforcement interventions to target English or Language Arts skills, which were comprised of reading, writing, and spelling in combination.	English Language Arts was added as a possible subject in coding.

Appendix B
Screening Training Manual

Contingencies in Classrooms

▪

Screening Interscorer Agreement Training Manual

Updated August 5, 2022

Investigators:
Taylor Hitchings, Syracuse University
Dr. Bridget O. Hier, Syracuse University
Dr. Tanya L. Eckert, Syracuse University

ISA Coding Team:
Samantha Maguire, Syracuse University
Kaytlin Nelson, Syracuse University

About the Project

This project is a systematic review of the literature that has examined the effects of providing reinforcement based on students' direct academic performances. The only subjects being examined are reading, writing, spelling, and math.

Title: Contingencies in Classrooms: Contingencies in Classrooms: A Systematic Review of Contingent Reinforcement for Academic Performance

Purpose: To identify the study characteristics, participant demographics, assessment procedures, and intervention components that have been examined in the relevant literature in this area.

Research Questions:

1.) What are the basic publication and methodological characteristics of the studies that have explored contingency-based reinforcement interventions for academic performance?
2.) What are the demographic characteristics of participants who have been included in studies examining contingency-based reinforcement interventions?
3.) What percentage of studies included Can't Do/Won't Do assessments to inform intervention selection?
4.) What percentage of studies used preference and reinforcer assessments to identify effective rewards?
5.) What percentage of studies used percentile shaping, stable percentage improvements, or randomized and unknown contingency components to determine and communicate reinforcement criteria? Among those that used stable percentage improvements, how much did students' performances need to improve across sessions according to the reinforcement criterion?
6.) What percentage of studies used individual contingencies versus group contingencies? Of the interventions that were administered in a group-wide format, what percentage were implemented using independent, dependent, and interdependent formats?
7.) How frequently has each specific academic subject (i.e., math, reading, writing, or spelling) been directly targeted with contingency-based reinforcement interventions, and what academic skills have been targeted within these subject areas?

Dissemination: Will be turned into a manuscript and submitted for publication to *JPBI*, potential submission to APBS annual convention

ISA General Information and Tips

General Info

Overall, there will be approximately 146 studies that need to be scored as part of the initial screening and ancestral review. On average, screening takes about 80 seconds per study, though this varies based on the length and type of study being screened.

Software Used: Google Drive and Covidence (through any web browser)
Ideal Deadline: Screening completed by September 16th, 2022

General Expectations: Because most screening will be done during summer when people may be on vacation or wish to take time off, specific assignments will not be given on a weekly basis. However, with an ideal deadline of 9/16/22 (assuming actual coding starts around 8/8), an average of 24 studies should be screened per week. This means that an average of 12 studies per week should be completed by **each coder**.

Communication and Meetings: The primary means of communication throughout screening will be email. I am always available at thitchin@syr.edu, and I aim to respond to all questions or other emails within (at most) 12-24 hours. If either of you would like to meet to discuss anything over other methods, I am happy to schedule Zoom meetings.

Each Monday during screening, I will review the coding that was completed during the prior week and email specific feedback to each coder. If more in-depth feedback is necessary, I will request to meet over Zoom with one or both coders.

Broad Coding Tips

- You do not need to read the entire study! Nearly all relevant information will be found in the methods section, typically under the "sample," "procedures/treatments," and "dependent variables" sub-headers. In rare cases, you may need to examine the results tables to examine how data were analyzed, primarily in relation to the "Subject" variable

- Although variables need to be coded in order for ISA purposes, it saves time to jump around in the spreadsheet itself as variables appear. I recommend coding all you can from the abstract, then jumping into the methods section for the remaining variables.
    - If the study is ultimately excluded, simply erase the subsequent variables that have already been coded

- If you are unsure about how to code a particular variable for a study (or it appears to be an odd case), the FAQ's section of the coding manual often includes specific examples of hypothetical studies and their coding, which may be relevant. Otherwise, you're always welcome to email me!

Software

Google Drive

All coding will be completed in a shared Google Drive folder entitled "Contingencies in Classrooms - Thesis." The Drive folder will be shared with your SU email address, unless you prefer otherwise.

Subfolders will be used to differentiate between proficiency checks and formal ISA screening. The proficiency spreadsheet files are called "Thesis Proficiency ISA_[YOURNAME]," and are separated by coder. There is a single spreadsheet ("Thesis Screening ISA") that is used for formal coding by all coders.

The shared Google Drive also holds resources relevant to the project, including the ISA training manual, example coding, and the coding manual.

Covidence

Link: https://www.covidence.org/

Covidence is a systematic review software that can be accessed on any web browser. It holds a list of all studies that need to be screened, as well as links to each study's full text. You will enter your coding decisions in Covidence during formal screening. Complete procedures can be found in the "Proficiency Procedures" and "Formal Screening Procedures" sections of this manual.

Troubleshooting

XML Error

This error sometimes occurs when you click the blue link to the full text PDF, but it does not open. Instead, in a new tab, you receive the message:

This XML file does not appear to have any style information associated with it. The document tree is shown below.

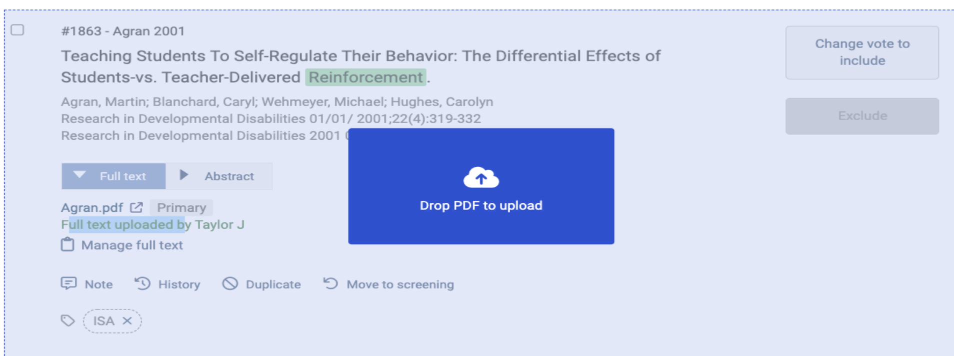

```
▼<Error>
  <Code>AccessDenied</Code>
  <Message>Request has expired</Message>
  <Expires>2022-06-04T20:24:12Z</Expires>
  <ServerTime>2022-06-05T17:16:46Z</ServerTime>
  <RequestId>34EE47D72D6YPWRM</RequestId>
  <HostId>0WIZzbzqj41l1eM8Npfc9i3QXuDbeLy8bHNznl4i031+0/eUuzK7/j08xKyctocXiPUzgC9+4Wo=</HostId>
</Error>
```


This error message occasionally occurs when many studies have been screened in a row. Once it happens, you will likely continue to get this error message for each subsequent full text you try to access. There are two options to solve this:

1.) Refresh the page link (I recommend this in cases where you repeatedly see the error message).
    a. NOTE: You will have to redo the “Filter” and “Relevancy” steps each time you refresh
2.) Select “Manage full text” underneath the link, and click the blue full text link from the pop-up box

“Drop PDF to upload”

This error can also occur when you click the blue link for a study’s full text. It happens when Covidence believes you were trying to drag the link, rather than click it. It looks like this:



Once you get this error, you will not be able to access that study’s full text. To fix this, refresh the page (and again do the filter and relevancy sorting procedures). You can also just skip it and continue screening the other studies—the next time someone logs in for screening, the full text will be available again

Accidentally clicking the incorrect decision or reason for exclusion

This is fine! Simply shoot me an email and let me know what study it was and what the correct coding should have been. I can easily fix any accidental coding.

Proficiency Procedures

To achieve proficiency, each coder must score at least 90% agreement on a series of ten practice articles. If 90% is not achieved on the first round, scorers will complete practice coding until 90% ISA is achieved on sample items.

When you screen, you will need to have both Google Drive and Covidence open. Through Covidence, you'll have access to studies' full texts. Actual coding will occur in the Google Drive spreadsheet.

Google Drive

To access proficiency spreadsheet

(1) Open "Contingencies in Classrooms – Thesis" shared folder on Google Drive
(2) Click "Screening" subfolder
(3) Click "Proficiency" subfolder
(4) Open the spreadsheet called "Thesis Proficiency ISA_[YOURNAME]"

Covidence

Link: https://www.covidence.org/

1.) After initial account set-up, click the three straight gray lines in the top right corner of the Covidence homepage and select "Sign in" from the dropdown menu
2.) Sign in
3.) Select "Contingent Reinforcement for Academic Performance" under "Your reviews"
4.) Select "Full text review" to open another dropdown menu
5.) Click "Continue" underneath the number of studies that can still be screened
6.) A list of studies will appear on the page. The next step is to filter these studies so that the only studies that appear are those that are used for the proficiency check

a. Click the gray button "Filter" → Click the text box where you can type under "ADD A FILTER" → Select "Pro" → Click BLUE button "Filter" → Click the gray "Filter" button to hide the dropdown menu
b. Select the gray "Relevancy" dropdown menu → Select "Author"

All remaining studies are those that need to be completed for proficiency. You'll notice that they all have the tag "Pro" marked in a dotted blue circle at the bottom of each entry.

Once you've accessed the full texts, please complete the proficiency coding in the "Thesis Proficiency ISA_[YOURNAME]" file.

(During formal screening, you'll enter your inclusion decisions into Covidence after spreadsheet coding, but we ignore this step during the proficiency stage.)

Feel free to email me or text with any questions or concerns. ISA proficiency will be reviewed, and feedback will be provided as needed.

Formal Screening Procedures

When you screen, you will need to have both Google Drive and Covidence open. Through Covidence, you'll have access to studies' full texts. Actual coding will occur in the Google Drive spreadsheet.

Google Drive

To access formal coding spreadsheet

(1) Open "Contingencies in Classrooms – Thesis" shared folder on Google Drive
(2) Click "Screening" subfolder
(3) Click "Formal Screening" subfolder
(4) Open the spreadsheet called "Thesis Screening ISA" → This is a shared document across all coders

Covidence

Link: https://www.covidence.org/

1.) After initial account set-up, click the three straight gray lines in the top right corner of the Covidence homepage and select "Sign in" from the dropdown menu
2.) Sign in

3.) Select “Contingent Reinforcement for Academic Performance” under “Your reviews”
4.) Select “Full text review” to open another dropdown menu
5.) Click “Continue” underneath the number of studies that can still be screened
6.) A list of studies will appear on the page. The next step is to filter these studies so that the only studies that appear are those that are used for the proficiency check
    - **a.** Click the gray button “Filter” → Click the text box where you can type under “ADD A FILTER” → Select “ISA” → Click BLUE button “Filter” → Click the gray “Filter” button to hide the dropdown menu
    - **b.** Select the gray “Relevancy” dropdown menu → Select “Author”

All remaining studies are those that need to be completed for formal screening. You’ll notice that they all have the tag “ISA” marked in a dotted blue circle at the bottom of each entry.

<u>Entering Your Decision into Covidence</u>

Once all spreadsheet coding has been completed and the “Decision” variable has been determined, return to the Covidence tab where all studies are listed.

- a. If the decision in the spreadsheet was “0,” click the blue button “Exclude.” Covidence will prompt you to click on the reason for exclusion. Select whichever variable was coded as a “0” in your spreadsheet
- b. If all variables were coded as a “1,” click the blue button “Include”

The study has been successfully screened! Its entry will disappear from the list of studies tagged “ISA.”

Background Resources

The “Contingencies in Classrooms - Thesis” Google Drive folder also includes PDFs of three seminal articles that provide more in-depth information about contingency-based reinforcement. These articles may be useful resources to gain a greater understanding of the theory behind contingency-based reinforcement interventions, critiques of the intervention, the intervention’s history in the school psychology field, and examples of real or hypothetical instances where these interventions have been implemented.

The included studies are:

Gresham, F. M., & Gresham, G. N. (1982). Interdependent, dependent, and independent group contingencies for controlling disruptive behavior. *The Journal of Special Education*, *16*(1), 101-110.

Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis*, *8*(3), 341.

Skinner, C. H., Williams, R. L., & Neddenriep, C. E. (2004). Using interdependent group-oriented reinforcement to enhance academic performance in general education classrooms. *School Psychology Review*, *33*(3), 384-397.

In terms of application to this specific project, I would especially urge you to review Skinner, Williams, and Neddenriep (2004). This article speaks in-depth about academic performance contingencies and the different forms that they take. (C. H. Skinner also wrote a similar chapter in NASP’s *Best Practices in School Psychology IV,* entitled “Best Practices in Contingency Management: Application of Individual and Group Contingencies in Educational Settings.”)

Appendix C
Data Extraction Training Manual

Contingencies in Classrooms

Data Extraction Training Manual

Updated September 28, 2022

Investigators:
Taylor Hitchings, Syracuse University
Dr. Bridget O. Hier, Syracuse University
Dr. Tanya L. Eckert, Syracuse University

ISA Coding Team:
Dr. Tanya Eckert, Syracuse University
Taylor Hitchings, Syracuse University

About the Project

This project is a systematic review of the literature that has examined the effects of providing reinforcement based on students' direct academic performances. The only subjects being examined are reading, writing, spelling, and math.

Title: Contingencies in Classrooms: Contingencies in Classrooms: A Systematic Review of Contingent Reinforcement for Academic Performance

Purpose: To identify the study characteristics, participant demographics, assessment procedures, and intervention components that have been examined in the relevant literature in this area.

Research Questions:

1.) What are the basic publication and methodological characteristics of the studies that have explored contingency-based reinforcement interventions for academic performance?
2.) What are the demographic characteristics of participants who have been included in studies examining contingency-based reinforcement interventions?
3.) What percentage of studies included Can't Do/Won't Do assessments to inform intervention selection?
4.) What percentage of studies used preference and reinforcer assessments to identify effective rewards?
5.) What percentage of studies used percentile shaping, stable percentage improvements, or randomized and unknown contingency components to determine and communicate reinforcement criteria? Among those that used stable percentage improvements, how much did students' performances need to improve across sessions according to the reinforcement criterion?
6.) What percentage of studies used individual contingencies versus group contingencies? Of the interventions that were administered in a group-wide format, what percentage were implemented using independent, dependent, and interdependent formats?
7.) How frequently has each specific academic subject (i.e., math, reading, writing, or spelling) been directly targeted with contingency-based reinforcement interventions, and what academic skills have been targeted within these subject areas?

Dissemination: Will be turned into a manuscript and submitted for publication to *JPBI*, potential submission to APBS annual convention

Task: 98 studies met inclusion criteria during screening and will undergo data extraction. Of these, 33 studies were randomly selected for double scoring to calculate ISA.

Software

Google Drive

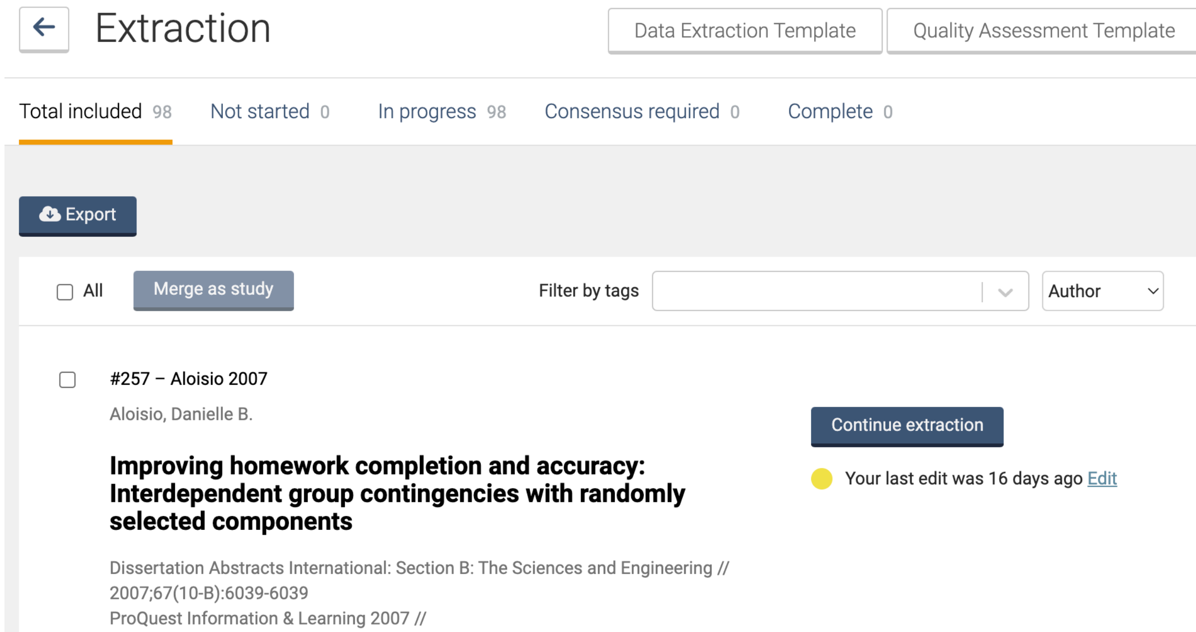
All coding will be completed in a shared Google Drive folder entitled "Contingencies in Classrooms - Thesis." The Drive folder will be shared with your SU email address, unless you prefer otherwise. Documents specific to data extraction can be accessed in the "Data Extraction" sub-folder.

The shared Google Drive also holds resources relevant to the project, including the data extraction training manual and the data extraction coding manual.

Covidence

Link: https://www.covidence.org/

Covidence is a systematic review software that can be accessed on any web browser. It holds a list of all studies that need to be screened, as well as links to each study's full text. You will enter your coding decisions in Covidence using the data extraction template provided. Complete procedures can be found in the "Proficiency Procedures" and "Data Extraction Procedures" sections of this manual.



Troubleshooting

Full text will not load

When a study is opened for data extraction, the full text will occasionally fail to open alongside the data extraction template. This seems to occur across different internet browsers. When this occurs, I recommend returning to the data extraction page where all studies are listed (picture above).

From here, select the white "View full text" button then click on the blue PDF link for your study of interest. The study's full text will then open in a separate tab and should load as normal.

Proficiency Procedures

To achieve proficiency, each coder must score at least 90% agreement on a series of five practice articles. If 90% is not achieved on the first round, scorers will complete practice coding until 90% ISA is achieved on sample items.

Practice and proficiency coding are done through the templates in Google Drive, while actual data extraction will occur in Covidence.

Google Drive

To access proficiency template:

(5) Open "Contingencies in Classrooms – Thesis" shared folder on Google Drive
(6) Click "Data Extraction" subfolder
(7) Click "Data Extraction Training and Proficiency" subfolder
(8) Open the PDF called "Data Extraction Template – Editable."
(9) Download the document and open in a PDF reading software
(10) Complete the template form using pre-added text boxes and marks
(11) Save and email to thitchin@syr.edu or upload to Google Drive
(12) Repeat the process with the next study

Covidence

Link: https://www.covidence.org/

7.) After initial account set-up, click the three straight gray lines in the top right corner of the Covidence homepage and select "Sign in" from the dropdown menu
8.) Sign in
9.) Select "Contingent Reinforcement for Academic Performance" under "Your reviews"
10.) Select "Extraction" to open another dropdown menu
11.) Click "Continue" underneath the number of studies that can still be extracted
12.) A list of studies will appear on the page. The next step is to filter these studies so that the only studies that appear are those that are used for the proficiency check
    a. Click the text box next to "Filter by tags" and select "Extraction.Proficiency"

All remaining studies are those that need to be completed for proficiency. You'll notice that they all have the tag "Extraction.Proficiency" marked in a dotted blue circle at the bottom of each entry. There may potentially be other tags there as well, but these can be ignored.

Once you've accessed the full texts, please complete the proficiency coding using the Data Extraction Template document on Google Drive as explained above.

During actual data extraction, you'll enter your extracted data into Covidence itself using an identical template.

Feel free to email me or text with any questions or concerns. Data extraction proficiency will be reviewed, and feedback will be provided as needed.

Data Extraction Procedures

When you screen, you will only need to have Covidence open. As mentioned above, I recommend having Covidence open in two separate windows to make data extraction more efficient. The extraction training manual and coding manual will be available in

Covidence

Link: https://www.covidence.org/

7.) Sign in
8.) Select "Contingent Reinforcement for Academic Performance" under "Your reviews"
9.) Select "Extraction" to open another dropdown menu
10.) Click "Continue" underneath the number of studies that can still be screened
11.) A list of studies will appear on the page. The next step is to filter these studies so that the only studies that appear are those that are used for the proficiency check
    a. Click the text box next to "Filter by tags" and select "Extraction.ISA"

All remaining studies are those that need to be completed for data extraction ISA. You'll notice that they all have the tag "Extraction.ISA" marked in a dotted blue circle at the bottom of each entry.

<u>Entering Your Decision into Covidence</u>

Coding can be saved mid-template if you need to stop partway through by selecting the gray "Save" button. Once all data have been extracted and the template is completely filled in, select the blue "Send for Consensus" button.

Background Resources

The "Resources" Google Drive folder also includes PDFs of three seminal articles that provide more in-depth information about contingency-based reinforcement. These articles may be useful resources to gain a greater understanding of the theory behind contingency-based reinforcement interventions, critiques of the intervention, the intervention's history in the school psychology field, and examples of real or hypothetical instances where these interventions have been implemented.

The included studies are:

Gresham, F. M., & Gresham, G. N. (1982). Interdependent, dependent, and independent group contingencies for controlling disruptive behavior. *The Journal of Special Education*, *16*(1), 101-110.

Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis*, *8*(3), 341.

Skinner, C. H., Williams, R. L., & Neddenriep, C. E. (2004). Using interdependent group-oriented reinforcement to enhance academic performance in general education classrooms. *School Psychology Review*, *33*(3), 384-397.

In terms of application to this specific project, I would especially urge you to review Skinner, Williams, and Neddenriep (2004). This article speaks in-depth about academic performance contingencies and the different forms that they take. (C. H. Skinner also wrote a similar chapter in NASP's *Best Practices in School Psychology IV,* entitled "Best Practices in Contingency Management: Application of Individual and Group Contingencies in Educational Setting

References

*Studies marked with an asterisk were included in the systematic review.

Akin-Little, K. A., Eckert, T. L., Lovett, B. J., & Little, S. G. (2004). Extrinsic reinforcement in the classroom: Bribery or best practice. *School Psychology Review*, *33*(3), 344-362. https://doi.org/10.1080/02796015.2004.12086253

Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher*, *49*(3), 198-211. https://doi.org/10.3102%2F0013189X20902110

Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behaviour Research and Therapy*, *31*(6), 621-631. https://doi.org/10.1016/0005-7967(93)90115-B

*Aloisio, D. B. (2006). *Improving homework completion and accuracy: Interdependent group contingencies with randomly selected components* (Publication No. 3239790) [Doctoral dissertation, Hofstra University]. ProQuest Dissertations & Theses Global.

*Alric, J. M., Bray, M. A., Kehle, T. J., Chafouleas, S. M., & Theodore, L. A. (2007). A comparison of independent, interdependent, and dependent group contingencies with randomized reinforcers to increase reading fluency. *Canadian Journal of School Psychology*, *22*(1), 81–93. https://doi.org/10.1177%2F0829573507301254

*Amado, R. S. (1982). *Some collateral effects of four academic reinforcement regimes (Publication No. 8209540* [Doctoral dissertation, University of Illinois, Urbana-Champaign]. ProQuest Dissertations & Theses Global.

American Psychological Association [APA]. (2020). *Publication manual of the American Psychological Association* (7th ed.). https://doi.org/10.1037/0000165-000

American Psychological Association [APA]. (2022). *APA Thesaurus of Psychological Index Terms*. April 19, 2022 from https://www.apa.org/pubs/databases/training/thesaurus

*Andersen, M. N., Daly III, E. J., & Young, N. D. (2013). Examination of a one-trial brief experimental analysis to identify reading fluency interventions. *Psychology in the Schools*, *50*(4), 403–414. https://doi.org/10.1002/pits.21682

*Ardi, V. (1989). *A comparative study examining the effects of two treatments on the disruptive behavior and reading performance of behavior disordered children* (Publication No. 8904184) [Doctoral Dissertation, Boston College]. ProQuest Dissertations & Theses Global.

Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding curriculum-based measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology*, *51*(1), 1-18. https://doi.org/10.1016/j.jsp.2012.09.004

Athens, E. S., Vollmer, T. R., & Pipkin, C. C. S. P. (2007). Shaping academic task engagement with percentile schedules. *Journal of Applied Behavior Analysis*, *40*(3), 475-488. https://doi.org/10.1901/jaba.2007.40-475

*Auge, M. (2021). *The effects of an interdependent group contingency on student writing* [Doctoral dissertation, University of Tennessee, Knoxville]. University of Tennessee, Knoxville Digital Archive. https://trace.tennessee.edu/cgi/viewcontent.cgi?article=8146&context=utk_graddiss

Axelrod, S., McGregor, G., Sherman, J., & Hamlet, C. (1987). Effects of video games as reinforcers for computerized addition performance. *Journal of Special Education Technology*, *9*(1), 1-8. https://doi.org/10.1177/016264348700900101

* Ayllon, T., Layman, D., & Burke, S. (1972). Disruptive behavior and reinforcement of academic performance. *The Psychological Record*, *22*(3), 315–323.

* Ayllon, T., Layman, D., & Kandel, H. J. (1975). A behavioral-educational alternative to drug control of hyperactive children. *Journal of Applied Behavior Analysis*, *8*(2), 137–146. https://doi.org/10.1901/jaba.1975.8-137

*Ayllon, T., & Roberts, M. D. (1974). Eliminating discipline problems by strengthening academic performance. *Journal of Applied Behavior Analysis*, *7*(1), 71–76. https://doi.org/10.1901/jaba.1974.7-71

*Baker, M. (2013). *The effects of using random independent and interdependent group reinforcers for homework completion* [Master's thesis, Western Illinois University].

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based Writing-to-Learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, *74*, 29–58. https://doi.org/10.3102%2F00346543074001029

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology*, *1*(3), 311-320. https://doi.org/10.1037%2F1089-2680.1.3.311

*Bear, G. G., & Richards, H. C. (1980). An interdependent group-oriented contingency system for improving academic performance. *School Psychology Review*, *9*(2), 190–193.

*Bennett, M. M. (2006). *An interdependent group contingency with mystery motivators to increase spelling performance* [Doctoral dissertation, Hofstra University].

*Benowitz, M. L., & Busse, T. V. (1976). Effects of material incentives on classroom learning over a four-week period. *Journal of Educational Psychology*, *68*(1), 57. https://doi-org.libezproxy2.syr.edu/10.1037/0022-0663.68.1.57

*Benowitz, M. L., & Rosenfeld, J. G. (1973). Three types of incentives and the classroom learning of middle- and lower-class children. *Psychology in the Schools*, *10*(1), 79–83. https://doi.org/10.1002/1520-6807(197301)10:1<79::AID-PITS2310100113>3.0.CO;2-6

*Billingsley, F. F. (1977). The effects of self-and externally-imposed schedules of reinforcement on oral reading performance. *Journal of Learning Disabilities*, *10*(9), 549–559. https://doi.org/10.1177/002221947701000904

*Bit, S. (1981). *A study of the effects of reward structures on academic achievement and sociometric status of Cambodian students* (Publication No. 8125893) [Doctoral dissertation, University of San Francisco]. ProQuest Dissertations & Theses Global.

Bonfiglio, C. M., Daly III, E. J., Martens, B. K., Lin, L. H. R., & Corsaut, S. (2004). An experimental analysis of reading interventions: Generalization across instructional strategies, time, and passages. *Journal of Applied Behavior Analysis*, *37*(1), 111-114. https://doi.org/10.1901/jaba.2004.37-111

Bowman-Perrott, L., Burke, M. D., Zaini, S., Zhang, N., & Vannest, K. (2016). Promoting positive behavior using the Good Behavior Game: A meta-analysis of single-case research. *Journal of Positive Behavior Interventions*, *18*(3), 180–190. https://doi.org/10.1177%2F1098300715592355

Breaux, R. P., Langberg, J. M., Bourchtein, E., Eadeh, H.-M., Molitor, S. J., & Smith, Z. R. (2019). Brief homework intervention for adolescents with ADHD: Trajectories and predictors of response. *School Psychology*, *34*(2), 201–211. https://doi.apa.org/doi/10.1037/spq0000287

*Brigham, T. A., Finfrock, S. R., Breunig, M. K., & Bushell Jr., D. (1972). The use of programmed materials in the analysis of academic contingencies. *Journal of Applied Behavior Analysis*, *5*(2), 177–182. https://doi.org/10.1901/jaba.1972.5-177

*Brooks, R. B., & Snow, D. L. (1972). Two case illustrations of the use of behavior-modification techniques in the school setting. *Behavior Therapy*, *3*(1), 100–103. https://doi.org/10.1016/S0005-7894(72)80057-1

*Broughton, S. F. (1983). The lack of vicarious effects with contingent teacher attention for academic performance. *Child and Family Behavior Therapy*, *5*(1), 25–36.

*Broughton, S. F., & Lahey, B. B. (1978). Direct and collateral effects of positive reinforcement, response cost, and mixed contingencies for academic performance. *Journal of School Psychology*, *16*(2), 126–136. https://doi.org/10.1016/0022-4405(78)90051-1

*Brownell, K. D., Colletti, G., Ersner-Hershfield, R., Hershfield, S. M., & Terence Wilson, G. (1977). Self-control in school children: Stringency and leniency in self-determined and externally imposed performance standards. *Behavior Therapy*, *8*(3), 442–445. https://doi.org/10.1016/S0005-7894(77)80080-4

Burton, B. N. (2012). *The effects of choice versus no-choice of task sequence on task engagement*. The University of Southern Mississippi.

Busk, P. L., & Serlin, R. (1992). Meta-analysis for single case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single case research design and analysis: New directions for psychology and education* (pp. 187-212). Hillsdale, NJ: Lawrence Erlbaum.

Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*, *64*(3), 363-423. https://doi.org/10.3102%2F00346543064003363

*Carlson, R. M., Hotchkiss, J. M., & Walker, H. M. (1970). *A report of the experienced teacher fellowship program on the effects of reinforcement on regular classroom academic performance*. U.S. Department of Health, Education, and Welfare. https://files.eric.ed.gov/fulltext/ED044358.pdf

*Carson, P. M., & Eckert, T. L. (2003). An experimental analysis of mathematics instructional components: Examining the effects of student-selected versus empirically-selected interventions. *Journal of Behavioral Education*, *12*(1), 35–54. https://doi.org/10.1023/A:1022370305486

Casillas, A., Robbins, S., Allen, J., Kuo, Y. L., Hanson, M. A., & Schmeiser, C. (2012). Predicting early academic failure in high school from prior academic achievement, psychosocial characteristics, and behavior. *Journal of Educational Psychology*, *104*, 407–420. https://doi.apa.org/doi/10.1037/a0027180

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin*, *140*(4), 980–1008. https://doi.apa.org/doi/10.1037/a0035661

Chadwick, B. A., & Day, R. C. (1971). Systematic reinforcement: Academic performance of underachieving students. *Journal of Applied Behavior Analysis*, *4*(4), 311–319. https://doi.org/10.1901/jaba.1971.4-311

Chaffee, R. K., Briesch, A. M., Volpe, R. J., Johnson, A. H., & Dudley, L. (2020). Effects of a class-wide positive peer reporting intervention on middle school student behavior. *Behavioral Disorders*, *45*(4), 224-237. https://doi-org.libezproxy2.syr.edu/10.1177/0198742919881112

*Christ, T. J., & Schanding, G. T., Jr. (2007). Curriculum-based measures of computational skills: A comparison of group performance in novel, reward, and neutral conditions. *School Psychology Review*, *36*(1), 147–158.

Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, *47*(5), 291–314. https://doi.org/10.1016/j.jsp.2009.04.002

Codding, R. S., Chan-Iannetta, L., George, S., Ferreira, K., & Volpe, R. (2011). Early number skills: Examining the effects of class-wide interventions on kindergarten performance. *School Psychology Quarterly*, *26*(1), 85. https://doi.org/10.1037/a0022661

*Coleman, R. (1970). A pilot demonstration of the utility of reinforcement techniques in trainable programs. *Education and Training of the Mentally Retarded*, *5*(2), 68. https://www.jstor.org/stable/23876241

Coogan, B. A., Kehle, T. J., Bray, M. A., & Chafouleas, S. M. (2007). Group contingencies, randomization of reinforcers, and criteria for reinforcement, self-monitoring, and peer feedback on reducing inappropriate classroom behavior. *School Psychology Quarterly*, *22*(4), 540–556. https://doi.org/10.1037/1045-3830.22.4.540

*Cracco, J. J. (2006). *The effects of interdependent group contingencies on homework completion and accuracy of special needs adolescents* [Doctoral dissertation, Hofstra University].

Crouch, P. L., Gresham, F. M., & Wright, W. R. (1985). Interdependent and independent group contingencies with immediate and delayed reinforcement for controlling classroom behavior. *Journal of School Psychology*, *23*(2), 177–187. https://doi.org/10.1016/0022-4405(85)90008-1

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7-29.

https://doi.org/10.1177%2F0956797613504966

*Dalton, A. J., Rubino, C. A., & Hislop, M. W. (1973). Some effects of token rewards on school achievement of children with Down's syndrome. *Journal of Applied Behavior Analysis*, *6*(2), 251–259. https://doi.org/10.1901/jaba.1973.6-251

*Daly, E. J. I., Persampieri, M., McCurdy, M., & Gortmaker, V. (2005). Generating reading interventions through experimental analysis of academic skills: Demonstration and empirical evaluation. *School Psychology Review*, *34*(3), 395–414. https://doi.org/10.1016/j.jsp.2009.02.003

Daly III, E. J., Wells, N. J., Swanger-Gagné, M. S., Carr, J. E., Kunz, G. M., & Taylor, A. M. (2009). Evaluation of the multiple-stimulus without replacement preference assessment method using activities as stimuli. *Journal of Applied Behavior Analysis*, *42*(3), 563-574. https://doi.org/10.1901/jaba.2009.42-563

Daly III, E. J., Witt, J.C., Martens, B. K., & Dool, E.J. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review*, *26*(4), 554-574. https://doi.org/10.1080/02796015.1997.12085886

de Brey, C., Musu, L., McFarland, J., Wilkinson-Flicker, S., Diliberti, M., Zhang, A., Branstetter, C., Wang, X., National Center for Education Statistics (ED), & American Institutes for Research (AIR). (2019). Status and trends in the education of racial and ethnic groups 2018. NCES 2019-038. In *National Center for Education Statistics*. National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019038

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627–668. https://doi.org/10.1037/0033-2909.125.6.627

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, *29*(4), 519-533. https://doi.org/10.1901/jaba.1996.29-519

*Denison, A. J. (2013). *Utility of an error analysis and performance deficit assessment for selecting brief interventions to increase math fluency* [Master's Thesis, Utah State University]. Utah State University Digital Archive. https://doi.org/10.26076/4859-beb2

Deshais, M. A., Fisher, A. B., & Kahng, S. (2018). A preliminary investigation of a randomized dependent group contingency for hallway transitions. *Education and Treatment of Children*, *41*(1), 49–63. https://doi.org/10.1353/etc.2018.0002

Deshais, M. A., Fisher, A. B., & Kahng, S. (2019). A comparison of group contingencies on academic compliance. *Journal of Applied Behavior Analysis*, *52*(1), 116–131. https://doi.org/10.1002/jaba.505

*Dickerson, E. A., & Creedon, C. F. (1981). Self-selection of standards by children: The relative effectiveness of pupil-selected and teacher-selected standards of performance. *Journal of Applied Behavior Analysis*, *14*(4), 425–433. https://doi.org/10.1901/jaba.1981.14-425

*Diedrick, L. A. (2009). *A pictorial reinforcement intervention targeting homework submission, completion, and accuracy rates of students in a restrictive classroom environment* (Publication No. 3322199) [Doctoral dissertation, Indiana State University].

Dion, E., Roux, C., Landry, D., Fuchs, D., Wehby, J., & Dupéré, V. (2011). Improving attention and preventing reading difficulties among low-income first-graders: A randomized study. *Prevention Science*, *12*(1), 70-79. https://doi.org/10.1007/s11121-010-0182-5

DiPerna, J. C., Volpe, R. J., & Elliott, S. N. (2002). A model of academic enablers and elementary reading/language arts achievement. *School Psychology Review*, *31*(3), 298-312. https://doi-org.libezproxy2.syr.edu/10.1080/02796015.2002.12086157

DiPerna, J. C., Volpe, R. J., & Elliott, S. N. (2005). A model of academic enablers and mathematics achievement in the elementary grades. *Journal of School Psychology*, *43*(5), 379-392. https://doi.org/10.1016/j.jsp.2005.09.002

*Doherty, M. A. (1981). *Effects of reinforcement and motivational orientation on children's academic behavior as a function of self-determined and externally imposed contingencies* (Order No. 8124562) [Doctoral dissertation, Temple University].

Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., Laudolff, J., Turkkan, J. S., Ford, C., & Wheeler, L. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology, 14*(3), 317–345. https://doi.org/10.1016/0193-3973(93)90013-L

*Duhon, G. J., Noell, G. H., Witt, J. C., Freeland, J. T., Dufrene, B. A., & Gilbertson, D. N. (2004). Identifying academic skill and performance deficits: The experimental analysis of brief assessments of academic skills. General Articles. *School Psychology Review*, *33*(3), 429–443. https://doi.org/10.1080/02796015.2004.12086260

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., Japel, C., & American Psychological Association, W. D. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.apa.org/doi/10.1037/0012-1649.43.6.1428

*Eckert, T. L., Ardoin, S. P., Daisey, D. M., & Scarola, M. D. (2000). Empirically evaluating the effectiveness of reading interventions: The use of brief experimental analysis and single case designs. *Psychology in the Schools*, *37*(5), 463–473. https://doi.org/10.1002/1520-6807(200009)37:5<463::AID-PITS6>3.0.CO;2-X

Eckert, T. L., Ardoin, S. P., Daly, E. J., III, & Martens, B. K. (2002). Improving oral reading fluency: A brief experimental analysis of combining an antecedent intervention with consequences. *Journal of Applied Behavior Analysis*, *35*(3), 271–281. https://doi.org/10.1901/jaba.2002.35-271

Fabiano, G. A., Pelham Jr, W. E., Karmazin, K., Kreher, J., Panahon, C. J., & Carlson, C. (2008). A group contingency program to improve the behavior of elementary school students in a cafeteria. *Behavior Modification*, *32*(1), 121-132. https://doi.org/10.1177/0145445507308577

*Felixbrod, J. J., & O'Leary, K. D. (1973). Effects of reinforcement on children's academic behavior as a function of self-determined and externally imposed contingencies. *Journal of Applied Behavior Analysis*, *6*(2), 241–250. https://doi.org/10.1901/jaba.1973.6-241

*Felixbrod, J., & O'Leary, K. D. (1974). Self-determination of academic standards by children: Toward freedom from external control. *Journal of Educational Psychology*, *66*(6), 845–850. https://doi.org/10.1037/h0021526

Ferneza, P., Jabot, M., & Maheady, L. (2013). The effects of group contingencies with randomized components on the homework accuracy of an 11th-and 12th-grade general science class. *Journal of Evidence-Based Practices for Schools*, *13*(2), 100-119. Retrieved from https://link.gale.com/apps/doc/A500501582/AONE?u=nysl_ce_syr&sid=AONE&xid=ad571d7d

*Ferritor, D. E., Buckholdt, D., Hamblin, R. L., & Smith, L. (1972). The noneffects of contingent reinforcement for attending behavior on work accomplished. *Journal of Applied Behavior Analysis*, *5*(1), 7–17. https://doi.org/10.1901/jaba.1972.5-7

Fields, R. (2014). Towards the National Assessment of Educational Progress (NAEP) as an indicator of academic preparedness for college and job training. *National Assessment Governing Board*. https://files.eric.ed.gov/fulltext/ED582999.pdf

Fiscella, K., & Kitzman, H. (2009). Disparities in academic achievement and health: the intersection of child education and health policy. *Pediatrics*, *123*(3), 1073-1080. https://doi.org/10.1542/peds.2008-0533

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis*, *25*(2), 491-498. https://dx.doi.org/10.1901%2Fjaba.1992.25-491

*Fowler, R. E., Thomas, L., & Santogrossi, D. A. (1977). Using contingent access to activities to increase accurate responding. Education and Treatment of Children, 1(1), 5–8. https://www.jstor.org/stable/42898404

*Freeland, J. T., & Noell, G. H. (1999). Maintaining Accurate Math Responses in Elementary School Students: The Effects of Delayed Intermittent Reinforcement and Programming Common Stimuli. *Journal of Applied Behavior Analysis*, *32*(2), 211–215. https://doi.org/10.1901/jaba.1999.32-211

Fritz, J. N., Roath, C. T., Shoemaker, P. T., Edwards, A. B., Hussein, L. A., Villante, N. K., Langlinais, C. A., & Rettig, L. A. (2020). Validity of the multiple-stimulus without replacement preference assessment for edible items. *Journal of Applied Behavior*

Analysis. https://doi.org/10.1002/jaba.703

Froiland, J. M., & Worrell, F. C. (2016). Intrinsic motivation, learning goals, engagement, and achievement in a diverse high school. *Psychology in the Schools*, *53*(3), 321-336. https://doi.org/10.1002/pits.21901

Galbicka, G. (1994). Shaping in the 21st century: Moving percentile schedules into applied settings. *Journal of Applied Behavior Analysis*, *27*(4), 739-760. https://dx.doi.org/10.1901%2Fjaba.1994.27-739

Gansle, K. A., Noell, G. H., & Freeland, J. T. (2002). Can't Jane read or won't Jane read? An analysis of pre-reading skills designed to differentiate skill deficits from performance deficits. *The Behavior Analyst Today*, *3*(2), 161–165. http://dx.doi.org/10.1037/h0099976

Gest, S. D., & Gest, J. M. (2005). Reading tutoring for students at academic and behavioral risk: Effects on time-on-task in the classroom. *Education and Treatment of Children*, *28*(1), 25–47. Retrieved from https://www.jstor.org/stable/42899826

*Gilbert, S. (1990). *Effects of competition, group reinforcement, and individual reinforcement on achievement in spelling* (Publication No. 9017170) [Doctoral dissertation, Hofstra University]. ProQuest Dissertations & Theses Global.

Gilbertson, D., Witt, J. C., Duhon, G., & Dufrene, B. (2008). Using brief assessments to select math fluency and on-task behavior interventions: An investigation of treatment utility. *Education and Treatment of Children*, *31*(2), 167-181. https://doi.org/10.1353/etc.0.0023

Gobel, P. (2008). Student off-task behavior and motivation in the CALL classroom. *International Journal of Pedagogies and Learning, 4*, 4-18. https://doi.org/10.5172/ijpl.4.4.4

Graham, S. (1982). Composition research and practice: A unified approach. *Focus on Exceptional Children*, *14*, 1-16. https://doi.org/10.17161/foec.v14i8.7447

Graham, S., MacArthur, C. A., & Fitzgerald, J. (2007). Introduction. In Graham, S., MacArthuer, C.A., & Fitzgerald, J. (Eds.), *Best practices in writing instruction* (1-9). Guilford Publications.

Graplin, S. L., & Kranzler, J. H. [Eds.] (2018). *School psychology: Professional issues and practices*. NY: Springer Publishing.

Green, C. W., Reid, D. H., White, L. K., Halford, R. C., Brittain, D. P., & Gardner, S. M. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. *Journal of Applied Behavior Analysis*, *21*(1), 31-43. https://doi.org/10.1901/jaba.1988.21-31

Gresham, F. M. (1981). Assessment of children's social skills. *Journal of School Psychology*. *19,* 120-133. https://doi.org/10.1016/0022-4405(81)90054-6

Gresham, F. M., & Gresham, G. N. (1982). Interdependent, dependent, and independent group contingencies for controlling disruptive behavior. *Journal of Special Education*, *16*(1), 101–110. https://doi.org/10.1177%2F002246698201600110

Grigorenko, E. L., Jarvin, L., Diffley III, R., Goodyear, J., Shanahan, E. J., & Sternberg, R. J. (2009). Are SSATS and GPA enough? A theory-based approach to predicting academic success in secondary school. *Journal of Educational Psychology*, *101*(4), 964-981. https://doi.org/10.1037/a0015906

Gross, T. J., Duhon, G. J., Shutte, G., & Rowland, J. E. (2016). A comparison of group-oriented contingencies for addition fluency. *Contemporary School Psychology*, *20*(2), 130–141. http://dx.doi.org/10.1007/s40688-015-0054-x

Haahr, M. (2022). *Random.org: True random number service*. https://www.random.org

*Hardy, C. M. (2014). *A comparison of behavioral and academically-focused goals within the mystery motivator: Effects on disruptive behavior and academic performance* [Doctoral Dissertation, University of Southern Mississippi]. University of Southern Mississippi Digital Archive. https://aquila.usm.edu/dissertations/226/

*Hargis, D. Z. (2012). *Increasing homework completion and accuracy among mathematics students using the Jars Game* [Master's project, State University of New York at Fredonia]. State University of New York Digital Archive. https://soar.suny.edu/handle/20.500.12648/476

Haring, N. G., & Eaton, M.D. (1978). Systematic instructional procedures: An instructional hierarchy. In N.G. Haring, T.C. Lovitt, M.D. Eaton & C.L. Hansen (Eds.), *The fourth R: Research in the classroom* (pp. 23-40). Columbus, OH: Merrill.

*Harris, V. W., & Sherman, J. A. (1974). Homework assignments, consequences, and classroom performance in social studies and mathematics. *Journal of Applied Behavior Analysis*, *7*(4), 505–519. https://doi.org/10.1901/jaba.1974.7-505

*Hauserman, N., & McIntire, R. (1969). Training elementary reading skills through reinforcement and fading techniques. *Proceedings of the Annual Convention of the American Psychological Association*. 669–670.

*Hay, W. H., Hay, L. R., & Nelson, R. O. (1977). Direct and collateral changes in on-task and academic behavior resulting from on-task versus academic contingencies. *Behavior Therapy*, *8*(3), 431–441. https://doi.org/10.1016/S0005-7894(77)80079-8

Heering, P. W., & Wilder, D. A. (2006). The use of dependent group contingencies to increase on-task behavior in two general education classrooms. *Education and Treatment of*

Children, *29*, 459–467. Retrieved from https://www.jstor.org/stable/42899896

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

*Hoff, N. T. (2020). *An evaluation of a reinforcer validation assessment* (Publication No. 22589051) [Doctoral dissertation, University of Nebraska Lincoln]. https://digitalcommons.unl.edu/dissertations/AAI22589051/

*Hofstadter-Duke, K. L., & Daly, E. J. (2015). Identifying controlling variables for math computation fluency through experimental analysis: The interaction of stimulus control and reinforcing consequences. *Behavior Modification*, *39*(2), 342–364. https://doi.org/10.1177/0145445514559928

*Holt, G. L. (1971). Effect of reinforcement contingencies in increasing programmed reading and mathematics behaviors in first-grade children. *Journal of Experimental Child Psychology*, *12*(3), 362. https://doi.org/10.1016/0022-0965(71)90031-2

*Hopkins, B. L., Schutte, R. C., & Garton, K. L. (1971). The effects of access to a playroom on the rate and quality of printing and writing of first and second-grade students. *Journal of Applied Behavior Analysis*, *4*(2), 77–87. https://doi.org/10.1901/jaba.1971.4-77

Hussar, B., Zhang, J., Hein, S., Wang, K., Roberts, A., Cui, J., Smith, M., Mann, F. B., Barmer, A., Dilig, R., National Center for Education Statistics (ED), & American Institutes for Research (AIR). (2020). The condition of education 2020. NCES 2020-144. In *National Center for Education Statistics*. National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020144

Ialongo, N. S., Werthamer, L., Kellam, S. G., Brown, C. H., Wang, S., & Lin, Y. (1999).

Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression, and antisocial behavior. *American Journal of Community Psychology*, *27*(5), 599-641. https://doi.org/10.1023/a:1022137920532

January, S. A. A., & Klingbeil, D. A. (2020). Universal screening in grades K-2: A systematic review and meta-analysis of early reading curriculum-based measures. *Journal of School Psychology*, *82*, 103-122. https://doi.org/10.1016/j.jsp.2020.08.007

Kelshaw-Levering, K., Sterling-Turner, H. E., Henry, J. R., & Skinner, C. H. (2000). Randomized interdependent group contingencies: Group reinforcement with a twist. *Psychology in the Schools, 37*(6), 523–533. https://doi.org/10.1002/1520-6807(200011)37:6<523::AID-PITS5>3.0.CO;2-W

King, S. A., & Kostewicz, D. E. (2014). Choice-based stimulus preference assessment for children with or at-risk for emotional disturbance in educational settings. *Education and Treatment of Children*, 531-558. https://doi.org/10.1353/etc.2014.0026

*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*(1), 79–84. https://doi.org/10.1901/jaba.1972.5-79

Klingbeil, D. A., Dart, E. H., & Schramm, A. L. (2019). A systematic review of function-modified check-in/check-out. *Journal of Positive Behavior Interventions*, *21*(2), 77-92. https://doi-org.libezproxy2.syr.edu/10.1177/1098300718778032

Koenig, E. A., Eckert, T. L., & Hier, B. O. (2016). Using performance feedback and goal setting to improve elementary students' writing fluency: A randomized controlled trial. *School Psychology Review*, *45*(3), 275–295. https://doi-org.libezproxy2.syr.edu/10.17105/SPR45-3.275-295

Kohn, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Houghton, Mifflin and Company.

Kowalewicz, E. A., & Coffee, G. (2014). Mystery Motivator: A Tier 1 classroom behavioral intervention. *School Psychology Quarterly*, *29*(2), 138. https://doi.org/10.1037/spq0000030

*Lahey, B. B., McNees, M. P., & Brown, C. C. (1973). Modification of Deficits in Reading for Comprehension. *Journal of Applied Behavior Analysis*, *6*(3), 475–480. https://doi.org/10.1901/jaba.1973.6-475

*Landy, K. (2013). *The effects of the Tree Jars Intervention on the homework completion and accuracy of middle school mathematics students* [Master's project, State University of New York at Fredonia]. State University of New York Digital Archive.

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*(2). https://doi.org/10.3102%2F0034654315617832

Lentz, F. E. (1988). Effective reading interventions in the regular classroom. In J. L. Graden, J. Zins, & M. J. Curtis [Eds.] *Alternative educational delivery systems: Enhancing instructional options for all students* (pp. 351-370). Washington, DC: National Association of School Psychologists.

Ling, S. M., & Barnett, D. W. (2013). Increasing preschool student engagement during group learning activities using a group contingency. *Topics in Early Childhood Special Education, 33*(3), 186–196. https://doi.org/10.1177/0271121413484595

Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis, 8,* 341-347. https://doi.org/10.1901/jaba.1975.8-

341

*Little, S. G., Akin-Little, A., & Newman-Eig, L. M. (2010). Effects on homework completion and accuracy of varied and constant reinforcement within an interdependent group contingency system. *Journal of Applied School Psychology*, *26*(2), 115–131. https://doi.org/10.1080/15377900903471989

Little, S. G., Akin-Little, A., & O'Neill, K. (2015). Group contingency interventions with children—1980-2010: A meta-analysis. *Behavior Modification*, *39*(2), 322-341. https://doi.org/10.1177%2F0145445514554393

Long, A. C. J., Miller, F. G., & Upright, J. J. (2019). Classroom management for ethnic–racial minority students: A meta-analysis of single-case design studies. *School Psychology*, *34*(1), 1–13. https://doi.org/10.1037/spq0000305

*Lynch, A., Theodore, L. A., Bray, M. A., & Kehle, T. J. (2009). A comparison of group-oriented contingencies and randomized reinforcers to improve homework completion and accuracy for students with disabilities. *School Psychology Review*, *38*(3), 307–324.

*Madaus, M. M. R., Kehle, T. J., Madaus, J., & Bray, M. A. (2003). Mystery motivator as an intervention to promote homework completion and accuracy. *School Psychology International*, *24*(4), 369–377. https://doi.org/10.1177/01430343030244001

*Madsen, C. K., & Forsythe, J. L. (1973). Effect of contingent music listening on increases of mathematical responses. *Journal of Research in Music Education*, *21*(2), 176. https://doi.org/10.2307/3344594

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*(5), 625-654.

https://doi.org/10.1016/j.jsp.2012.06.001

Maggin, D. M., Pustejovsky, J. E., & Johnson, A. H. (2017). A meta-analysis of school-based group contingency interventions for students with challenging behavior: An update. *Remedial and Special Education*, *38*(6), 353–370. https://doi.org/10.1177/0741932517716900

Malecki, C. K., & Elliot, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly*, *17*(1), 1–23. https://doi-org.libezproxy2.syr.edu/10.1521/scpq.17.1.1.19902

*Marholin II, D., & Steinman, W. M. (1977). Stimulus control in the classroom as a function of the behavior reinforced. *Journal of Applied Behavior Analysis*, *10*(3), 465–478. https://doi.org/10.1901/jaba.1977.10-465

Martens, B. K., Lochner, D. G., & Kelly, S. Q. (1992). The effects of variable-interval reinforcement on academic engagement: A demonstration of matching theory. *Journal of Applied Behavior Analysis*, *25*(1), 143–151. https://doi.org/10.1901/jaba.1992.25-143

McCurdy, M., Skinner, C., McClurg, V., Whitsitt, L., & Moore, T. (2020). Bonus rewards for everyone: Enhancing mathematics performance with supplemental interdependent group contingencies. *Preventing School Failure*, *64*(1), 77–88. https://doi.org/10.1080/1045988X.2019.1678008

*McEvoy, M. A., & Brady, M. P. (1988). Contingent access to play materials as an academic motivator for autistic and behavior disordered children. *Education and Treatment of Children*, *11*(1), 5–18. https://www.jstor.org/stable/42899046

*McLaughlin, T. F. & And Others. (1980). The effects of individual and group contingencies on spelling performance for a special education class. *B. C. Journal of Special Education*, *4*(3), 263–270.

*Melton, J. E. (1970). *An investigation of the differential probability principle as a reinforcement technique for facilitating academic performance* (Publication No. 6922288) [Master's thesis, University of Nebraska, Lincoln]. ProQuest Dissertations & Theses Global.

*Metallo, S. (2015). *Using a dependent group contingency to increase homework completion and accuracy in a general education classroom* [Master's thesis, Louisiana State University]. Louisiana State University Digital Archive. https://digitalcommons.lsu.edu/gradschool_theses/20/

Miltenberger, R. G. (2016). *Behavior modification: Principles and procedures* (6th ed.). Cengage Learning.

Moffett, L., & Morrison, F. J. (2020). Off-task behavior in kindergarten: Relations to executive function and academic achievement. *Journal of Educational Psychology*, *112*(5), 938–955. https://doi-org.libezproxy2.syr.edu/10.1037/edu0000397

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*, Article e1000097. https://doi.org/10.1371/journal.pmed1000097

Moore, D. W., Prebble, S., Robertson, J., Waetford, R., & Anderson, A. (2001). Self-recording with goal setting: A self-management programme for the classroom. *Educational Psychology*, *21*(3), 255–265. https://doi-org.libezproxy2.syr.edu/10.1080/01443410120065469

Murphy, M., Mathis, E., Weaver, A. D., & Dart, E. (2019). Using Antecedent-Based Strategies to Address Motivation in Behavioral Interventions. *Communique*, *48*(2), 1-23. Retrieved from https://apps.nasponline.org/resources-and-publications/periodicals/cq-index-list.aspx

Murphy, K. A., Theodore, L. A., Aloiso, D., Alric-Edwards, J. M., & Hughes, T. L. (2007). Interdependent group contingency and mystery motivators to reduce preschool disruptive behavior. *Psychology in the Schools*, *44*(1), 53-63. https://doi.org/10.1002/pits.20205

National Center for Educational Statistics. (2007). *The nation's report card: Writing 2007 (NCES 2008–468)*. Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2012a). *The Nation's Report Card: Writing 2011* (NCES 2012–470). U.S. Department of Education.

National Center for Education Statistics. (2012b). *Schools and staffing survey: Public teacher questionnaire, selected years 2011-12*. Washington, DC: U.S. Department of Education.

National Center for Educational Statistics. (2015). *The Nation's Report Card: 2014 U.S. History, Geography, and Civics at Grade 8* (NCES 2015-112). U.S. Department of Education.

National Center for Educational Statistics. (2018). *An Overview of NAEP*. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2013455

National Center for Educational Statistics. (2020). *Intended Meaning of NAEP*. Retrieved from https://nces.ed.gov/nationsreportcard/guides/

National Commission on Writing. (2003). The neglected "R": The need for a writing revolution. *Report of The National Commission on Writing in America's Schools and Colleges*. College Entrance Examination Board. Retrieved from https://files.eric.ed.gov/fulltext/ED475856.pdf

National Commission on Writing. (2005). Writing: A powerful message from state government. *Report of The National Commission on Writing in America's Schools and Colleges*.

College Entrance Examination Board. Retrieved from https://archive.nwp.org/cs/public/download/nwp_file/21492/powerful-message-from-state.pdf?x-r=pcfile_d

Newstrom, J., McLaughlin, T. F., & Sweeney, W. J. (1999). The effects of contingency contracting to improve the mechanics of written language with a middle school student with behavior disorders. *Child & Family Behavior Therapy*, *21*(1), 39-48. https://doi.org/10.1300/j019v21n01_03

Niehaus, K., Rudasill, K. M., & Adelson, J. L. (2012). Self-efficacy, intrinsic motivation, and academic outcomes among Latino middle school students participating in an after-school program. *Hispanic Journal of Behavioral Sciences*, *34*(1), 118-136. https://doi.org/10.1177%2F0739986311424275

*Noell, G. H., Freeland, J. T., Witt, J. C., & Gansle, K. A. (2001). Using brief assessments to identify effective interventions for individual students. *Journal of School Psychology*, *39*(4), 335–355. https://doi.org/10.1016/S0022-4405(01)00072-3

Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. T. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly*, *12*(1), 77–88. https://doi.org/10.1037/h0088949

Northup, J., George, T., Jones, K., Broussard, C., & Vollmer, T. R. (1996). A comparison of reinforcer assessment methods: The utility of verbal and pictorial choice procedures. *Journal of Applied Behavior Analysis, 29*(2), 201–212. https://doi.org/10.1901/jaba.1996.29-201

*O'Connor, M. A., & Daly, E. J. I. (2018). Selecting effective intervention strategies for escapemaintained academic-performance problems: Consider giving 'em a break! *Journal of School Psychology*, *66*, 41–53. https://doi.org/10.1016/j.jsp.2017.09.003

Pace, G. M., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis*, *18*(3), 249-255. https://doi.org/10.1901/jaba.1985.18-249

*Panahon, C. J., & Martens, B. K. (2013). A comparison of noncontingent plus contingent reinforcement to contingent reinforcement alone on students' academic performance. *Journal of Behavioral Education*, *22*(1), 37–49. https://doi.org/10.1007/s10864-012-9157-x

Park, J., Kim, S., Kim, A., & Yi, M. Y. (2019). Learning to be better at the game: Performance vs. completion contingent reward for game-based learning. *Computers & Education*, *139*, 1–15. https://doi.org/10.1016/j.compedu.2019.04.016

Pigott, H. E., Fantuzzo, J. W., & Clement, P. W. (1986). The effects of reciprocal peer tutoring and group contingencies on the academic performance of elementary school children. Journal of Applied Behavior Analysis, *19*(1), 93-98. https://doi.org/10.1901/jaba.1986.19-93

*Pipkin, C. St. P., Winters, S. M., & Diller, J. W. (2007). Effects of instruction, goals, and reinforcement on academic behavior: Assessing skill versus reinforcement deficits. *Journal of Early and Intensive Behavior Intervention*, *4*(4), 648–657. https://doi.org/10.1037/h0100398

Pokorski, E. A., Barton, E. E., & Ledford, J. R. (2017). A review of the use of group contingencies in preschool settings. *Topics in Early Childhood Special Education*, *36*(4), 230–241. https://doi.org/10.1177/0271121416649935

Popkin, J., & Skinner, C. H. (2003). Enhancing academic performance in a classroom serving students with serious emotional disturbance: Interdependent group contingencies with randomly selected components. *School Psychology Review*, *32*(2), 282–295. https://doi.org/10.1080/02796015.2003.12086199

Radley, K. C., Dart, E. H., Battaglia, A. A., & Ford, W. B. (2019). A comparison of two procedures for assessing preference in a classroom setting. *Behavior Analysis in Practice*, *12*(1), 95-104. https://doi.org/10.1007/s40617-018-0244-x

*Ralston, S. B. (2012). *Effect of dependent group contingencies on homework completion and accuracy* (Publication No. 3476608) [Doctoral dissertation, University of Connecticut]. University of Connecticut Digital Archive. https://opencommons.uconn.edu/dissertations/AAI3476608/

Reed, D. D., & Azulay, R. L. (2011). A Microsoft Excel® 2010 based tool for calculating interobserver agreement. *Behavior Analysis in Practice*, *4*(2), 45-52. https://doi-org.libezproxy2.syr.edu/10.1007/BF03391783

*Reinhardt, D., Theodore, L. A., Bray, M. A., & Kehle, T. J. (2009). Improving homework accuracy: Interdependent group contingencies and randomized components. *Psychology in the Schools*, *46*(5), 471–488. https://doi.org/10.1002/pits.20391

*Reisener, C. D. (2009). *Selecting effective mathematics interventions in the RTI process via brief experimental analyses* [Doctoral dissertation, University of Southern Mississippi]. University of Southern Mississippi Digital Archive. https://aquila.usm.edu/dissertations/1072/

*Reiss, D. J. & And Others. (1974). *The effects of contingent use of recess activities on the academic performance of a third grade classroom* [Paper presentation]. Annual Meeting of the American Educational Research Association, April 1974, Chicago, Illinois. https://files.eric.ed.gov/fulltext/ED091371.pdf

Resetar, J. L., & Noell, G. H. (2008). Evaluating preference assessments for use in the general education population. *Journal of Applied Behavior Analysis*, *41*(3), 447-451. https://doi.org/10.1901/jaba.2008.41-447

*Rieth, H. J., Polsgrove, L., Raia, S., Patterson, N., & Buchman, K. (1977). The use of free time to increase the reading achievement of three students placed in programs for behavior disordered children. *Behavioral Disorders*, *3*(1), 45–54. https://doi.org/10.1177/019874297700300107

Roane, H. S., Vollmer, T. R., Ringdahl, J. E., & Marcus, B. A. (1998). Evaluation of a brief stimulus preference assessment. *Journal of Applied Behavior Analysis*, *31*(4), 605-620. https://doi.org/10.1901/jaba.1998.31-605

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261–288. https://doi.org/10.1037/0033-2909.130.2.261

*Robin, H. L. (2014). *The effects of mystery motivators on homework completion and accuracy of a 10th grade mathematics class* [Master's project, State University of New York at Fredonia]. State University of New York Digital Archive. https://soar.suny.edu/handle/20.500.12648/412

Romeo, F. F. (1998). The negative effects of using a group contingency system of classroom management. *Journal of Instructional Psychology, 25,* 130-134.

*Rosenberg, M. S., Sindelar, P. T., & Stedt, J. (1985). The effects of supplemental on-task contingencies on the acquisition of simple and difficult academic tasks. *The Journal of Special Education*, *19*(2), 189–203. https://doi.org/10.1177/002246698501900206

*Rosenfeld, G. W. (1972). Some effects of reinforcement on achievement and behavior in a regular classroom. *Journal of Educational Psychology*, *63*(3), 189–193. https://doi.org/10.1037/h0032650

*Ross, P. A., & Braden, J. P. (1991). The effects of token reinforcement versus cognitive behavior modification on learning-disabled students' math skills. *Psychology in the Schools*, *28*(3), 247–256. https://doi.org/10.1002/1520-6807(199107)28:3<247::AID-PITS2310280310>3.0.CO;2-E

Roth, P. L., & Clarke, R. L. (1998). Meta-analyzing the relation between grades and salary. *Journal of Vocational Behavior*, *53*(3), 386-400. https://doi.org/10.1006/jvbe.1997.1621

*Rothberg, C. I. (1973). *Ethnic factors and the effectiveness of individual and group contingencies upon performance* [Doctoral dissertation, University of New Mexico]. University of New Mexico Digital Archive. https://digitalrepository.unm.edu/educ_ifce_etds/90/

*Schellenberg, T., Skok, R., & McLaughlin, T. F. (1991). The effects of contingent free time on homework completion in English with senior high school English students. *Child & Family Behavior Therapy, 13*, 1–11. https://doi.org/10.1300/J019v13n03_01

*Schunk, D. H. (1983). Reward contingencies and the development of children's skills and self-efficacy. *Journal of Educational Psychology*, *75*(4), 511–518. https://doi.org/10.1037/0022-0663.75.4.511

Schunk, D. H. (1984). Enhancing self-efficacy and achievement through rewards and goals:

Motivational and informational effects. *Journal of Educational Research*, *78*(1), 29–34. https://doi.org/10.1080/00220671.1984.10885568

*Scott, K. C., Skinner, C. H., Moore, T. C., McCurdy, M., Ciancio, D., & Cihak, D. F. (2017). Evaluating and comparing the effects of group contingencies on mathematics accuracy in a first-grade classroom: Class average criteria versus unknown small-group average criteria. *School Psychology Review*, *46*(3), 262–271. https://doi.org/10.17105/SPR-2017-0037.V46-3

Shapiro, E. S. (2011). *Academic skills problems: Direct assessment and intervention* (4th ed.). Guilford Press.

*Shapiro, E. S., & Goldberg, R. (1986). A comparison of group contingencies for increasing spelling performance among sixth grade students. *School Psychology Review*, *15*(4), 546-557. https://doi-org.libezproxy2.syr.edu/10.1080/02796015.1986.12085256

*Sharp, S. R., & Skinner, C. H. (2004). Using interdependent group contingencies with randomly selected criteria and paired reading to enhance class-wide reading performance. *Journal of Applied School Psychology*, *20*(2), 29–45. https://doi.org/10.1300/J370v20n02_03

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology, 70*, 747-770. https://doi.org/10.1146/annurev-psych-010418-102803

*Simon, S. J., Ayllon, T., & Milan, M. A. (1982). Behavioral compensation: Contrastlike effects in the classroom. *Behavior Modification*, *6*(3), 407–420. https://doi.org/10.1177/014544558263006

Skinner, C. H., Williams, R. L., & Neddenriep, C. E. (2004). Using interdependent group-oriented reinforcement to enhance academic performance in general education classrooms. *School Psychology Review*, *33*, 384–397. https://doi.org/10.1080/02796015.2004.12086255

*Slavin, R. E. (1978). Student teams and comparison among equals: Effects on academic performance and student attitudes. *Journal of Educational Psychology*, *70*(4), 532–538. https://doi.org/10.1037/0022-0663.70.4.532

*Speltz, M. L., Shimamura, J. W., & McReynolds, W. T. (1982). Procedural variations in group contingencies: Effects on children's academic and social behaviors. *Journal of Applied Behavior Analysis*, *15*(4), 533–544. https://doi.org/10.1901/jaba.1982.15-533

Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review, 26*(3), 333–368. https://doi.org/10.1080/02796015.1997.12085871

Stipek, D. J. (1996). Motivation and instruction. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology*. (pp. 85–113). Macmillan Library Reference USA.

*Strandy, C., McLaughlin, T. F., & Hunsaker, D. (1979). Free time as a reinforcer for assignment completion with high school special education students. *Education and Treatment of Children*, *2*(4), 271–277. http://www.jstor.org/stable/42900275

*Suter, C. L. (1993). *The effects of supplemental performance and on-task contingencies on the acquisition of math skills for elementary school students with behavioral disorders, students with attention deficit disorders, and students without disabilities* [Doctoral dissertation, University of North Texas]. University of North Texas Digital Archive.

https://digital.library.unt.edu/ark:/67531/metadc278442/

*Swain, J. C., & McLaughlin, T. F. (1998). The effects of bonus contingencies in a classwide token program on math accuracy with middle-school students with behavioral disorders. *Behavioral Interventions*, *13*(1), 11–19. https://doi.org/10.1002/(SICI)1099-078X(199802)13:1<11::AID-BIN6>3.0.CO;2-1

*Taffel, S. J., & O'Leary, K. D. (1976). Reinforcing Math with More Math: Choosing Special Academic Activities as a Reward for Academic Performance. *Journal of Educational Psychology*, *68*(5), 579.

*Terry, M. N., Deck, D., Huelecki, M. B., Huelacki, M. B., & Santogrossi, D. A. (1978). Increasing arithmetic output of a fourth-grade student. *Education and Treatment of Children*, *1*(2), 25–29. http://www.jstor.org/stable/42940485

Theodore, L. A., Bray, M. A., & Kehle, T. J. (2004). A comparative study of group contingencies and randomized reinforcers to reduce disruptive classroom behavior. *School Psychology Quarterly*, *19*(3), 253–271. https://doi.org/10.1521/scpq.19.3.253.40280

*Theodore, L. A., DioGuardi, R. J., Hughes, T. L., Aloiso, D., Carlo, M., & Eccles, D. (2009). A class-wide intervention for improving homework performance. *Journal of Educational & Psychological Consultation*, *19*(4), 275–299. https://doi.org/10.1080/10474410902888657

*Thomson, E. W., & Galloway, C. G. (1970). Material Reinforcement and Success in Spelling. *The Elementary School Journal, 70*(7), 395—398. https://doi.org/10.1086/460599

*Turco, T. L., & Elliott, S. N. (1990). Acceptability and effectiveness of group contingencies for improving spelling achievement. *Journal of School Psychology*, *28*(1), 27-37. https://doi.org/10.1016/0022-4405(90)90034-5

VanDerHeyden, A. M., & Witt, J. C. (2008). Best practices in Can't Do/Won't Do assessment.

Best Practices in School Psychology V, 2, 195-208.

*Velazquez, N. (2014). *The effects of the three jars game on the academic performance of a 5th grade inclusion class* [Master's project, State University of New York at Fredonia]. State University of New York Digital Archive. https://soar.suny.edu/handle/20.500.12648/470

Vidoni, C., & Ward, P. (2006). Effects of a dependent group-oriented contingency on middle school physical education students' fair play behaviors. *Journal of Behavioral Education*, *15*(2), 80–91. https://doi.org/10.1007/s10864-006-9012-z

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazalewski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis*, *26*, 9–21. https://doi.org/10.1901/jaba.1993.26-9

Warmbold-Brann, K., Burns, M. K., Preast, J. L., Taylor, C. N., & Aguilar, L. N. (2017). Meta-Analysis of the effects of academic interventions and modifications on student behavior outcomes. *School Psychology Quarterly*, *32*(3), 291–305. https://doi.org/10.1037/spq0000207

*Weekley, A. L. W. (1980). *Effects of group, individual, and no contingencies of reinforcement on the arithmetic performance of Navajo and Hopi students* [Doctoral dissertation, Ohio State University]. Ohio State University Digital Archive. http://rave.ohiolink.edu/etdc/view?acc_num=osu1487154578121428

*Wilder, V. K. (2011). *The efficacy of contingency models of reinforcement on group expectations and reading achievement* (Publication No. 3461676) [Doctoral dissertation, Walden University].

Williamson, B. D., Campbell-Whatley, G. D., & Lo, Y. (2009). Using a random dependent group contingency to increase on-task behaviors of high school students with high incidence disabilities. *Psychology in the Schools*, *46*(10), 1074–1083. https://doi.org/10.1002/pits.20445

*Winn, J. B. (2005). *The effects of known versus unknown criteria on journal writing performance of elementary students using an independent group contingency* [Doctoral dissertation, University of Tennessee, Knoxville]. University of Tennessee, Knoxville Digital Archive. https://trace.tennessee.edu/utk_graddiss/2313/

*Wodarski, J. S., Hamblin, R. L., Buckholdt, D. R., & Ferritor, D. E. (1973). Individual consequences versus different shared consequences contingent on the performance of low-achieving group members. *Journal of Applied Social Psychology*, *3*(3), 276–290. https://doi.org/10.1111/j.1559-1816.1973.tb02803.x

Wolfe, K., Pyle, D., Charlton, C. T., Sabey, C. V., Lund, E. M., & Ross, S. W. (2016). A systematic review of the empirical support for check-in check-out. *Journal of Positive Behavior Interventions*, *18*(2), 74-88. https://doi-org.libezproxy2.syr.edu/10.1177/1098300715595957

*Yarbrough, C., Charboneau, M., & Wapnick, J. (1977). Music as Reinforcement for Correct Math and Attending in Ability Assigned Math Classes. *Journal of Music Therapy*, *14*(2), 77–88. https://doi.org/10.1093/jmt/14.2.77

VITA

TAYLOR HITCHINGS

EDUCATION

Syracuse University (Syracuse, NY) 2019 – Present
School Psychology Doctoral Program
GPA: 4.0
Primary Advisor: Tanya Eckert, Ph.D.

University at Buffalo (Buffalo, NY)
Counseling and School Psychology Doctoral Program 2018 - 2019
GPA: 3.8
Primary Advisor: Bridget Hier, PhD

University of Colorado Boulder (Boulder, CO) 2013 - 2017
B.A. in Psychology and English—Literature Concentration
GPA: 3.85
Honors: Dean's List (2013-2017), Honors Program Member

RESEARCH EXPERIENCES

Syracuse University, TRAC and RITES (Syracuse, NY) 2019 - Current
Supervisor: Tanya Eckert, PhD (current); Bridget Hier, PhD (previous)

University at Buffalo, RITES (Buffalo, NY) 2018 - 2019
Supervisor: Bridget Hier, PhD

Colorado State University, Dept. of Psychology (Fort Collins, CO) 2017 - 2018
Supervisor: Bradley Conner, PhD
Post-Baccalaureate Volunteer

PUBLICATIONS

Hier, B.O., MacKenzie, C., Ash, T., Maguire, S., Nelson, K., Helminen, E., Watts, E. A., Matsuba, E., Masters, E., Finelli, C. C., Circe, J. J., **Hitchings, T.J.**, Goldstein, A., & Sullivan, W. (2022). *Effects of the Good Behavior Game on Students' Academic Engagement in Remote Classrooms During the COVID-19 Pandemic* [Accepted for publication]. Department of Psychology, Syracuse University.

Hier, B. O., **Hitchings, T. J.**, Ardoin, S. P., & Circe, J. J. (2022). *Effects of writing topic choice on elementary students' text generation* [In progress].

Hier, B. O., Truckenmiller, A. J., Cameron, C. E., Boyd, M. P., & **Hitchings, T. J.** (2022). *Guiding elementary students' planning during the writing process through verbal prompts: A randomized controlled trial* [In progress].

PRESENTATIONS

Eckert, T. L., Hier, B. O., Williams, N. L., Circe, J. J., **Hitchings, T. J.,** Watts, E. A., & Finelli, C. C., & Ardoin, S. P. (2021, February). *Association between writing productivity, writing apprehension, and grit* [Poster presentation]. National Association of School Psychologists, Virtual Convention.

Hier, B. O., Datchuk, S. M., Watts, E. A., Circe, J. J., **Hitchings, T. J.**, Goldstein, A., & Finelli, C. C. (2021, February). *Validity and classification accuracy of curriculum-based measurement of sentence construction* [Poster presentation]. National Association of School Psychologists, Virtual Convention.

Hitchings, T. J., Hier, B. O., Watts, E. A., & Finelli, C. C. (2021, February). *The relationship between second-grade students' off-task behavior and text generation* [Poster presentation]. National Association of School Psychologists, Virtual Convention.

Hier, B. O., **Hitchings, T. J.**, Ardoin, S. P., Goldstein, A., R., & Watts, E. A. (2020, February). *The effect of topic choice on second-grade students' writing performance* [Poster presentation]. National Association of School Psychologists Annual Convention, Baltimore, MD, United States.

GRANTS

Student Research Grant 2020
Association for Positive Behavior Support
"Reinforcement in the Classroom: A Meta-Analysis of Group Contingencies"

PROFESSIONAL EXPERIENCES

Colorado Center for Assessment and Counseling (Fort Collins, CO) 2021 - 2022
Psychometrist

Psychological Services Center at Syracuse University 2020 - 2021
Clinician

Trumpet Behavioral Health (Fort Collins, CO) 2017 - 2018
Behavioral Therapist

TEACHING EXPERIENCES

Syracuse University (Syracuse, NY)
Instructor of Record 2020

Syracuse University (Syracuse, NY)
Teaching Assistant 2019 - 2020