# A REVIEW OF THE LITERATURE ON RESPONSE CARDS AMONG STUDENTS WITH DISABILITIES

---

A Thesis
Submitted to
the Temple University Graduate Board

---

in Partial Fulfillment
of the Requirements for the Degree of
MASTER OF SCIENCE
IN APPLIED BEHAVIOR ANALYSIS

---

by
Laura E. Kuntzmann
August 2021

Thesis Approvals:
Matthew Tincani, Ph.D., BCBA-D, Professor, Special Education and Applied Behavior Analysis, Temple University

Examining Committee Members:
Saul Axelrod, Ph.D., BCBA-D, Retired Professor, Special Education and Applied Behavior Analysis, Temple University
Art Dowdy, Ph.D., BCBA-D, Professor, Special Education and Applied Behavior Analysis, Temple University
Matthew Tincani, Ph.D., BCBA-D, Professor, Special Education and Applied Behavior Analysis, Temple University
Jason Travers , Ph.D., BCBA-D, Professor, Special Education and Applied Behavior Analysis, Temple University

## ABSTRACT

Response cards are low-cost, low-tech instructional tools designed to engage learners while simultaneously aiding educators in evaluating student comprehension. The use of such a tool in educational settings has increased over the last several decades, as evidenced by the growing body of research evaluating their impact in the classroom. Though several researchers have reviewed the literature on response cards to date, the purpose of the present study was to conduct a more current and comprehensive systematic review of the literature focused on the use of response cards specifically among learners who have been diagnosed with one or more disabilities. A multi-step search procedure revealed 15 relevant studies which met inclusion criteria. A total of 141 participants with and without disabilities ranging in age from five to 19 were included in the present study. Outcomes of interest included active responding, correct responding, on-task behavior, inappropriate behavior, and academic achievement. Results of the review extend the findings of Randolph (2007), Horn (2010), and Schnorr et. al. (2016), indicating the continued effectiveness of response cards in yielding desirable outcomes, both behaviorally and academically. Increases in nearly all dependent variables including active responding, correct responding, on-task behavior, and academic achievement were reported. However, outcomes related to decreasing inappropriate behavior were variable. There exist several limitations within the present study, and recommendations for future research are plentiful. Those recommendations as well as implications for use are discussed.

*Keywords:* response cards, disability, learning

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER 1

# A REVIEW OF THE LITERATURE

The use of response cards as an academic and behavioral intervention for students with (and without) disabilities has become common practice in educational settings (Horn, 2010; Randolph, 2007). As compared to more traditional means of active participation in the classroom (e.g. hand raising, cold calling), the use of response cards allows for a more effective and efficient learning experience for both students and teachers (Tincani & Twyman, 2016). While hand raising is a cost-efficient and universally understood strategy for student participation, it allows for the active participation and evaluation of just one student in the classroom at a time. On the contrary, the use of response cards engages all learners in the learning process while simultaneously offering teachers real-time, formative feedback on student understanding at the individual level (Horn, 2010).

## Active Student Responding

Active student responding (ASR) entails engaging students in the learning process through active and observable participation. Heward (1994) defined ASR more technically as "an observable response made to an instructional antecedent" (p. 290). In his discussion of ASR, Heward acknowledged that learning can, and certainly does, occur privately (i.e., without an observer being able to detect the process), and that the process of learning privately and independently has its place in the classroom. However, for an educator to truly assess comprehension of a skill or concept among students, responding must, at some point, become public and observable (i.e. saying or emitting a response that reflects student understanding). The first step to achieve higher rates of active (and

accurate) student responding is to increase the actual number of opportunities that students have to respond to teacher-directed questions or cues. According to Hattie (2012) and Rosenshine (2012), it is likely that as opportunities to respond (OTR) increase, so, too, will ASR.

Examples of ASR vary depending on setting, subject matter, and student profiles. Most examples of ASR include some form of written or spoken responses, though any observable demonstration of understanding following an instructional cue would qualify. In a Science or English Language Arts classroom, for example, ASR might present itself quite traditionally (i.e. students raise their hands and respond verbally to a teacher-posed question). In a Family and Consumer Sciences class, however, ASR might take on a more hands-on approach and include, for example, a room full of students identifying and using a specific cooking utensil or sewing machine. Just as the needs of learners evolve, so, too, must the means by which educators seek engagement.

## Response Cards

As inclusion of students with disabilities into general education classes increases (Schnorr et. al., 2016), so, too, must the development of strategies to engage learners in those settings irrespective of their ability (Ault & Horn, 2018; Nagro et. al., 2016). According to Nagro et. al. (2016), teachers have at their disposal a host of low and high-tech student response systems to engage students in the learning process and increase ASR, one of which is response cards. Though the advantages of response cards are far reaching for learners of varying demographics and abilities, (Berrong et. al., 2007; Helf, 2015; Heward, 1997), much of the research on response cards has focused on students with disabilities.

Heward (1997) defined response cards as “cards, signs or items that are used by students to indicate their response to a question or problem presented by the teacher” (p. 43). Among their many advantages, these tools are cost efficient and low-tech, making them accessible no matter a school’s socio-economic status (Herring & Woolsey, 2020; Horn, 2010; Randolph 2007; Schnorr et.al., 2016). Equally convenient, the form of these response cards can vary greatly from one classroom to the next and can be easily manipulated to serve the specific needs of the learners in an academic setting (Randolph, 2007). The simplest form of response card is a write-on card on which students are expected to construct their own response. These are used when answer choices are not provided by the teacher and may include dry erase boards, index cards, or electronic devices such as tablet computers. More advanced response cards, however, can be used when response choices are limited to those provided by the teacher. These can be the same basic, blank cards that were previously mentioned, or they can be pre-printed to include pre-determined answer choices or universal answer options (e.g., true/false, multiple choice letterings). No matter their form, response cards all serve the same function as a means to increase opportunities to respond, active student responding, on-task behavior, and academic achievement.

Among the first to evaluate the effects of response cards specifically among students with disabilities was Davis et. al. (2004). Davis et. al. investigated whether response cards affected academic and off-task responding for four participants in a self-contained English classroom. Of the four participants, two were seventh-grade girls who received special education services for identified learning disabilities as well as English as a Second Language (ESL) services. The remaining two participants were in eighth

grade, one of whom was a boy with a learning disability. The other 8th grade participant was a girl with a traumatic brain injury. Results indicated that all students demonstrated increased academic responding during response card conditions when compared to baseline conditions. These findings mirrored the findings of previous preliminary research on response cards. Interestingly, however, the effects of the intervention on off-task or disruptive behaviors was inconclusive. In addition, social validity data of the intervention contradicted that of previous studies, indicating a preference for hand-raising as a means of responding. However, these results were attributed to the difficulties of several participants, specifically those who received ESL services, to communicate their responses given the language differences. For these reasons, further research among students receiving ESL services was suggested.

Additional studies were later conducted on the effects of response cards among students with disabilities, including Horn et al. (2006) and Berrong et. al. (2007). Both studies evaluated the effects of response cards on the behavior of students with disabilities; the dependent variables in both studies were percent of active responding, percent of on-task behavior, and frequency of disruptive behaviors. The intention of both studies was to determine whether response cards increased desirable behaviors while simultaneously decreasing undesirable behaviors. Horn et. al. (2006) included three middle school-aged participants within their study, all of whom were ages 12- to 15-years-old and received instruction in a self-contained classroom. Similarly, Berrong et. al. (2007) included eight students between the ages of 10- and 12-years-old who were also instructed in a self-contained setting. In both studies, all students were diagnosed as having moderate or severe disabilities. Results of these studies extended previous

findings relating to student responding, affirming the effectiveness of response cards as an intervention for increasing participation in the classroom. Additionally (and contradicting to previous research), both studies highlighted a simultaneous increase in on-task behavior and decrease in disruptive behaviors following response card conditions. Results indicate the use of response cards during instructional time could have a considerable impact not only on active responding but on other student behaviors as well.

Randolph (2007) analyzed the body of response card research in a meta-analysis focused specifically on student engagement, academic achievement, and on/off-task behavior. This meta-analysis sought to extend the findings of his own previous work conducted in 2005 on the same topic. Included in this analysis were 18 studies, 11 of which had been published first as dissertations or master's theses. Study characteristics within this analysis included type of treatment, type and place of publication, academic area, grade level, study design, level of analysis, ceiling effects, and interrater reliability/treatment fidelity. Results were indicative of the intervention's efficacy. When compared to traditional hand-raising, response cards outperformed on all of the dependent variables included in the review, regardless of the model used. Additionally, when the effects of different types of response cards (write-on vs pre-printed) were evaluated, there was no significant difference between types of response cards in terms of their impact on student engagement and academic achievement. This conclusion illustrates educators' flexibility and autonomy when using response cards as an intervention. Also, the majority of students who participated in the included studies preferred using response cards as opposed to engaging in hand-raising behavior, which is

certainly worth consideration among educational practitioners. Despite the indication that response cards are likely to influence many areas of student behavior and performance, Randolph failed to evaluate the quality of studies included in his review and therefore conclusions regarding whether this intervention can be considered evidence-based are limited.

More recently, Horn (2010) conducted a review on the published, empirical literature focused on using response cards as a means of active responding for students identified with a disability. Also analyzed within the review was evidence which might classify response cards as an evidence-based practice. There were six studies included in the review, all of which used response cards as a means of engaging students in active responding and were coded based on participant description, study design, and results. Horn concluded that in each of the studies, students exhibited an increased rate of accurate responding. In addition, half of the studies reported an increase in on-task behavior and a decrease in the occurrence of inappropriate behavior. Based on preliminary criteria outlined in Horner et. al. (2005), Horn concluded that response cards could be considered an evidenced-based practice in educational settings. These preliminary guidelines included a minimum of five studies demonstrating experimental control, investigations conducted by a variety of researchers across multiple settings, and investigations conducted with a minimum of 20 total participants. However, a quality assessment of all studies was again omitted from this review, indicating a need for an additional systematic review that accounts for the quality of reviewed studies in determining whether response cards are an evidence-based intervention.

Most recently, Schnorr et. al. (2016) published a review of the literature that focused specifically on response cards as a strategy for increasing opportunities to respond, and, much like Horn (2010), also evaluated the extent to which the intervention could be classified evidence-based. Schnorr et. al. (2016) identified six studies which met their inclusion criteria. Participants within all six studies included elementary-aged students with and without disabilities, and the studies were coded based on purpose, participants, setting, design, dependent variable(s), type of response card used, and results. All studies were also assessed according to a 20-item single-case quality indicator checklist outlined by the National Secondary Transition Technical Assistance Center (NSTTAC). This checklist was developed based on the criteria set forth by Horner et. al. (2005) and included a dichotomous rating scale (allowing "Yes" or "No" responses). This clarification regarding the system in which quality indicators were rated stems from the work of Cook et. al. (2009) who concluded that rating systems involving more than two possible values resulted in lower rates of interrater reliability. The purpose of this quality indicator assessment was to determine whether the use of response card research justified classifying the intervention as an evidence-based practice.

Of the six studies included in Schnorr et.al. (2016), two were determined to be of high quality (i.e. met all 20 quality indicators) and the remaining four were determined to be of acceptable quality. The results of this review indicate that the use of response cards at the elementary level *did* lead to increased opportunities to respond, and based on the quality indicator checklist results, could be classified an evidence-based practice with a moderate level of evidence. However, the authors of the review acknowledged several limitations of their work, including that only studies conducted at the elementary level

were analyzed and only one dependent variable, OTR, was reviewed. Schnorr et. al. suggested future researchers investigate the level of evidence for using response cards to increase (or decrease) dependent variables other than OTR, as well as research focused on determining the effectiveness of the intervention on more specific populations (e.g. individuals with disabilities, English Language Learners, etc) and in different settings (e.g. inclusive settings, self-contained settings).

Several themes emerge from the existing literature on response cards, one of which is the multitude of functions they serve within the classroom. As was previously mentioned, a major function of response cards in the academic setting is to increase opportunities to respond (OTR), but even more than just providing opportunities to respond, the use of response cards has also been shown to increase active student responding (Schnorr et. al. 2016). And while active responding does not necessarily ensure accurate responding, studies have also yielded results indicating an increase in academic achievement among students following the implementation of active response strategies like response cards (Rosenshine 2012; Owiny et. al., 2018). A final potential benefit of response cards is their ability to increase on-task behavior during instructional time, making the intervention effective not only for academic purposes but for behavior management purposes as well.

Equally notable was the discovery that response cards were found to be effective among participants with varying disabilities. From diagnoses of Intellectual and/or Learning Disabilities to Hearing Impairments and even to Traumatic Brain Injuries, studies have included participants with a multitude of disabilities or deficits. The overwhelming majority of participants in those studies seem to have benefited from the

use a response cards, indicating the intervention's broad applicability. In addition, individuals who engage in disruptive and off-task behaviors, whether identified as having a disability or not, have also been reported to benefit greatly from the use of response cards as an intervention (Heward 1994), highlighting the intervention's potential to increase desirable behaviors and also reduce those that are undesirable simultaneously.

Finally, response cards have been effective to increase desired behaviors across levels (elementary school, middle school, high school) and settings (inclusive, self-contained) (Randolph, 2007; Horn; 2010). Not just an engaging tool for a specific age group or demographic, this intervention has shown that it can be adapted and implemented across virtually any environment and still result in increased responding, advancements in academic achievement, and greater instances of on-task behavior.

However, despite the emergence of these themes in the existing literature on response cards and their reported success in previous analyses, no recent systematic review has addressed these variables, making it necessary to re-evaluate the efficacy of such an intervention.

## Research Questions

Systematic reviews of response cards as an academic and behavioral intervention highlight their effectiveness in serving a variety of functions, across a wide range of ability levels, and within a multitude of settings. However, no review to date has evaluated all these factors together, nor has any review analyzed any one of these variables within the last five years. Additionally, existing reviews all highlight limitations to their research and the scope of studies included within their work. Most recently, Schnorr et. al. (2016) highlighted the need for further analysis of research among middle

and high school-aged participants and on dependent variables other than OTR in order to establish response cards as an evidence-based practice among broader populations and in serving a greater number of functions. The purpose of the current review is to extend the findings of Randolph (2007), Horn (2010), and Schnorr et. al. (2016) and provide focused directions for future research. Specifically, this systematic review will: (a) describe the extent to which response cards have been reported to influence student behaviors, both desirable and undesirable (i.e. active responding, correct responding, on-task behavior, inappropriate behavior); (b) describe the extent to which response cards have been reported to increase academic achievement among students; (c) determine the extent to which response card research incorporates features of high-quality single-case research; and (d) discuss opportunities for future research.

# CHAPTER 2

# METHOD

This systematic review identified studies that evaluated the use of response cards among students with (and without) disabilities in the school setting. The same extraction procedures were used for this review as those described in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et.al., 2009). Following extraction, the resulting articles were coded according to the article coding categories described in the following sections.

## Search Strategy

Article extraction began in September of 2020 and concluded in November of the same year. Google Scholar and APA PsycInfo were the databases used throughout the search process. Seven individual searches were conducted by the author of this review within each database, all of which included the phrase “response cards” along with one of seven terms intended to limit hits to those including some type of disability category. The following terms, all with Boolean operators and truncation, were paired with the aforementioned search term: disab*, autism, deaf, disturb*, impair*, injury, ADHD. These seven search terms were selected so as to ensure the greatest chance of capturing all studies which evaluated response cards among individuals with any of the thirteen IDEA disability categories (e.g. the term *impair* would likely yield results which included participants with hearing, orthopedic, other health, speech/language, and/or visual impairments). This preliminary search yielded 7,683 records, though only 2,250 remained following the removal of all duplicates.

## Selection of Studies

Following the preliminary search and the removal of all duplicate articles, the remaining 2,250 articles went through a preliminary screening based solely on title and abstract. The title and/or abstract of each article must have referenced the intervention and must also have mentioned individuals with disabilities as being part of the selected group of participants within the study. Following exclusion of articles when it was clear by their abstracts that they did not meet the established criteria, the remaining studies were each read and screened using pre-determined inclusion criteria. In order to be included in this review, studies must have specified the active manipulation of response cards as the sole intervention; however, no specifications were made about the dependent variable upon which the response cards were intended to have an effect. Only single-subject studies were included in the systematic review and all studies must have had at least one baseline and one intervention phase. All studies were required to have included at least one participant with a disability, making studies involving both students with and without disabilities acceptable for inclusion. The year of publication was not considered at any point throughout the search or evaluation process. It should also be noted that while theses and dissertations were included within this review, any instances in which a thesis or dissertation re-appeared as a published article resulted in the inclusion only of the published study. A second rater also participated in the screening process, reading 100% of the articles with 100% agreement.

## Study Coding

In addition to citation, the extracted articles were coded by the author of this review according to the categories below, and in cases where insufficient information

was provided by authors of studies, that category was labeled as "___ unknown." This coding is shown in Table 1.

### *Participants*

The number of participants as well as their races, genders, ages, and disability classifications were coded for each study included within the review. The possible values for race were Caucasian, African American, Asian or Asian American, and Hispanic. The possible values for genders were male or female. Disability classifications were abbreviated within Table 1 with a Note to define each abbreviation. No information was assumed, so in cases where insufficient information was provided by the authors, that category was labeled as "___ unknown."

### *Setting*

The settings in which each study took place were coded according to type of learning environment, geographical region, level of the school in which the study was conducted, and type of instruction that was implemented during data collection. The possible values for type of learning environment were inclusive and self-contained, and the possible values for geographical region were urban, suburban, and rural. Additionally, the possible values for level of the school were elementary school, middle school, and high school. Specification about school classification (e.g., public vs special school) were not included in the coding process. Again, no information was assumed (i.e., grade levels were not translated into common age ranges, etc.), so in cases where insufficient information was provided by the authors, that category was labeled as "___ unknown."

## *Design*

The experimental design was also coded for each study included within the review. The majority of studies followed an ABAB design; however, designs were coded directly as they were reported within each article to acknowledge all variations and alternative designs.

## *Dependent Variables*

The dependent variable(s) upon which the intervention was intended to have an effect were coded for each study. The possible values were Active Responding, Correct Responding, On-Task Behavior, Inappropriate Behavior, and Academic Achievement.

### *Active Responding*

For the purposes of this review, active responding was defined as any instance in which participants completed the assigned response task using the appropriate method (i.e., raising their hand during hand-raising conditions, manipulating response cards during response card conditions, etc.). While referred to in some studies by a different name (e.g., participation, active student responding (ASR), etc.), any variable described in similar terms was coded as having measured active responding as a dependent variable. Response tasks were not specified when selecting studies for review, nor were response tasks coded as they were not considered relevant to this study. Though many studies noted the amount of time within which participants were required to complete the assigned response task, that factor was also not considered relevant enough to have been coded for this review.

### *Correct Responding*

For the purposes of this review, correct responding was defined nearly identically to active responding, with the additional criteria of an accurate response on the part of the participant. Participant responses were not recorded if they were not accurate, nor were they recorded if the participant did not use the appropriate method to complete the response task (i.e., raising their hand during hand-raising conditions, manipulating response cards during response card conditions).

### *On-Task Behavior*

For the purposes of this review, on task behavior was defined as the active engagement of participants in the teacher's instructions, questions, or explanation of response task as well as physical attentiveness to whomever was speaking to the class (e.g., teacher, classmate, etc.). Though referred to in some studies by a different name (e.g., active engagement, academic engagement, etc.), any variable described in similar terms was coded as having measured on-task behavior as a dependent variable.

### *Inappropriate Behavior*

For the purposes of this review, inappropriate behavior was defined as any instance in which a participant was out of seat without permission, talking or yelling without being called upon by the instructor, placing their hands on other students within the class, or engaging in any other activities that did not directly relate to the given response task. Though referred to in some studies by a different name (e.g., off-task behavior), any variable described in similar terms was coded as having measured inappropriate behavior as a dependent variable.

### *Academic Achievement*

For the purposes of this review, academic achievement was defined as the number or percentage of correct responses out of a total number of possible responses during any sort of assessment or probe. Information regarding the types of questions within each assessment was not coded as it was not considered relevant to this review. Again, though referred to in some studies by a different name (e.g., skill acquisition), any variable described in similar terms was coded as having measured academic achievement as a dependent variable.

It should be noted that one study analyzed the effects of response cards on an additional dependent variable not noted above. This variable was that of opportunities to respond, which differed slightly in definition from active responding. Opportunities to respond was defined in the related study as any instance in which a participant completed the assigned response task using the appropriate method *and* was called on by the teacher to share their response. Given that only one study made this distinction between such similar dependent variables and that no discrepancies were observed between the two variables, it was decided that coding of that category was unnecessary.

## *Independent Variable*

The form of the response cards within each study was also coded by any one or combination of three possible values: pre-printed, write-on, and/or manipulative. All three possible values were either reported by the authors of each study or were deduced by the author of this review based on descriptions of student responding.

### *Pre-Printed*

For the purposes of this review, pre-printed response cards were defined as any response card device on which possible answer choices had been pre-populated by the instructor. Pre-populated answer choices included, but were not limited to, multiple choice options, "True/False" statements, image choices, etc. Acceptable student use of pre-printed response cards included receptive identification of the correct response within a card or of the correct card within an assortment of cards.

### *Write-On*

For the purposes of this review, write-on response cards were defined as any response card device on which written responses are required. These write-on response cards included, but were not limited to, dry-erase boards, laminated papers, index cards, electronic devices with some sort of drawing capability, etc. Acceptable student use of write-on response cards included expressive identification of the correct response via filling in a blank or fully writing the desired response.

### *Manipulative*

For the purposes of this review, manipulative response cards were defined as any tangible object or device with which responses can be displayed or demonstrated. Manipulative response cards included, but were not limited to, flip cards (i.e., an object consisting of flaps that could be rotated to indicate a letter or number combination), etc.

## *Results*

Mean data values were coded for each study included within this review. Specifically, mean data values among all participants (as opposed to mean data values for each individual participant) were coded. Mean data values were extracted directly from

each study. In cases when mean data values for the entire group of participants were not provided but individual mean data values for each participant were reported, the author of the review calculated mean data values for the group. The purpose of this calculation was so that comparisons of like-data could be made across studies. In cases where neither group nor individual mean data values were reported, narrative results were extracted and summarized from within the study.

### *Quality of Study*

Horner et. al. (2005) offer a series of 21 quality indicators for quantifying the rigor and credibility of single-case research. Those indicators are divided into seven categories: descriptions of participants and settings (three indicators), dependent variable measurement (five indicators), independent variable measurement (three indicators), baseline (two indicators), experimental control (three indicators), external validity (one indicator), and social validity (five indicators). However, only 20 of the prescribed 21 indicators were deemed relevant to this review. The remaining indicator – *social validity is enhanced by implementation of the IV over extended time period, by typical intervention agents in typical physical and social contexts* – was not scored. This indicator was excluded from the checklist because it was determined that the remaining indicators within the final category of the checklist were more likely to be reported and would adequately represent the social validity of the intervention.

Following initial coding, each study was scored on these 20 quality indicators according to a dichotomous rating scale (allowing only responses of "Yes" or "No"). This rating scale was selected over alternatives (e.g., a four-value rating scales including "Partial" or "N/A" responses) in order to yield the most accurate and agreeable results.

According to Cook et. al. (2009), rating systems involving more than two possible values have historically resulted in lower rates of interrater reliability. For every instance in which a study was rated as meeting a quality indicator (i.e., being ranked “Yes”), one point was awarded. Conversely, studies were awarded zero points for being ranked “No” within a quality indicator.

Following this determination of scoring guidelines and the completion of all ratings, a percent of quality indicators met was calculated for each study. Points awarded to each study were added together, divided by the total items scored, and multiplied by 100 (e.g., if a study was awarded 14 points out of 20 applicable indicators, the percent would be calculated as 14/20 * 100 = 70%).

## Interrater Agreement

To assess interrater agreement of the article coding, the author and a second rater separately coded three of the 15 articles (20%). All three selected articles were chosen at random in order to ensure reliability. Following individual coding, results were compared and any disagreements were resolved until interobserver agreement was 100% for each coding category.

To assess interobserver agreement of the quality indicators checklist, the author and second rater separately scored three of the 15 articles (20%). Again, all three articles were selected randomly to ensure reliability. Within the rating process, both raters independently awarded points to each study according to the procedures outlined previously. Following individual ratings, results were compared and interobserver agreement was calculated by dividing the agreements of each indicator by the number of

agreements plus disagreements. The resulting value was then multiplied by 100 and converted it to a percentage. Interobserver agreement was 100%.

# CHAPTER 3

## RESULTS

Results of the PRISMA search and article extraction are shown in Figure 1. Of 7,683 articles identified through initial database searches, 5,433 articles were found to have been duplicates, leaving 2,250 articles for further screening. Another 2,203 articles were then removed because titles and/or abstracts indicated that they did not meet the inclusion criteria. Following a full-text screening of each of the remaining 37 articles, 22 were subsequently screened out. The resulting 15 articles were coded for the systematic review.

Although no criteria were established regarding year of publication, the earliest study included within this review was published in 2004. Of the 15 articles, only one article was published between 2000 and 2005, four were published between 2006 and 2010, three were published between 2011 and 2015, and the remaining seven were published between 2016-2021.

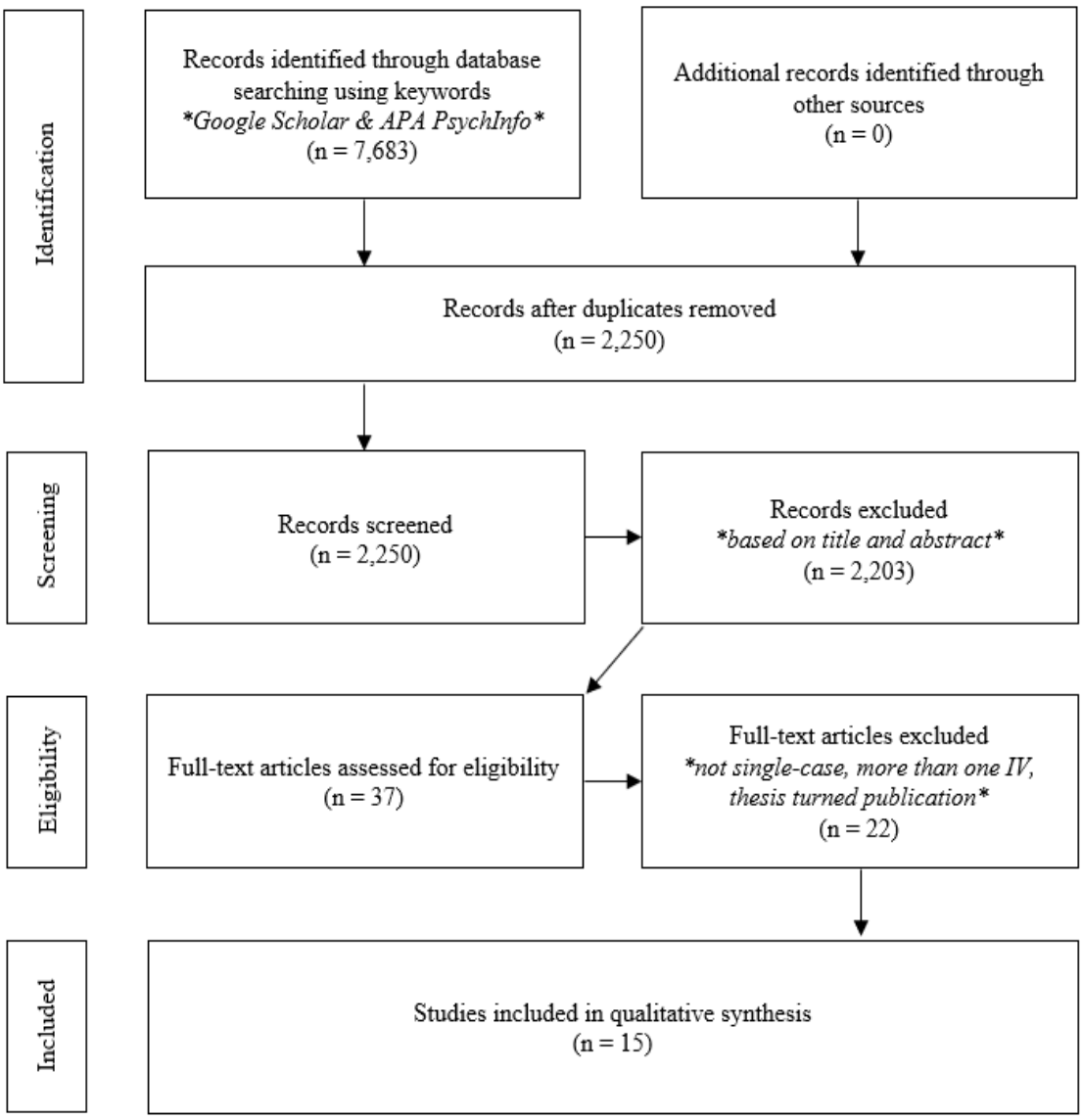


**Figure 1.** PRISMA Flow Diagram

**Table 1**

*Overview of Response Card Studies*

| Reference | Participants | Setting | Design | Dependent Variable(s) | Independent Variable | Results | Quality of Study |
|---|---|---|---|---|---|---|---|
| Berrong et. al. (2007) | Eight students, all with moderate to severe disabilities; one physical disability, one hearing impairment; three female, five male; races unknown; ages 10-12 | Self-contained special education classroom; elementary school; during "Calendar" instruction | ABAB design | Active Responding, On-Task Behavior, Inappropriate Behavior | Pre-printed RC | Active Responding: M%: HR1 (M = 21.7%), RC1 (M = 58.8%), HR2 (M = 28.7%), RC2 (M = 56.3%). On-Task Behavior: M%: HR1 (M = 35.7%), RC1 (M = 79.4%), HR2 (M = 36.9%), RC2 (M = 71.5%). Inappropriate Behavior: M rate: HR1 (M = 0.77), RC1 (M = 0.40), HR2 (M = 0.89), RC2 (M = 0.21). | 95% |
| Boden (2015) | Five students, all with MoID; one African American male, age 11; one African American male, age 15; one Caucasian female, age 15; two African American females, age 13 | Self-contained special education classrooms (2); urban middle school; during Math instruction | Yoked Multiple-Baseline Across Dyads design (with an embedded reversal) | Active Responding, On-Task Behavior, Academic Achievement | Pre-printed RC | Active Responding: M%: HR1 (M = 77.9%), RC1 (M = 97.7%), HR2 (M = 87.3%), RC2 (M = 99.4%). On-Task Behavior: M%: HR1 (M = 43.1%), RC1 (M = 68.0%), HR2 (M = 53.5%), RC2 (M = 71.0%). Academic Achievement: M%: HR1 (M = 22.5%), RC1 (M = 42.3%), HR2 (M = 61.5%), RC2 (M = 75.2%). | 85% |
| Bondy & Tincani (2018) | Three students; one Caucasian male with ASD, age 7; one Asian American male with Down Syndrome/OHI, age 7; one Black male with ASD/ADHD, age 7 | Self-contained special education classroom; suburban elementary school; during a variety of instruction | ABAB design | Active Responding, Correct Responding | Pre-printed RC | Active Responding: M%: HR1 (M = 32.6%), RC1 (M = 90.9%), HR2 (M = 33.2%), RC2 (M = 84.7%). Correct Responding: M%: HR1 (M = 6.4%), RC1 (M = 52.7%), HR2 (M = 5.0%), RC2 (M = 56.2%). | 95% |

**Table 1**
*(continued)*

| Reference | Participants | Setting | Design | Dependent Variable(s) | Independent Variable | Results | Quality of Study |
|---|---|---|---|---|---|---|---|
| Cakiroglu (2014) | Four students, all with MID; one Caucasian female, age 11; one Caucasian male, age 9; one Caucasian male, age 12; one Caucasian female with a physical disability, age 9 | Self-contained special education classroom; foreign (Turkey) elementary school; during Soc Studies instruction | ABAB design | Active Responding, Correct Responding | Pre-printed RC / Write-on | Active Responding: M%: HR1 (M = 27.5%), RC1 (M = 91.5%), HR2 (M = 22.6%), RC2 (M = 92.7%). Correct Responding: M%: HR1-2 (M = 15.5%), RC1-2 (M = 76.5%). | 100% |
| Clarke (2010) | Seven students, all with MID/SLI; two Caucasian females with ADHD, age 9; one Caucasian male, age 8; one Caucasian male with physical disability, age 8; one Caucasian male with ADHD, age 9; one Caucasian female with hearing impairment, age 11; one Caucasian male, age 11 | Inclusive general education classroom; rural elementary school; during Science and/or Social Studies instruction | ABAB design | Active Responding, Correct Responding, On-Task Behavior, Academic Achievement | Pre-printed RC | Active Responding: M%: HR1 (M = 9.9%), RC1 (M = 100.0%), HR2 (M = 3.9%), RC2 (M = 100.0%). Correct Responding: M%: HR1 (M = 1.9%), RC1 (M = 75.0%), HR2 (M = 0.7%), RC2 (M = 89.5%). On-Task Behavior: M%: HR1 (M = 90.4%), RC1 (M = 100.0%), HR2 (M = 92.9%), RC2 (M = 100.0%). Academic Achievement: M%: HR Multiple Choice (M = 38.6%), RC Multiple Choice (M = 55.7%), HR Open Ended (M = 42.9%), RC Open Ended (M = 53.8%). | 95% |
| Clarke et. al. (2016) | Five students, all with ID/SLI; two female, three male; races unknown; ages 8-9 | Inclusive general education classroom; rural elementary school; during Science and Social Studies instruction | ABAB design | Active Responding, On-Task Behavior | Pre-printed RC | Active Responding: M%: HR1 (M = 9.76%), RC1 (M = 100.0%), HR2 (M = 3.93%), RC2 (M = 100.0%). On-Task Behavior: M%: HR1 (M = 92.2%), RC1 (M = 100.0%), HR2 (M = 96.0%), RC2 (M = 100%). | 95% |

**Table 1**
*(continued)*

| Reference | Participants | Setting | Design | Dependent Variable(s) | Independent Variable | Results | Quality of Study |
|---|---|---|---|---|---|---|---|
| Davis & O'Neill (2004) | Four students; two females with SLD/receiving ESL services, races and ages unknown; one male with SLD, race and age unknown; one female with TBI, race and age unknown | Self-contained special education classroom; middle school; during English instruction | ABAB design | Active Responding, Correct Responding, Inappropriate Behavior | Write-on RC | Active Responding: M% not reported; however, RC resulted in increase in active responding for all students. Correct Responding: M%: HR1-2 (M = 74.0%), RC1-2 (M = 91.0%). Inappropriate Behavior: M% not reported; however, RC influenced off-task behavior for one student. | 75% |
| Davis (2020) | Nine students; seven Caucasian, one African American, one Asian; one male with MID, age 16; one female with ASD/ADD/behavior needs, age 16; one male with MID/physical disability, age 16; one male with ASD/ADD/SLI/behavior needs, age 17; one male with OHI/MID/ADD, age 16; one female with MID/SLI, age 16; one male with MoID/Down Syndrome, age 20; one female with MID/receiving ESL services, age 19; one female with MID/Down Syndrome, age 19 | Self-contained special education classroom; rural high school; during English instruction | ABAB design | Active Responding, On-Task Behavior, Inappropriate Behavior, Academic Achievement | Write-on RC | Active Responding: M%: HR1 (M = 19.2%), RC1 (M = 82.0%), HR2 (M = 17.6%), RC2 (M = 99.4%). On-Task Behavior: M%: HR1 (M = 50.4%), RC1 (M = 58.8%), HR2 (M = 47.8%), RC2 (M = 84.2%). Inappropriate Behavior: M%: HR1 (M = 49.6%), RC1 (M = 39.6%), HR2 (M = 52.2%), RC2 (M = 15.8%). Academic Achievement: M%: HR Pre-Test (M = 36.1%), RC Pre-Test (M = 27.9%), HR Post-Test (M = 43.8%), RC Post-Test (M = 54.7%). | 95% |
| Didion et. al. (2020) | Five students, all with EBD; all male; two Caucasian, three Black; ages 12-14 | Self-contained special education classroom; urban middle school; during Math instruction | ABAB design | On-Task Behavior, Active Responding | Write-on RC | On-Task Behavior: M%: HR1 (M = 59.7%), RC1 (M = 79.6%), HR2 (M = 46.4%), RC2 (M = 84.1%). Active Responding: M%: HR1 (M = 26.13%), RC1 (M = 82.09%), HR2 (M = 35.98%), RC2 (M = 82.6%). | 100% |

**Table 1**
*(continued)*

| Reference | Participants | Setting | Design | Dependent Variable(s) | Independent Variable | Results | Quality of Study |
|---|---|---|---|---|---|---|---|
| Duchaine et. al. (2018) | Six students; one male with EBD, age 18; one female with SLD, age 19; one male with ASD, age 16; one male with no disability, age 17; one female with no disability, age 17; one male with no disability, age 16; races unknown. | Inclusive general education classroom; suburban high school; during a variety of instruction | Alternating Treatment design | Active Responding, Academic Achievement | Write-on RC | Active Responding: M%: Baseline (M = 7.9%), HR (M = 8.95%), RC (M = 52.8%). Academic Achievement: M%: Baseline Probe (M = 56.0%), Alternating Treatment Probe (M = 64.6%). | 90% |
| Dudley (2013) | Five students; disabilities and genders unknown; all Caucasian; ages 11-12 | Inclusive general education classroom; rural elementary school; during Math instruction | Alternating Treatment design | Academic Achievement | Write-on RC | Academic Achievement: M%: HR (M = 58.0 %), RC (M = 84.0%) | 80% |
| Horn et. al. (2006) | Three students, all with moderate to severe disabilities; one female, age 12; one male, age 15; one male, age 12; races unknown | Self-contained special education classroom; rural middle school; during "Time Telling" instruction | ABAB design | Active Responding, Correct Responding, On-Task Behavior, Inappropriate Behavior | Manipulative RC | Active Responding: M%: HR1 (M = 54.0%), RC1 (M = 100.0%), HR2 (M = 64.5%), RC2 (M = 100.0%). Correct Responding: M%: HR1 (M = 60.0%), RC1 (M = 90.0%), HR2 (M = 66.6%), RC2 (M = 90.0%). On-Task Behavior: M%: HR1 (M = 69.6%), RC1 (M = 97.6%), HR2 (M = 88.08%), RC2 (M = 100.0%). Inappropriate Behavior: M rate of behaviors per minute: HR1 (M = 0.96), RC1 (M = 0.24), HR2 (M = 0.82), RC2 (M = 0.19). | 95% |

**Table 1**
*(continued)*

| Reference | Participants | Setting | Design | Dependent Variable(s) | Independent Variable | Results | Quality of Study |
|---|---|---|---|---|---|---|---|
| McCargo (2017) | Sixty-eight students; three with SLD, two with OHI, two with EBD, two with ASD, one with MDS, fifty-eight with no disability; 29 female, 39 male; 44 Caucasian, 23 African American, one Hispanic; ages unknown | Inclusive general education classrooms (3); high school; during a variety of instruction | ABCBC design | On-Task Behavior, Inappropriate Behavior, Off-Task Behavior | Pre-printed RC | On-Task Behavior: M%: Baseline (M = 55.7%), Increased OTR1 (M = 59.7%), RC1 (M = 62.7%), Increased OTR2 (M = 58.7%), RC2 (M = 48.7%). Inappropriate Behavior: M%: Baseline (M = 27.0%), Increased OTR1 (M = 25.7%), RC1 (M = 27.7%), Increased OTR2 (M = 29.3%), RC2 (M =38.0%). Off-Task Behavior: M%: Baseline (M = 16.0%), Increased OTR1 (M = 13.0%), RC1 (M = 8.0%), Increased OTR2 (M = 11.0%), RC2 (M = 11.7%). | 90% |
| Müllerke et. al. (2019) | Five students, all with SLD; one Mongolian male, age 12; one Serbian male, age 15; one Russian male, age 13; one German female, age 13; one German female, age 14 | Self-contained special education school; foreign (Germany) rural school; during Math instruction | Multiple-Baseline (ABA) Across Participants design | Active Responding, Academic Achievement | Write-on RC | Active Responding: M% out of 15: HR1 (M = 13.93%), RC1 (M = 97.73%), HR2 (M = 25.8%). Academic Achievement: M%: HR1 (M = 84.8%), RC1 (M = 96.4%), HR2 (M = 75.2%) | 90% |
| Wood et. al. (2009) | Four students; one Multiracial female with no disability, 6 years old; one Caucasian male with no disability, 5 years old; one Caucasian male with SLD/SLI, age 6; one Caucasian female with DD, 6 years old. | Inclusive general education classroom; rural elementary school; during "Calendar" instruction | ABAB design | Active Responding, Inappropriate Behavior | Pre-printed RC | Active Responding: M%: HR1 (M = 5.67%), RC1 (M = 97.87%), HR2 (M = 9.77%), RC2 (M = 94.5%). Inappropriate Behavior: M%: HR1 (M = 70.36%), RC1 (M = 1.84%), HR2 (M = 53.6%), RC2 (M = 6.74%). | 100% |

*Note:* EBD = Emotional & Behavioral Disorder; ADHD = Attention Deficit Hyperactivity Disorder; ASD = Autism Spectrum Disorder; OHI = Other Health Impairment; MoID = Moderate Intellectual Disability; MID = Mild Intellectual Disability; SLI = Speech/Language Impairment; SLD = Specific Learning Disability; ID = Intellectual Disability; MDS = Multiple Disabilities; DD = Developmental Delay; ESL = English as a Second Language; RC = response cards

## Participants

A total of 141 participants were included across all studies within this systematic review. A breakdown of participant characteristics by category is shown in Figures 2-5.

### *Age Range*

Figure 2 illustrates the age ranges of individuals included in this review. Participants ranged in age from five to 19 years old; however, due to the variability in age ranges across school systems, participant ages were indicated in Figure 2 by school level (i.e. elementary-aged, middle school-aged, or high school-aged). Of the 141 participants, 36 were elementary-aged (25%), 22 were middle school-aged (16%), and 83 were high school-aged (59%).

It should be noted that one study in particular included 68 high school-aged participants, a number of participants much greater than any other study included in this review. The inclusion of this study heavily influenced the number of participants categorized as high school-aged and should be taken into consideration when analyzing response card research and age.

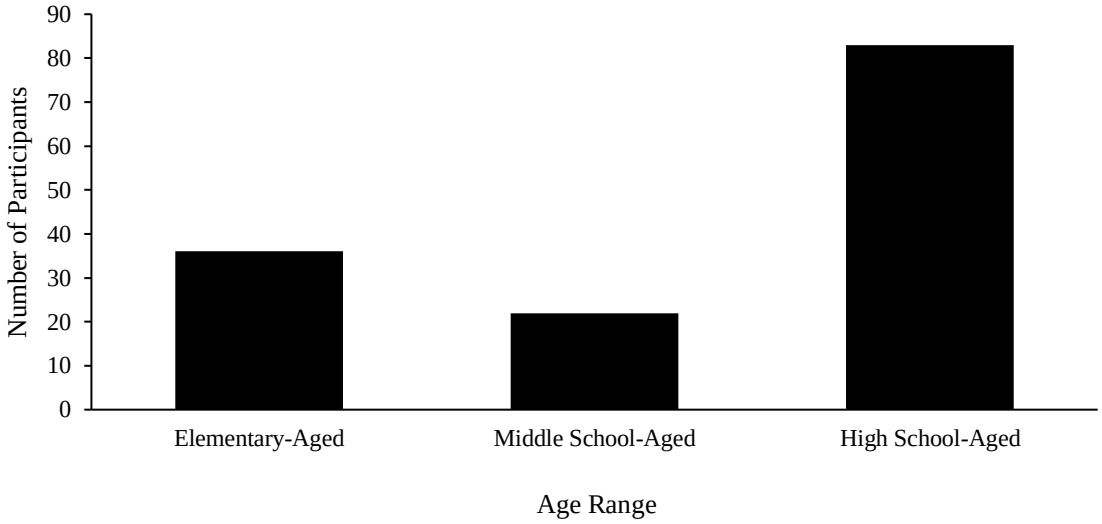


**Figure 2.** Age Range of Participants

### *Disability Category*

Figure 3 illustrates the disabilities under which all participants were classified. Several participants were reported to have both a primary and secondary disability, and in some cases, a tertiary disability. For all instances in which a participant was described as having more than one disability, all disabilities were recorded. It should also be noted that many participants included in this review did not have a diagnosed disability but participated in studies alongside students with a disability.

Twelve of the fourteen disability categories outlined within the Individuals with Disabilities Education Act (IDEA) were represented within this systematic review, including intellectual disability (ID) which is inclusive of mild, moderate, and severe diagnoses, speech or language impairment (SLI), specific learning disability (SLD), other health impairment (OHI) which is inclusive of attention deficit hyperactivity disorder (ADHD), emotional and behavioral disorders (EBD), autism spectrum disorder (ASD), physical disability, hearing impairment, multiple disabilities (MDS), traumatic brain injury (TBII), and developmental delay (DD). Also included was the category of English as a Second Language (ESL) as language barriers were deemed likely to impact intervention outcomes, though it should be noted that ESL is not a disability. For the purposes of this review, all participants whose disability was not specified were categorized as “Unknown Disability.”

Of the 141 participants, 63 were identified has having no disability (45%), 40 were identified has having ID (28%), 15 were identified as having SLI (11%), 13 were identified as having SLD (9%), 12 were identified as having OHI (including ADHD) (8%), 8 were identified as having EBD (6%), 7 were identified as having ASD (5%), 5

were identified as having a disability though it was not specified (3%), 4 were identified as having a physical disability (3%), 2 were identified has having a hearing impairment (1%), 1 was identified as having MDS (1%), 1 was identified as having a TBI (1%), and 1 was identified as having DD (1%). In addition, 3 participants were described as receiving ESL services (2%).

It should be noted that one study in particular included 68 participants, 58 of whom did not have a disability. The inclusion of this study heavily influenced the number of participants in general as well as the number of students labeled as having No Disability.

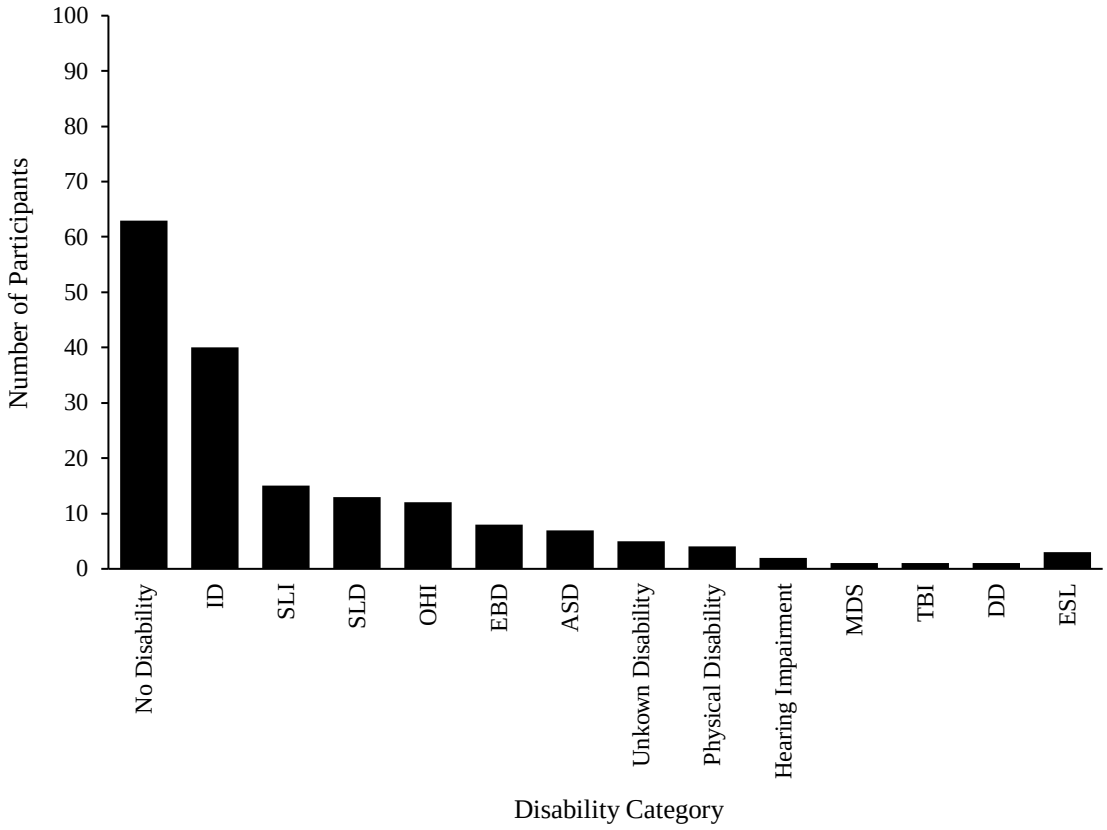


**Figure 3.** Disability Classification of Participants

### *Race*

Figure 4 illustrates the races of individuals included in this review. Five races were represented within the included studies including Caucasian, African-American or Black, Hispanic, Asian or Asian-American, and Multiracial. In addition, all participants whose race was not specified were categorized as "Unknown Race." Of the 141 participants, 74 were Caucasian (52%), 32 were African-American or Black (23%), 31 were not specified, (22%), 2 were Asian (1%), 1 was Hispanic (1%), and 1 was Multiracial (1%).

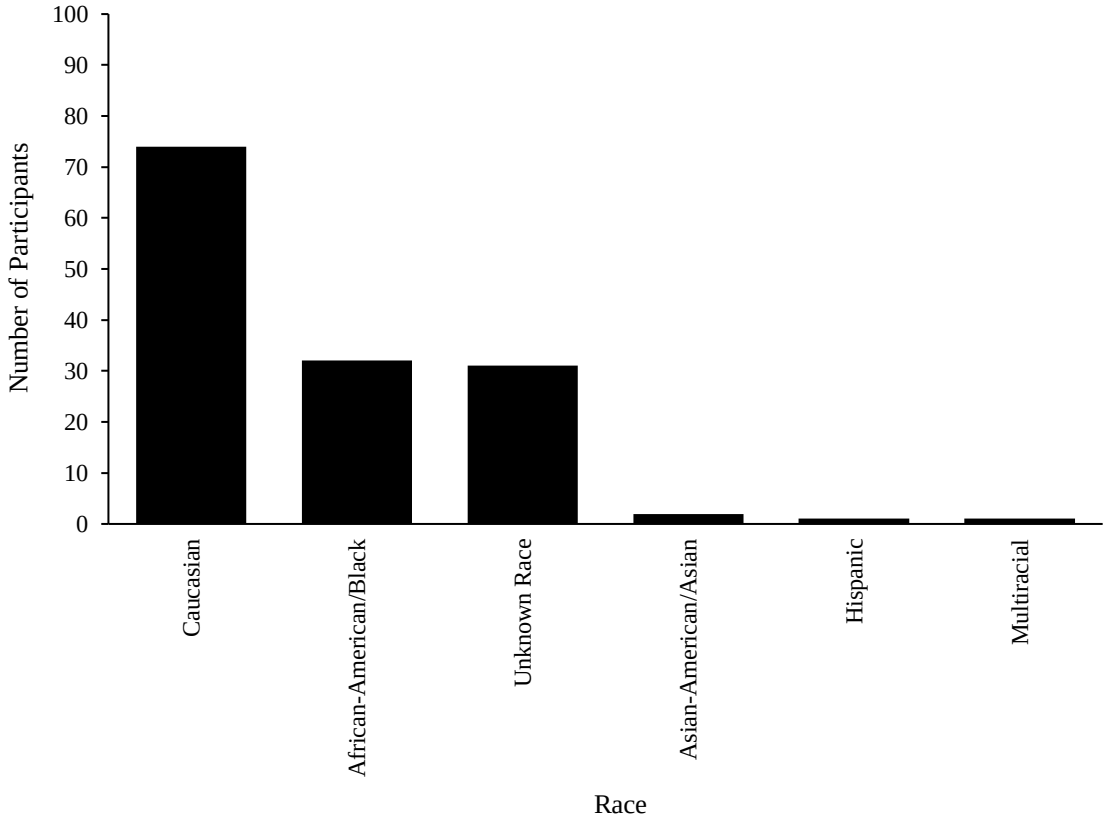


**Figure 4.** Race of Participants

### *Gender*

Figure 5 illustrates the genders of individuals included in this review. All genders were recorded as either male or female and all participants whose gender was not specified were categorized as "Unknown Gender." Of the 141 participants, 80 were male (57%), 56 were female (40%), and 5 were not specified (3%).

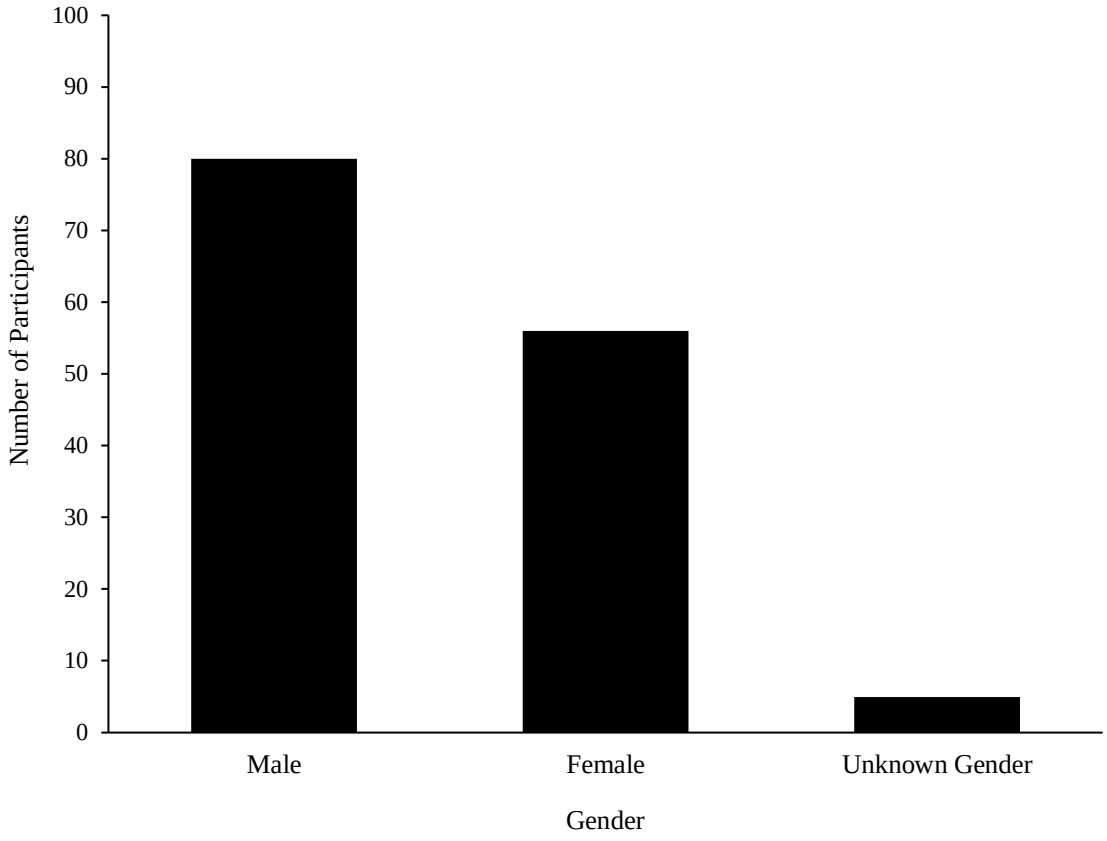


**Figure 5.** Gender of Participants

## Setting

The settings of all 15 studies included within this review were categorized in two ways: school level and learning environment. A breakdown of setting characteristics by category is shown in Figures 6-7.

### *School Level*

Figure 6 illustrates the school levels in which studies were conducted and includes elementary, middle, and high schools. Of the 15 studies, 7 took place in an elementary school setting (47%), 5 took place in a middle school setting (33%), and 3 took place in a high school setting (20%).

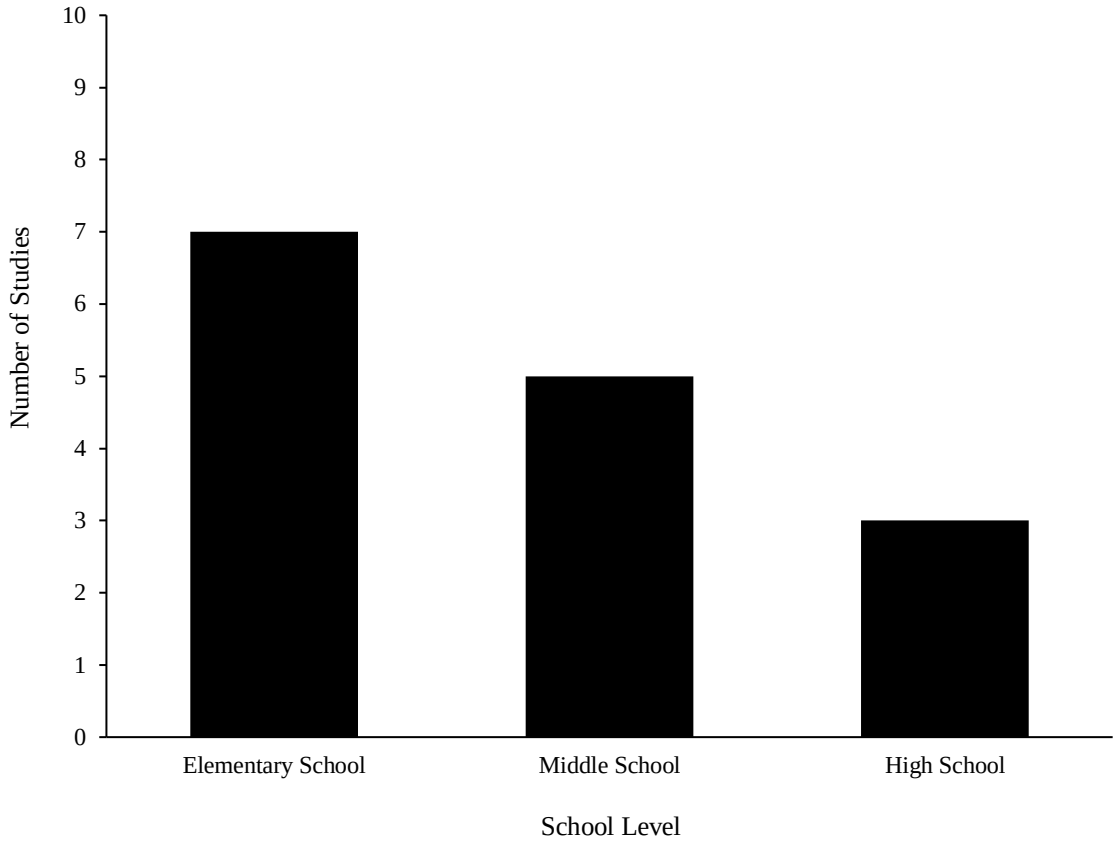


**Figure 6.** School Level Across Studies

### *Learning Environment*

Figure 7 illustrates the learning environment in which studies were conducted and includes either self-contained or inclusive. Of the 15 studies, 9 took place in a self-contained setting (60%) while 6 took place in an inclusive setting (40%).

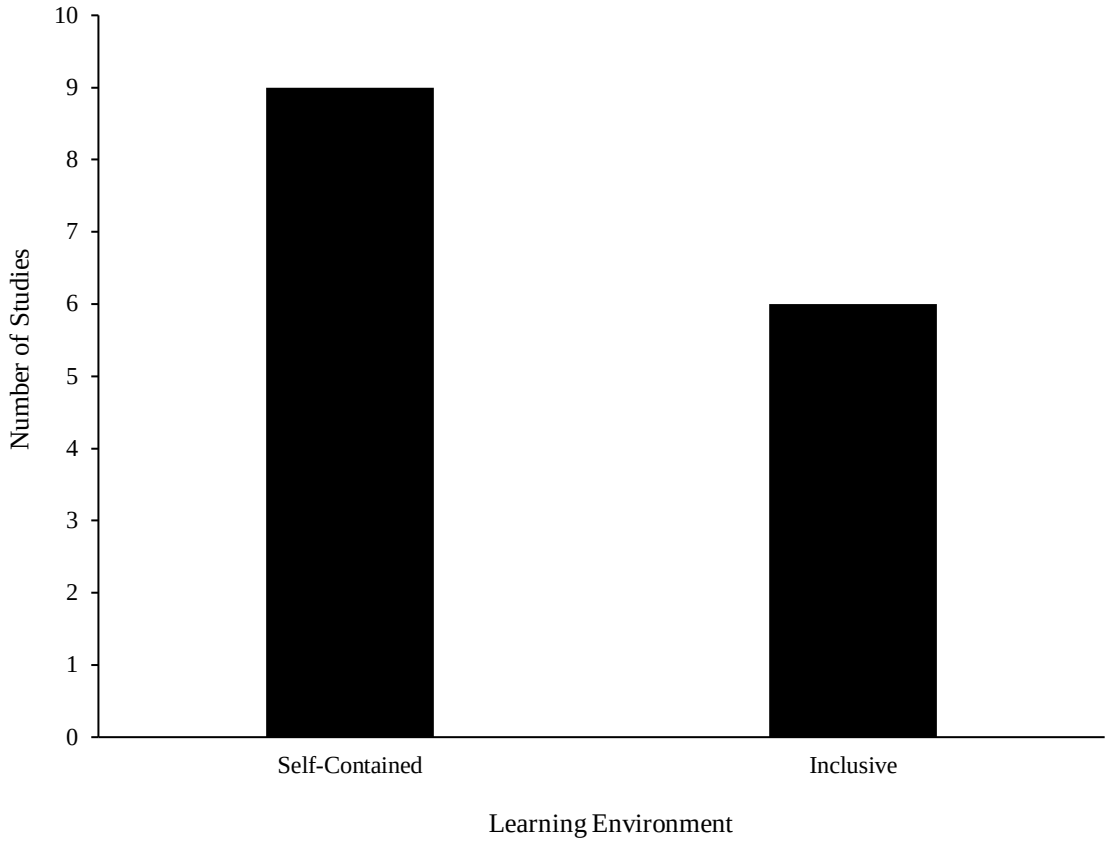


**Figure 7.** Learning Environment Across Studies

## Design

Figure 8 illustrates the experimental designs of all 15 studies included within this review. Designs include ABAB (reversal), multiple baseline (across dyads, across participants), and alternating treatments. Of the 15 studies, 11 utilized an ABAB reversal design (74%), 2 utilized an alternating treatments (13%), and 2 utilized a multiple baseline design (13%).

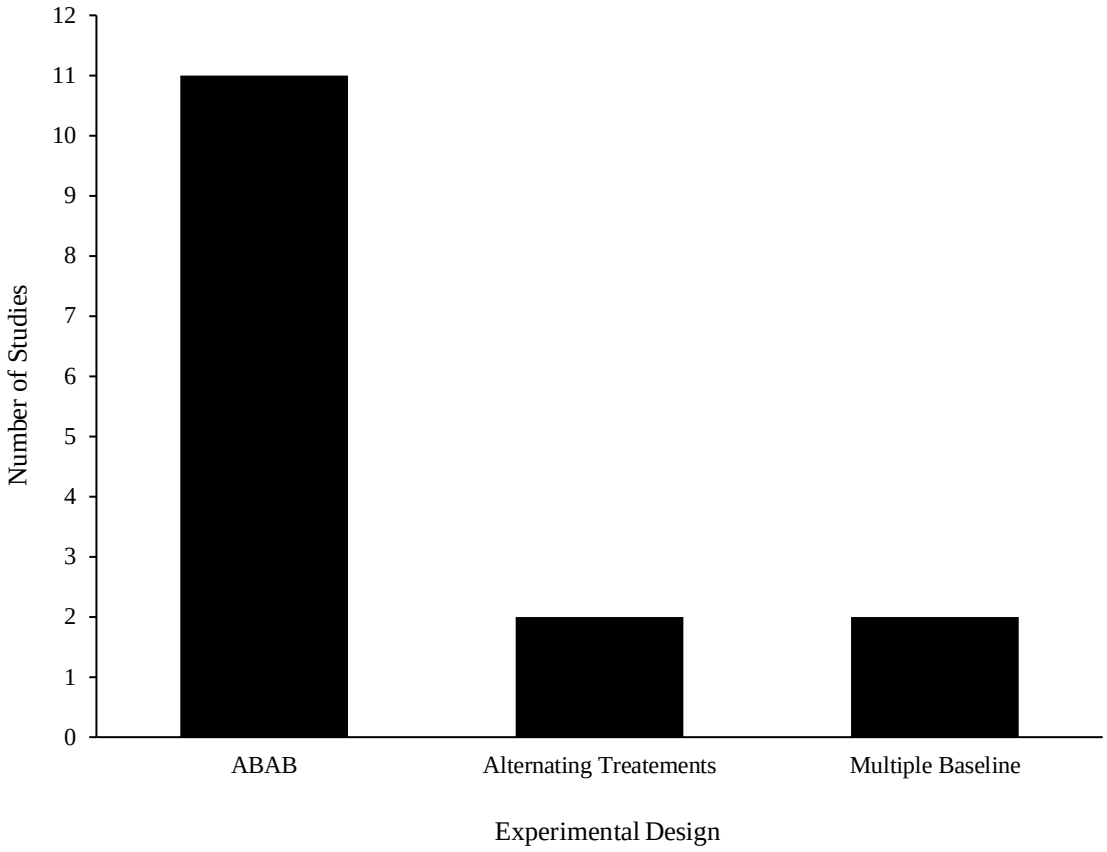


**Figure 8.** Experimental Design Across Studies

## Dependent Variables

Figure 9 illustrates the dependent variables which were manipulated within this review. The dependent variables include active responding, on-task behavior, academic achievement, inappropriate behavior, and correct responding. It should be noted that nearly all studies included within this review assessed the effects of response cards on more than one dependent variable. For all instances in which a study was described as having evaluated outcomes of more than one dependent variable, all variables were recorded.

Of the 15 studies, 13 evaluated the effects of the independent variable on active responding (87%), 8 evaluated the outcomes of on-task behavior (53%), 6 evaluated academic achievement following implementation of the response card intervention (40%), 6 recorded inappropriate behavior (40%), and 5 assessed correct responding as an additional criteria to active responding (33%).

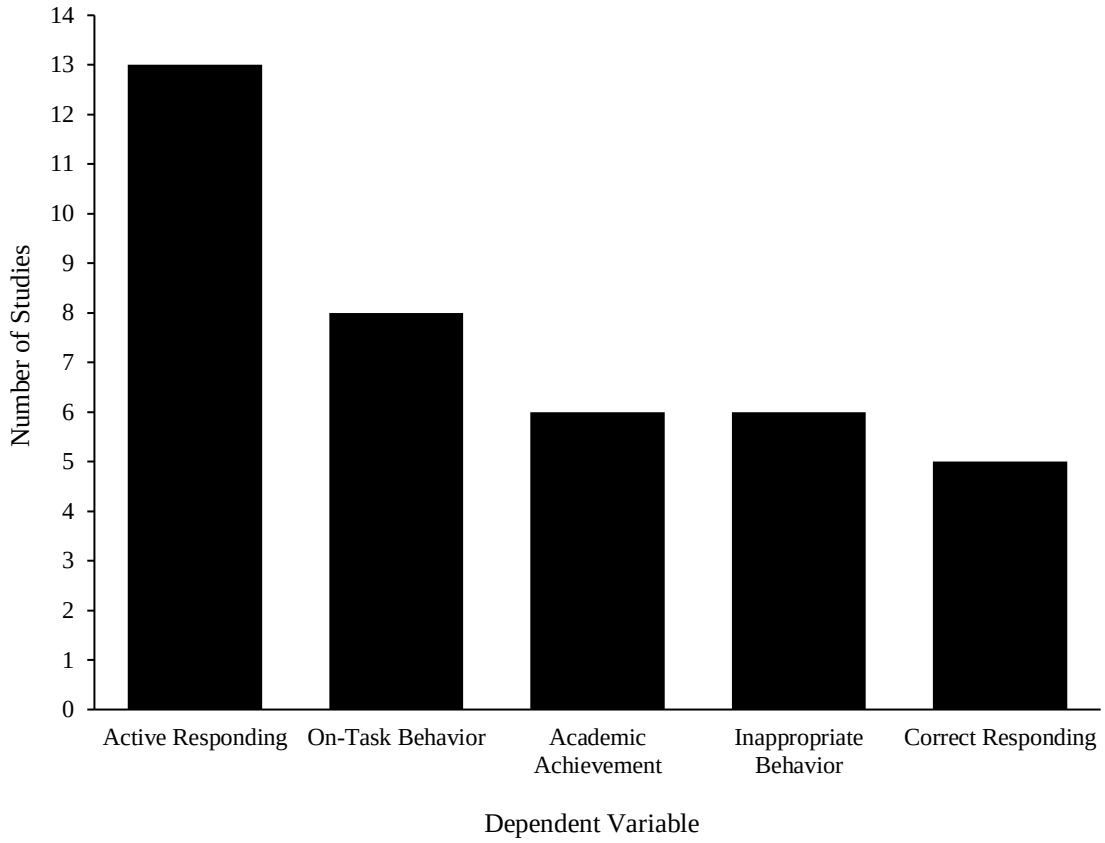


**Figure 9.** Dependent Variables Across Studies

## Independent Variable

Figure 10 illustrates the variety of response card formats included within this review. Formats include pre-printed, write-on, and manipulative. In one instance, multiple formats were used within a study; therefore, an additional category of

"Multiple" was included. Of the 15 studies, 7 included pre-printed response cards (46%), 6 included write-on response cards (40%), 1 included manipulative response cards (7%), and 1 utilized multiple formats (7%).

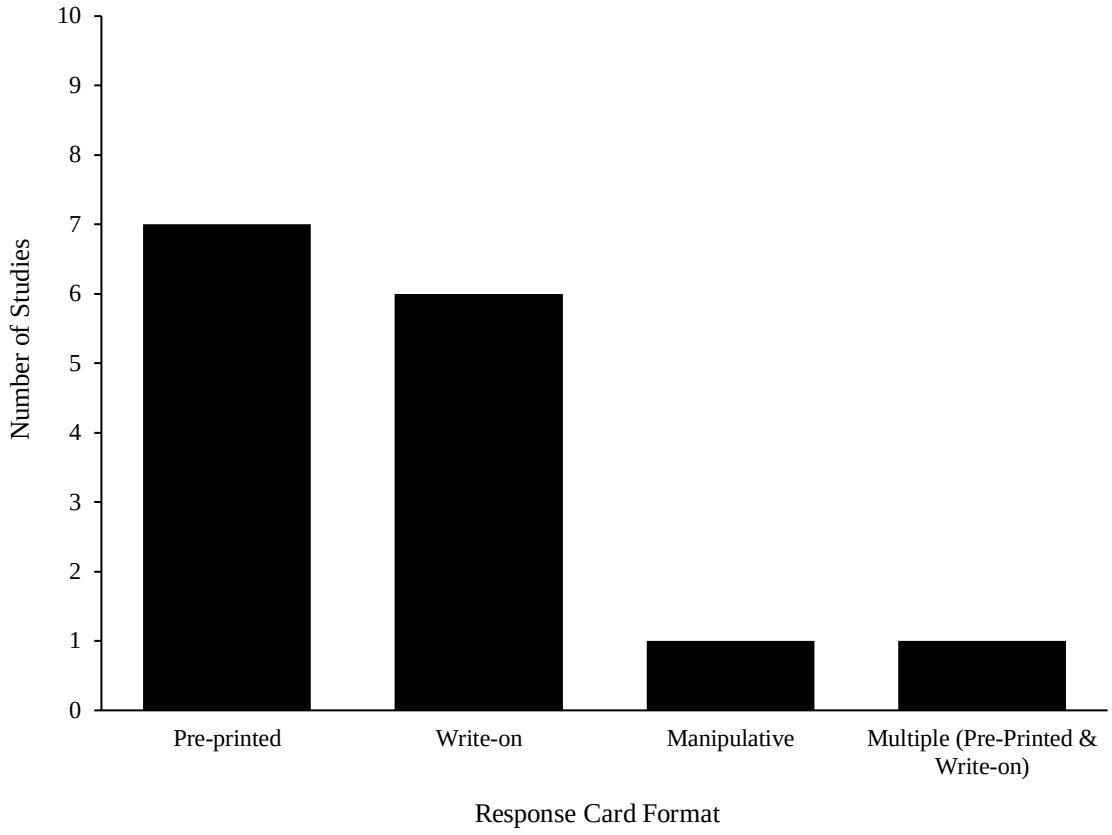


**Figure 10.** Response Card Format Across Studies

## Results

A total of five dependent variables were manipulated in the studies included within this review. A breakdown of outcomes for each is shown in Figures 11-14.

### *Active Responding*

Figure 11 illustrates the effect of response cards on active responding. Of the 15 studies included within this review, 13 evaluated active responding. In all 13 studies,

active responding was reported to have increased from hand raising (HR) conditions to response card (RC) conditions. However, only 12 of the 13 studies which manipulated active responding using the response card intervention actually produced data values which could be incorporated into the means found in Figure 11. The remaining study reported their results narratively.

In the first HR condition (HR1), an average of 24.68% of response tasks were completed across the 12 observed studies, as compared to 90.78% of response tasks in the first RC condition (RC1). Similarly, the second HR condition (HR2) yielded an average 28.5% active responding across all 12 studies, while 87.49% of response tasks were completed in the second RC condition (RC2).

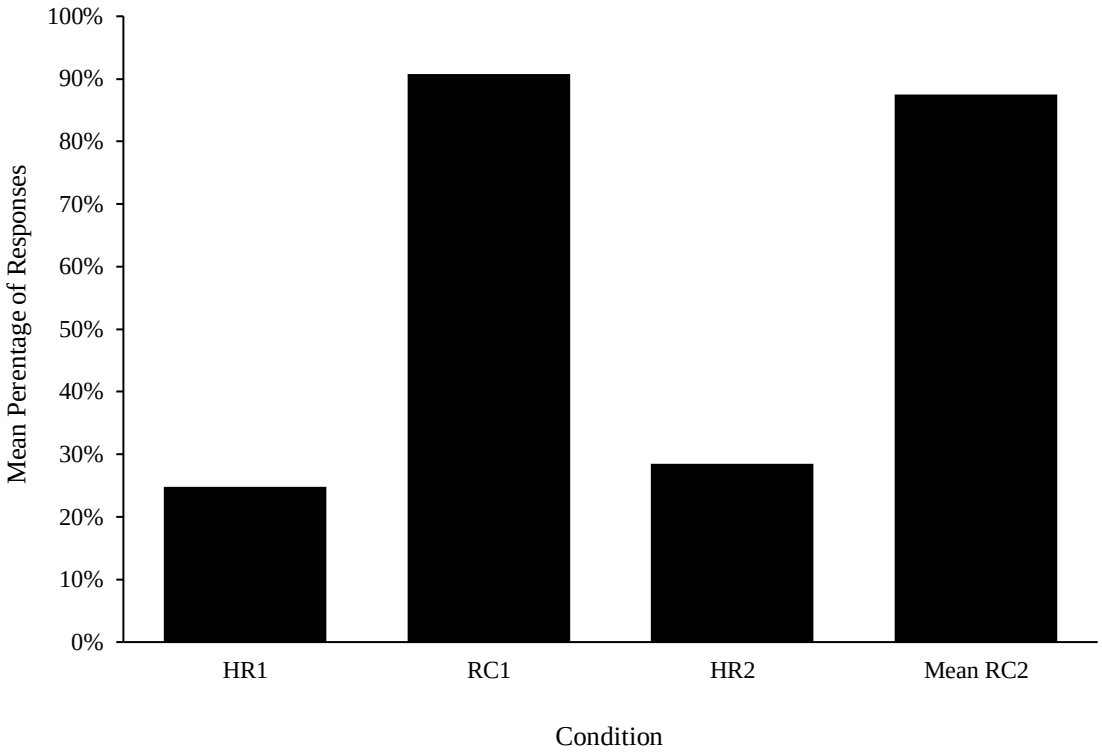


**Figure 11.** Active Responding Across Studies

### *Correct Responding*

Figure 12 illustrates the effect of response cards on correct responding. Of the 15 studies included within this review, five evaluated correct responding as an extension of active responding. In all five studies, correct responding was reported to have increased from hand raising (HR) conditions to response card (RC) conditions. However, only three of the five studies which manipulated correct responding using the response card intervention actually reported separate data values for the first set of conditions and the second. For the remaining two studies, averages for HR1 and HR2 were provided as well as averages for RC1 and RC2. In order to incorporate the data from those studies in these results, the average value reported for HR1-HR2 was used for *both* HR1 and HR2 in Figure 12. Similarly, the average value reported for RC1-RC2 was used for *both* RC1 and RC2 in Figure 12.

In the first HR condition (HR1), an average of 31.56% of response tasks were completed correctly across the five observed studies, as compared to 77.04% in the first RC condition (RC1). Similarly, the second HR condition (HR2) yielded an average 32.36% correct responding across all five studies, while 81.54% of response tasks were completed correctly in the second RC condition (RC2).

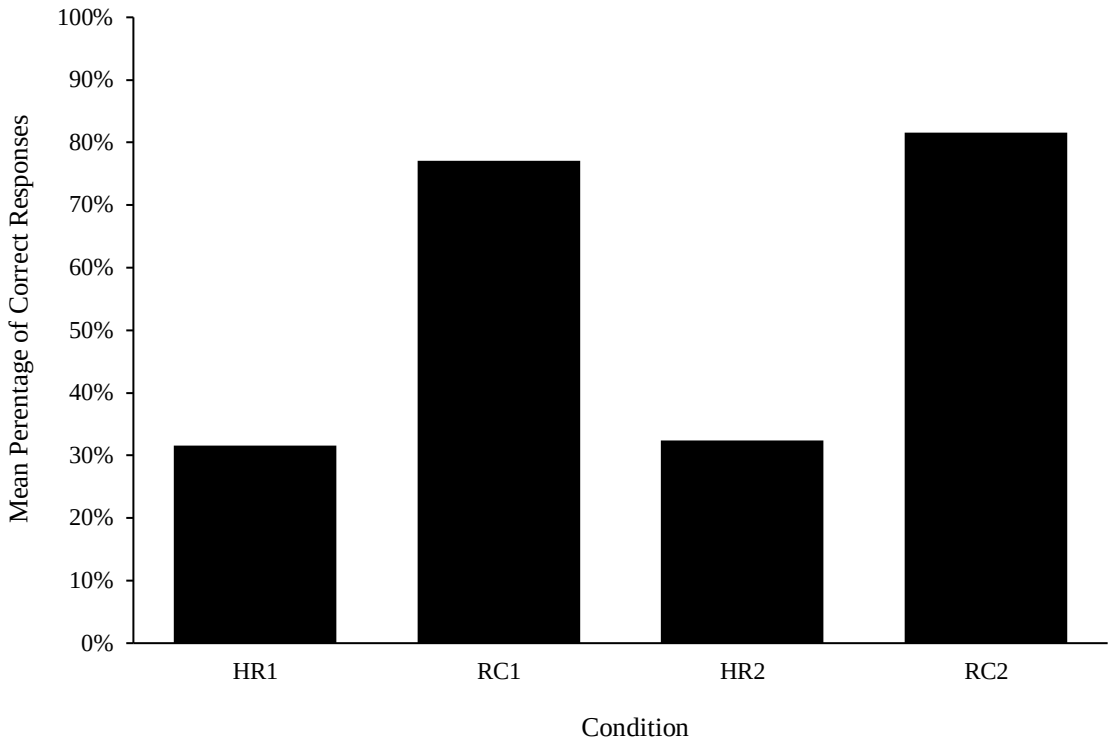


**Figure 12.** Correct Responding Across Studies

### *On-Task Behavior*

Figure 13 illustrates the effect of response cards on on-task behavior. Of the 15 studies included within this review, eight evaluated on-task behavior. In seven of the eight studies, time on-task was reported to have increased from hand raising (HR) conditions to response card (RC) conditions. In the remaining study, on-task behavior increased from HR1 to RC1 but decreased from HR2 to RC2.

In the first HR condition (HR1), participants were reported to have engaged in on-task behavior for an average of 62.6% of each session, as compared to 80.79% in the first RC condition (RC1). Similarly, the second HR condition (HR2) yielded an average of 65.04% time on-task, while participants were reported to have engaged in on-task behavior for an average of 73.57% of each session in the second RC condition (RC2).

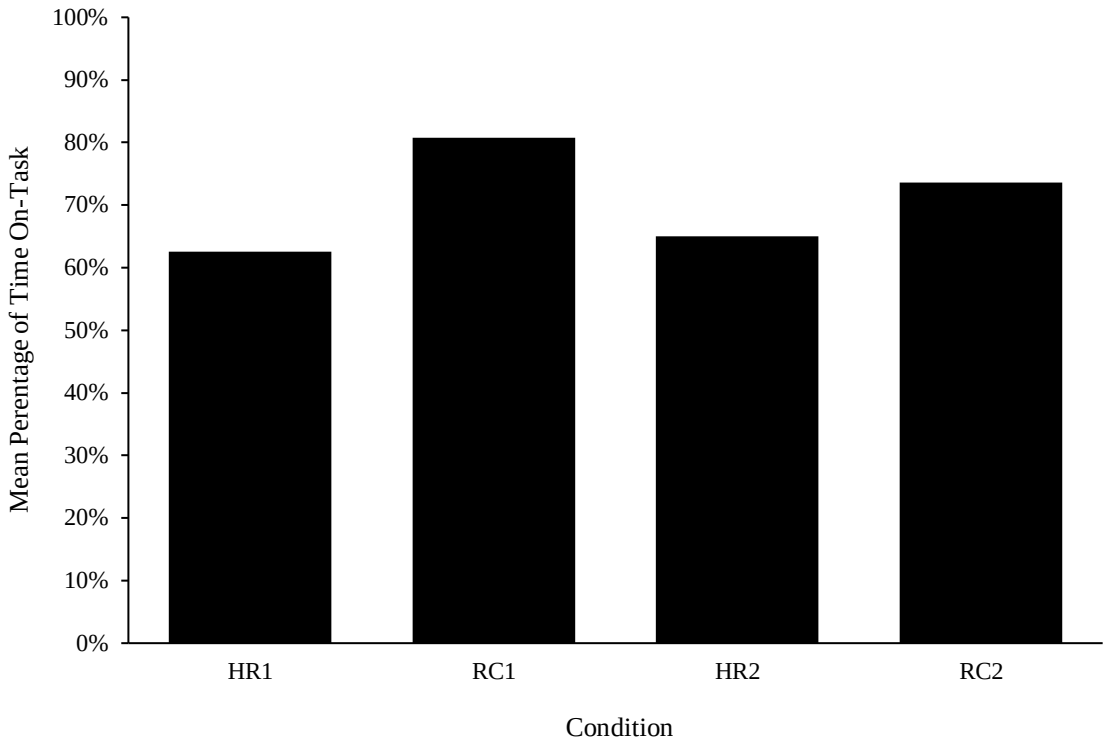


**Figure 13.** On-Task Behavior Across Studies

### *Inappropriate Behavior*

Of the 15 studies included within this review, six evaluated inappropriate behavior. However, due to the multitude of ways in which data were collected and reported for this dependent variable, there is no visual representation of this data. Rather, the results are described narratively below.

In the first of the six studies, inappropriate behavior was reported as the rate of inappropriate behaviors per minute during each condition. The mean rate of inappropriate behavior for the group was 0.77 during HR1 and 0.40 during RC1. Similarly, the mean rate of inappropriate behavior for the group was 0.89 during HR2 and 0.27 during RC2.

In the second of the six studies, inappropriate behavior was reported as the percentage of trials during which participants engaged in inappropriate behavior. Data

values were not recorded in this study and therefore cannot be reported for the purposes of this review; however, it was reported that only one of the four participants in this study displayed lower levels of inappropriate behavior during response card conditions.

In the third of the six studies, inappropriate behavior was reported using a 60-second momentary time sampling data recording system. The mean percentage of inappropriate behavior for the group was 49.6% during HR1 and 39.6% during RC1. Similarly, the mean rate of inappropriate behavior for the group was 52.2% during HR2 and 15.8% during RC2.

In the fourth of the six studies, inappropriate behavior was reported as the rate of inappropriate behaviors per minute during each condition. The mean rate of inappropriate behavior for the group was 0.96 during HR1 and 0.24 during RC1. Similarly, the mean rate of inappropriate behavior for the group was 0.82 during HR2 and 0.19 during RC2.

In the fifth of the six studies, inappropriate behavior was reported using a 10-second momentary time sampling data recording system. The mean percentage of inappropriate behavior for the group was 25.7% during HR1 and 27.7% during RC1. Similarly, the mean rate of inappropriate behavior for the group was 29.3% during HR2 and 38.0% during RC2.

In the last of the six studies, inappropriate behavior was reported using a 5s partial interval recording system for the first 6 min of each lesson. The mean percentage of intervals of inappropriate behavior for the group was 70.36% during HR1 and 1.84% during RC1. Similarly, the mean rate of inappropriate behavior for the group was 53.6% during HR2 and 6.74% during RC2.

### *Academic Achievement*

Of the 15 studies included within this review, six evaluated academic achievement via probes following both hand-raising conditions and response card conditions. However, much like the reporting of inappropriate behavior, academic achievement data was collected and reported in a multitude of ways. For that reason, there is no visual representation of this data. Rather, the results are described narratively below.

In the first of the six studies, academic achievement was observed and recorded via five-question assessment probes to assess skill acquisition within both hand-raising and response card conditions. The group of participants averaged 22.5% on assessment probes during HR1 and 42.3% during RC1. Similarly, the mean assessment probe score for the group was 61.5% during HR2 and 75.2% during RC2.

In the second of the six studies, academic achievement was assessed via pre-unit and post-unit tests. Both the pre-assessment and post-assessment were identical and were broken into two components: multiple choice and open-ended. The pre-unit multiple choice assessment was conducted during HR1 and resulted in an average score of 38.6%. The pre-unit open-ended assessment also occurring during the HR1 condition and resulted in an average score of 42.9%. Similarly, post-unit multiple choice and open-ended assessments were conducted during the RC1 condition and resulted in average scores of 55.7% and 53.8% respectively.

In the third of the six studies, academic achievement was observed and recorded via five-question pre- and post-test probes to assess skill acquisition within both hand-raising and response card conditions. The average pre-test score across HR conditions

was 36.1% while the average post-test score during the same conditions was 43.8%. Similarly, the mean pre-test score across RC conditions was 27.9% while the average post-test score during the same conditions was 54.7%.

In the fourth of the six studies, academic achievement was assessed via bi-weekly probes comprised of questions reflecting content learned during HR conditions as well as RC conditions. The average score among the group on questions taught during HR conditions was 56% while the average score on questions taught during RC conditions was 64.6%.

In the fifth of the six studies, academic achievement was observed and recorded via five to eight-question quizzes to assess daily skill acquisition within both hand-raising and response card conditions. The group of participants averaged 58% on quizzes during HR conditions and 84% during RC conditions.

In the last of the six studies, academic achievement was observed and recorded via weekly quizzes to assess skill acquisition within both hand-raising and response card conditions. The average quiz score among the group of participants was 84.8% during the HR1 condition and 96.4% during RC1. Similarly, the mean assessment probe score for the group was 75.2% during the HR2 condition.

## Quality of Study

Table 2 illustrates the results of the quality indicator assessment and quantifies the rigor and credibility of each single-case study included within this review. One value not represented in Table 2 is the percentage of quality indicators met for each study, which averaged 92% across studies, with a range of 75%-100%.

The first category of indicators on which studies were assessed was description of participants and setting. Of the 15 studies, 100% described participants with sufficient detail (n=15), though only 60% of studies described the process by which participants were selected (n=9). Additionally, 100% of studies described the setting in which the study took place with enough precision to allow for replication (n=15).

The next series of indicators pertained to the dependent variable. 100% of the studies assessed provided operational definitions of the dependent variables (n=15), measured the dependent variables in such a way that would generate a quantifiable index (n=15), described the measurement of dependent variables to allow for replication (n=15), measured the dependent variable over time (n=15), and collected reliability or IOA data on all dependent variables observed (n=15).

Following assessment on the dependent variables of each study, the independent variable was evaluated. Of the 15 studies, only 80% described the response card intervention with replicable precision (n=12), though 100% of studies systematically manipulated the independent variable under the control of the experimenter (n=15). In addition, 87% of studies measured the fidelity of implementation for the independent variable (n=13).

The next category of indicators on which studies were assessed was baseline information. Of the 15 studies, 93% included a baseline phase which provided measures of the dependent variable and established predictive patterns of responding (n=14). Furthermore, 87% of the studies included descriptions of baseline conditions which would allow for replication (n=13).

**Table 2**
*Quality Assessment of Response Card Studies*

| Quality Indicators | Percentage of Points Awarded Per Indicator (%) | |
|---|---|---|
| | Yes | No |
| ***Description of participants and settings:*** | | |
| 1. Participants are described with sufficient detail to allow others to select individuals with similar characteristics (e.g., age, gender, disability diagnosis) | 100% | 0% |
| 2. The process for selecting participants is described with replicable precision | 60% | 40% |
| 3. Critical features of the physical setting are described with sufficient precision to allow replication | 100% | 0% |
| ***Dependent Variables:*** | | |
| 4. Dependent variables are described with operational precision | 100% | 0% |
| 5. Each dependent variable is measured with a procedure that generates a quantifiable index | 100% | 0% |
| 6. Measurement of the dependent variable is valid and described with replicable precision | 100% | 0% |
| 7. DVs are measured repeatedly over time | 100% | 0% |
| 8. Data are collected on the reliability or IOA associated with each dependent variable and IOA levels meet minimal standards (e.g., IOA = 80%; Kappa= 60%) | 100% | 0% |
| ***Independent variables:*** | | |
| 9. Independent variable is described with replicable precision | 80% | 20% |
| 10. Independent variable is systematically manipulated and under the control of the experimenter | 100% | 0% |
| 11. Overt measurement of the fidelity of implementation for the independent variable is highly desirable | 87% | 13% |
| ***Baseline:*** | | |
| 12. The majority of single subject research studies will include: a baseline phase that provides repeated measures of a dependent variable and establishes a pattern of responding that can be used to predict the pattern of future performance, if introduction or manipulation of the IV did not occur | 93% | 7% |
| 13. Baseline condition are described with replicable precision | 87% | 13% |
| ***Experimental control/internal validity:*** | | |
| 14. The design provides at least three demonstrations of experimental effect at three different points in time | 93% | 7% |
| 15. The design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses) | 93% | 7% |
| 16. The results document a pattern that demonstrates experimental control | 73% | 27% |
| ***External validity:*** | | |
| 17. Experimental effects are replicated across participants, setting, or material to establish external validity | 100% | 0% |
| ***Social validity:*** | | |
| 18. The DV is socially important | 100% | 0% |
| 19. The magnitude of change in the DV resulting from the intervention is socially important | 67% | 33% |
| 20. Implementation of the IV is practical and cost effective | 100% | 0% |

Experimental control and internal validity were also assessed for quality. 93% of studies followed an experimental design which provided at least three demonstrations of experimental effect and which controlled for threats to internal validity (n=14). Additionally, 73% of studies demonstrated experimental control (n=11).

The final categories included within the quality assessment checklist included indicators related to external and social validity. 100% of studies reviewed replicated experimental effects to establish external validity (n=15), included dependent variables that were socially important (n=15), and involved the implementation of a practical and cost-effective independent variable (n=15). However, only 67% of studies indicated that the magnitude of change in the dependent variables resulting from the intervention were socially important (n=10).

# CHAPTER 4

# DISCUSSION

This review sought to extend the findings of Randolph (2007), Horn (2010), and Schnorr et. al. (2016). Specifically, the review investigated a) the extent to which response cards have been reported to influence student behaviors, both desirable and undesirable (i.e. active responding, correct responding, on-task behavior, inappropriate behavior); b) the extent to which response cards have been reported to increase academic achievement; c) the extent to which response card research incorporates features of high-quality single-case research; and d) opportunities for future research.

The first research question investigated within this review addressed the extent to which response cards have been reported to influence student behaviors, both desirable and undesirable. The student behaviors assessed and included within in this review include active responding, correct responding, on-task behavior, and inappropriate behavior. Analysis of the studies which assessed those dependent variables reveal fairly consistent findings for three of the four behaviors, including active responding, correct responding, and on-task behavior. However, reported outcomes relating to inappropriate behavior were much more variable and demonstrated much less effect.

Of all four behavior-based dependent variables evaluated within the present study, active responding is the only one that has also been evaluated in all prior systematic reviews on the same topic (though the variable of interest was referred to using different terminology in several of those publications). The results of this review extend the findings of Randolph (2007), Horn (2010), and Schnorr et. al. (2016), indicating that the use of response cards is highly effective in increasing active responding among students

and yields greater participation than the more traditional hand-raising method. Given that this systematic review included studies across age and school levels and in both self-contained and inclusive settings, these findings suggest that response cards are effective not only for a younger population of learners, as was most recently suggested by Schnorr et. al. (2016), but a wide range of students in a wide range of settings.

As an extension of active responding, this study analyzed the effect of the intervention on correct responding to determine whether the use of response cards would also result in more accurate responses during response card conditions. This measure differs from academic achievement as it asses only one piece of information at the exact moment of hand-raising or response-card responding, not large quantities of information following the intervention as is the case when measuring academic achievement. Of the three prior reviews, only Horn (2010) evaluated the impact of response cards on correct responding. The present study extends the findings of Horn (2010) and suggests that response cards can increase attempted and accurate responses.

The third variable evaluated within this review was that of on-task behavior. Only Horn (2010) focused specifically on the topic of on-task behavior in his previous review of the response card literature, and again, the results of this review mirror his findings. Though they yielded a much less significant effect, the use of response cards was reported to increase on-task behavior during response card conditions in nearly all studies of interest, suggesting that the tool might be considered more engaging and more preferred over other methods of instruction among students. This is likely the case because of the interactivity of the intervention for all learners within an instructional

group. The same level of interactivity cannot be said for the traditional hand-raising method which allows for the responding of only one student at a time.

Finally, outcomes relating to inappropriate behavior were analyzed. Unlike the outcomes of the three desirable behaviors previously discussed, the results of studies which evaluated the effect of response cards on inappropriate behavior were much more variable. Though Randolph (2007) concluded that levels of off-task behavior were significantly lower during response card conditions, such conclusions could not be drawn from the more current body of research. The reason for such a discrepancy is unknown but does warrant consideration during the instructional tool selection process. Though response cards have been reported to increase desirable behaviors in the classroom setting, educators might proceed more skeptically when selecting the intervention if they are doing so for the purposes of decreasing undesirable behavior.

The second research question investigated within this review addressed the extent to which response cards have been reported to increase academic achievement. Of the six studies which evaluated the effect of response cards on academic achievement, all six reported an increase in average assessment scores during response card conditions as opposed to baseline/hand-raising conditions. Analysis of the results of those studies reveal that the use of response cards during instructional periods are likely to result in greater skill acquisition than the traditional hand-raising method. These findings were consistent across studies, regardless of assessment length or format, and suggest that use of response cards as an instructional tool will not only increase student engagement but skill acquisition and comprehension of curriculum as well. This is a major finding as it

indicates that response cards are not only beneficial for short-term outcomes, but for long-term retention and academic success as well.

The third research question investigated within this review addressed the extent to which response card research incorporates features of high-quality single-case research. This research question was proposed as a stepping-stone toward determining whether response cards could be established as an evidence-based practice, a classification of which would even further validate the use of such a tool in the classroom.

Whether the present response card research incorporated features of high-quality single-case research was evaluated by means of the Horner et. al. (2005) checklist. Four of the seven categories within the checklist reveal high-quality and rigorous research, including the dependent variable category (100% of points awarded for all three indicators within the category), the baseline category (93% of points awarded for indicator number 12 and 87% of points awarded for indicator number 13), the experimental control/internal validity category (93% of points awarded for indicators 14 & 15 and 73% of points awarded for indicator 16), and the external validity category (100% of points awarded for indicator 17). However, three categories in particular revealed low ratings indicating lower-quality and less rigorous research: description of participants/setting (100% of points awarded for indicators 1 & 3 but only 60% of points awarded for indicator 2), independent variable (80% of points awarded for indicator 9, 100% of points awarded for indicator 10, and 87% of points awarded for indicator 11), and social validity (100% of points awarded for indicators 18 & 20 but only 67% of points awarded for indicator 19).

While high ratings in the first four categories endorse strong operational definitions of dependent variables and establishment of a strong baseline across studies, low ratings in the latter categories indicate a need for more precise explanations of the participant selection process, clearer definitions of the independent variable, more precise procedures for implementation, and a re-evaluation of whether the magnitude of outcomes of the intervention are socially important. Low social validity ratings are of particular concern as the selection of an intervention within the classroom depends largely on acceptability and feasibility. Further research on the social validity of the intervention and outcomes which it yields is necessary if response cards are to continue serving individuals with disabilities in the school setting.

Determination of an intervention as evidence-based goes beyond a quantification of the rigor of the single-case research studies of an intervention. Until this point, research has suggested response cards are evidence-based solely at the elementary level (Schnorr et. al., 2016), making the evaluation of the present research critical in expanding that classification. Though not addressed specifically within this review, Horner et. al. (2005) outline five standards which loosely align with their study quality indicator checklist and which can be used to establish a practice as evidence-based. Those five standards include (a) operational definition of the intervention (loosely aligned to indicator 9 on the checklist); (b) clearly defined context in which the practice is to be used (loosely aligned to indicators 1-3 on the checklist); (c) implementation of the intervention with fidelity (loosely aligned to indicator 11 on the checklist); (d) a functional relation between the intervention and change in dependent measures (loosely aligned to indicator 16 on the checklist); and (e) replication of the experimental effects

across a sufficient number of studies, researchers, and participants to allow confidence in the findings (loosely aligned to indicator 17 on the checklist, though this standard also calls for replication of effects across five or more studies with a total of at least 20 participants). As these standards were not applied to the results of the single-case research included within this review, further analysis of the literature is recommended to determine whether response cards can be established as a legitimate evidence-based practice.

The fourth and final research question investigated within this review addressed opportunities and directions for future research. Suggestions for future research will be discussed at the conclusion of this review.

## Limitations

Though this review yielded findings consistent with those of Randolph (2007), Horn (2010), and Schnorr et. al. (2016), there exist several limitations to the present study. One limitation of this review is the lack of attention to other, more technologically advanced, forms of response systems. Given that only the term *response cards* was included in preliminary database searches, all other response systems were excluded. Some potential alternatives to traditional response cards include clickers and online response platforms (Kahoot, Quizizz, Quizlet, etc.) which involve the use of a phone or laptop for responding. Even though these alternatives are less cost-efficient and may not be as accessible, further research on a broader scope of response systems is required to determine whether other formats yield consistent outcomes among the dependent variables of focus.

An additional limitation of this review is the way in which treatment effect was quantified for each study. Because responses within each condition were averaged across all studies as opposed to being reported individually, analysis within the present study does not take into account trend, level, and variability of behavior between conditions. However, such a limitation was counterbalanced by the inclusion of an experimental control indicator within the study quality checklist.

A final limitation to this study is that while study quality was assessed and all studies were deemed acceptably strong (at least 75% of points awarded for every study), standards which would have allowed for the establishment of response cards as evidence-based were not addressed. Though deemed an evidence-based practice among elementary-aged students by Schnorr et. al. (2016), response cards have yet to be determined an evidence-based practice on a larger, more encompassing scale. Future research should address this particular focus.

## Implications for Practice

Federal law requires that educators implement evidence-based classroom practices and interventions that are proven effective through scientific research. Results of this review offer several implications for educators. Similar to Horn (2010) and Schnorr et. al. (2005), the current review establishes response cards as a cost-effective and low-maintenance tool that can be used to impact student behavior and academic achievement. In addition, response cards can be used to evaluate student understanding in real-time and in large groups, as opposed to the more traditional method of hand-raising which allows for responding by only one student at a time. Not only does this maximize instructional time and make student understanding much more accessible to the teacher,

but it reduces the number of opportunities for students to demonstrate inappropriate or off-task behaviors.

Another implication for educators is that the studies analyzed in this review involved implementation of the intervention in both inclusive and self-contained learning environments. As many schools and districts across the country are promoting the inclusion of more and more students with disabilities in the general education environment, the results of this review are made even more relevant and practical. As an outcome of the present study, it is suggested that response cards can be used to serve a multitude of functions among an entire classroom of students, both with and without disabilities, or they can be used to help smaller groups of students, particularly those with disabilities, to access the general curriculum. Educators are encouraged to consider the use of response cards not only when planning large group instruction, but possibly as a tool during small group instruction depending on classroom makeup.

## Recommendations for Future Research

Further research is necessary to bridge some of the remaining gaps in the body of response card research. First and foremost, the quality of single-case research on response cards as an instructional tool or intervention must be held paramount. As was revealed in this study, more precise explanations of the participant selection process, clearer definitions of the independent variable and procedures for implementation, and a re-evaluation of the social significance of response cards and their outcomes are imperative if such an intervention is to continue serving schools to the fullest extent.

Future research is also necessary to further evaluate the effect of response cards specifically at the secondary level. Nearly half of the studies included within this review

were conducted at the elementary level, indicating a need for additional research on the tool within middle and high school settings. In addition, subject areas with which response cards have the greatest effect were not examined within this particular review and ought to be evaluated in future research.

Another area for future research is analysis of the effects of response *systems* on student behavior and academic achievement. A more encompassing search is necessary to determine if response systems, regardless of form, serve the same function as those evaluated in the present study. It is suggested, however, that this research focus not only the student outcomes of such tools, but the social validity of the intervention across stakeholders (e.g., students, teachers, administrators, parents, etc.) It is expected that as response systems become more costly and less accessible, the acceptance of such a tool may decrease at a comparable rate for some stakeholders, but not necessarily for all. The findings of such research would play a critical role in the selection of response cards as an instructional tool or intervention by educators and administrators.

A final area of future research is that of establishing response cards as evidence-based. Response cards have yet to be classified an evidence-based practice at all school levels and among all ability levels. Until response cards are established as evidence-based, their validity as an instructional tool will be in question and their benefits will not be as far-reaching. Future studies ought to pay careful attention to the quality of their single-case research, and future reviews should focus on using Horner's five standards to evaluate whether the quality and outcomes of existing research deem the practice evidence-based.

## REFERENCES

The references with asterisks were included in the systematic review.

Ault, M. J., & Horn, C. K. (2018). Increasing active engagement: Guidelines for using student response systems. *Journal of Special Education Technology, 33*(3), 207-216.

* Berrong, A. K., Schuster, J. W., Morse, T. E., & Collins, B. C. (2007). The effects of response cards on active participation and social behavior of students with moderate and severe disabilities. *Journal of Developmental and Physical Disabilities, 19*(3), 187-199.

* Boden, L. J. (2015). *Effects of Response Cards on Math Performance for Students with Moderate Intellectual Disability* (Doctoral dissertation).

* Bondy, A. H., & Tincani, M. (2018). Effects of response cards on students with autism spectrum disorder or intellectual disability. *Education and Training in Autism and Developmental Disabilities, 53*(1), 59-72.

* Cakiroglu, O. (2014). Effects of preprinted response cards on rates of academic response, opportunities to respond, and correct academic responses of students with mild intellectual disability. *Journal of Intellectual & Developmental Disability, 39*(1), 73-85.

* Clarke, L. S. (2010). *Use of picture response cards in the general education classroom to increase participation of students with mild mental disabilities* (Doctoral dissertation).

* Clarke, L. S., Haydon, T., Bauer, A., & Epperly, A. C. (2016). Inclusion of students with an intellectual disability in the general education classroom with the use of response cards. *Preventing School Failure, 60*(1), 35-42.

Cook, B. G., Tankersley, M., & Landrum, T. J. (2009). Determining evidence-based practices in special education. *Exceptional Children*, 75, 365–383.

* Davis, L. L., & O'Neill, R. E. (2004). Use of response cards with a group of students with learning disabilities including those for whom English is a second language. *Journal of Applied Behavior Analysis, 37*(2), 219-222.

* Davis, M. (2020). *The differential effects of hand-raising and digital response cards on active engagement of high school students with mild to moderate disabilities during literacy activities* (Master's Thesis).

* Didion, L. A., Toste, J. R., & Wehby, J. H. (2020). Response Cards to Increase Engagement and Active Participation of Middle School Students with EBD. *Remedial and Special Education, 41*(2), 111–123.

* Duchaine, E. L., Jolivette, K., Fredrick, L. D., & Alberto, P. A. (2018). Increase engagement and achievement with response cards: Science and mathematics inclusion classes. *Learning Disabilities: A Contemporary Journal, 16*(2), 157-176.

* Dudley, M. (2013). *The effects of response cards on students' math quiz performance* (Doctoral dissertation).

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York, NY: Routledge.

Helf, S. (2015). Increasing opportunities for student responding: Response cards in the classroom. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 88*(6), 182-184.

Herring, T. J., & Woolsey, M. L. (2020). Three suggested teaching strategies for students who are deaf or hard of hearing. *Support for Learning*, *35*(3), 346-358.

Heward, W. L. (1994). Three low-tech strategies for increasing the frequency of active student response during group instruction. In R. Gardner, III, D. Sainato, J. O. Cooper, T. Heron, W. L. Heward, J. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurable superior instruction* (pp. 283–320). Pacific Grove, CA: Brooks/Cole.

Heward, W. L. (1997). Four validated instructional strategies. *Behavior and Social Issues*, *7*(1), 43-51.

Horn, C. (2010). Response cards: An effective intervention for students with disabilities. *Education and Training in Autism and Developmental Disabilities, 45*(1), 116-123.

* Horn, C., Schuster, J., & Collins, B. (2006). Use of Response Cards to Teach Telling Time to Students with Moderate and Severe Disabilities. *Education and Training in Developmental Disabilities, 41*(4), 382-391.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional children, 71*(2), 165-179.

* McCargo, M. G. (2017). *The effects of Plickers as response cards on academic engagement behavior in high school students* (Master's Thesis).

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine, 6*(7), e1000097.

* Müllerke, N., Duchaine, E. L., Grünke, M., & Karnes, J. (2019). The Effects of a Response Card Intervention on the Active Participation in Math Lessons of Five Seventh Graders with Learning Disabilities. *Insights into Learning Disabilities, 16*(2), 107–120.

Nagro, S. A., Hooks, S. D., Fraser, D. W., & Cornelius, K. E. (2016). Whole-group response strategies to promote student engagement in inclusive classrooms. *Teaching Exceptional Children, 48*(5), 243-249.

Owiny, R. L., Spriggs, A. D., Sartini, E. C., & Mills, J. R. (2018). Evaluating response cards as evidence based. *Preventing School Failure: Alternative Education for Children and Youth, 62*(2), 59-72.

Randolph, J. J. (2005). A quantitative synthesis of the research on response cards on participation, academic achievement, behavioral disruptions, and student preference. In M-L. Julkunen (Ed.), *Learning and instruction in multiple contexts and settings III: Proceedings of the fifth Joensuu symposium on learning and instruction*. (Bulletins of the Faculty of Education, No. 96, pp. 149–165). Finland: University of Joensuu Press.

Randolph, J. J. (2007). Meta-Analysis of the Research on Response Cards: Effects on Test Achievement, Quiz Achievement, Participation, and Off-Task Behavior. *Journal of Positive Behavior Interventions*, *9*(2), 113–128.

Rosenshine, B. (2012). *Principles of instruction: Research-based strategies that all teachers should know.* American Educator, 36(1), 12–39.

Schnorr, C. I., Freeman-Green, S., & Test, D. W. (2016). Response Cards as a Strategy for Increasing Opportunities to Respond: An Examination of the Evidence. *Remedial and Special Education*, *37*(1), 41–54.

Tincani, M., & Twyman, J. S. (2016). Enhancing Engagement through Active Student Response. *Center on Innovations in Learning, Temple University*.

* Wood, C. L., Mabry, L. E., Kretlow, A. G., Lo, Y. Y., & Galloway, T. W. (2009). Effects of reprinted response cards on students' participation and off-task behavior in a rural kindergarten classroom. *Rural Special Education Quarterly, 28*(2), 39-47.