8-9-2022

Improving Reading Comprehension: An Investigation of Two Intervention Procedures and the Social Validity of the "Ask-Read-Tell" Reading Strategy

Quynh Doan Wellons
Georgia State University

ACCEPTANCE

This dissertation, IMPROVING READING COMPREHENSION: AN INVESTIGATION OF TWO INTERVENTION PROCEDURES AND THE SOCIAL VALIDITY OF THE "ASK-READ-TELL" READING STRATEGY , by QUYNH-NHU DOAN WELLONS, was prepared under the direction of the candidate's Dissertation Advisory Committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the College of Education & Human Development, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chairperson, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty.

_____
Andrew T. Roach, Ph.D.
Committee Chair

_____   _____
Claire Donehower, Ph.D.   Emily Graybill, Ph.D.
Committee Member   Committee Member

_____
Joel Meyers, Ph.D.
Committee Member

_____
Date

_____
Brian J. Dew, Ph.D.
Chairperson, Department of Counseling and
Psychological Services

_____
Paul A. Alberto, Ph.D.
Dean, College of Education &
Human Development

AUTHOR'S STATEMENT

By presenting this dissertation as a partial fulfillment of the requirements for the advanced degree from Georgia State University, I agree that the library of Georgia State University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote, to copy from, or to publish this dissertation may be granted by the professor under whose direction it was written, by the College of Education & Human Development's Director of Graduate Studies, or by me. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without my written permission.

Quynh-Nhu Doan Wellons

CURRICULUM VITAE

Quynh-Nhu Doan Wellons

ADDRESS: 105 Primrose Pass
                       Newnan, GA 30265

EDUCATION:

Ph.D.	2022	Georgia State University School Psychology
Ed.S.	2021	Georgia State University School Psychology
M.Ed.	2019	Georgia State University School Psychology
M.A.	1998	American University Psychology
B.A.	1995	University of Virginia Psychology

PROFESSIONAL EXPERIENCE:

2021-2022	Predoctoral Intern (APA-accredited) Tennessee Internship Consortium for Psychology/ Cherokee Health Systems, Knoxville, TN
2020-2021	Interning Therapist Marcus Autism Center, Atlanta, GA
2019-2021	School Psychology Ed.S. Intern Coweta County School System, Newnan, GA
2018-2019	School Psychology Practicum Student Coweta County School System, Newnan, GA
2017-2021	Graduate Research Assistant Center for Leadership in Disability at Georgia State University, Atlanta, GA
2017-2021	University Center for Excellence in Developmental Disabilities (UCEDD) Trainee Center for Leadership in Disability at Georgia State University, Atlanta, GA

PRESENTATIONS AND PUBLICATIONS:

Wellons, Q.D., Graybill, E., & Roach, A. T. (2022). Exploring social validity in the context of universal behavior screening in elementary schools. *Research and Practice in the Schools*, *9*(1), 1-20.

Wellons, Q. D., Roach, A., & Sanchez-Alvarez, S. (2022, February 15-18*). Is Social Validity an Afterthought in School Psychology Research?* [Paper presentation]. National Association of School Psychologists Annual Convention, Boston, MA, United States.

Wellons, Q. D., Roach, A., Sanchez-Alvarez, S, & McLendon, K. (2022, February 15-18). *The Impact of "ASK-READ-TELL" on Reading Comprehension* [Poster presentation]. National Association of School Psychologists Annual Convention, Boston, MA, United States.

Wellons, Q. D., Graybill, E., & Roach, A. (2021, February 23-26). *Exploring social validity of universal behavior screening in elementary schools* [Poster presentation]. National Association of School Psychologists Annual Convention, Virtual, United States.

Roach, A., Wellons, Q. D., Beacham, C., Crook, W., & Ducharme, D. (2021, February 23-26). *Still missing: Student voices in school psychology research 2010-2019* [Paper presentation]. National Association of School Psychologists Annual Convention, Virtual, United States.

Wellons, Q. D. (2020, March). *Overview and benefits of social emotional learning*. Eastside Elementary School, Senoia, GA, United States.

Wellons, Q. D., & Roach, A. (2020, February 18-21). *Monitoring student internet activity: Respecting privacy versus maintaining safety* [Paper presentation]. National Association of School Psychologists Annual Convention, Baltimore, MD, United States.

Wellons, Q. D., Floyd, H., Pimpalkhute, S., & Wilkins, S. (2020, January). *Threat assessment and school safety* [Paper presentation]. College of Education at Georgia State University, Atlanta, GA, United States.

Byars, J., Graybill, E., Wellons, Q., & Harper, L. (2020). Monitoring social media and technology use to prevent youth suicide and school violence. *Contemporary School Psychology*, *24*(3), 318-326.

Ducharme, D., Roach, A., Wellons, Q.D. (2020). The role of school psychologists in employment-focused transition services. *Journal of Applied School Psychology*.

Wellons, Q. D., Floyd, H., Pimpalkhute, S., Wilkins, S., & Varjas, K. (2019, February 26-March 1). *School shootings: Legal & ethical implications for school psychologists* [Paper presentation]. National Association of School Psychologists Annual Convention, Atlanta, GA, United States.

PROFESSIONAL SOCIETIES AND ORGANIZATIONS:

2021-Present	American Psychological Association, Division 16
2017-Present	National Association of School Psychologists
2018-Present	Georgia Association of School Psychologists
2017-Present	Student Affiliates in School Psychology Georgia State University Chapter

IMPROVING READING COMPREHENSION: AN INVESTIGATION OF TWO INTERVENTION PROCEDURES AND THE SOCIAL VALIDITY OF THE "ASK-READ-TELL" READING STRATEGY

by

QUYNH-NHU DOAN WELLONS

Under the Direction of Dr. Andrew T. Roach

ABSTRACT

The first chapter of this dissertation reports the results of a review of the practice and

application of social validity assessments in SCD studies within the field of school psychology.

Adopting similar procedures described in the Snodgrass et al. (2018) analysis, this review

systematically investigated the extent to which social validity was evaluated in school

psychology SCD studies, the characteristics of social validity assessments, the prevalence of

assessing total construct social validity (i.e., goals, procedures, and outcomes), and the extent to

which social validity findings were integrated with visual analyses to form conclusions about

interventions. Although the content analysis revealed several encouraging aspects of the current

state of social validity assessment, the majority of the findings suggested that there was still

much to be done to advance the rigor of social validity practice.

The second chapter reports the findings of a mixed-method study that utilized an

alternating treatments design to investigate the effects of McCallum's (2011) Ask-Read-Tell

(ART) procedure on the reading comprehension of three fourth-graders with a learning disability

in reading. Three conditions were used as the independent variable: (1) Control (baseline)

condition in which students did not use any prescribed strategies while reading, (2) ART

Condition: Students used the three-step cognitive strategy, and (3) ART + PD (Peer Discussion)

Condition: Students used the three-step ART sequence followed by a discussion of text with a

peer. Dependent variables included students' reading comprehension level (%C) and rate

(%C/M) as measured by students' comprehension performances on 400-word expository

passages. Both interventions had positive effects on reading performance as compared to the

baseline for all three participants. Results indicated that ART+PD led to the greatest gains in

rate and level for two participants, while ART resulted in greater gains for the third participant who presented with significant inattention and distractibility.

Finally, quantitative KIP and qualitative social validity data indicated that participants perceived the interventions as enjoyable, effective in improving their understanding of text, beneficial when doing schoolwork, and relevant in helping them achieve reading goals. By situating students as the primary respondents, the researcher hoped to underscore the importance of considering students' voices in determining the social validation of intervention effects.

IMPROVING READING COMPREHENSION: AN INVESTIGATION OF TWO

INTERVENTION PROCEDURES AND THE SOCIAL VALIDITY OF THE "ASK-READ-

TELL" READING STRATEGY

by

QUYNH-NHU DOAN WELLONS

A Dissertation

Presented in Partial Fulfillment of Requirements for the

Degree of

Doctor of Philosophy

in

School Psychology

in

Department of Counseling and Psychological Services

in

the College of Education & Human Development

Georgia State University

Atlanta, GA
2022

DEDICATION

I dedicate this body of work in memory of my beloved father. He exemplified courage, integrity, and grit. Pursuing this dream has been one of the biggest challenges of my life, and the belief that he still watches over me has given me strength along the way.

My appreciation goes to my extended family and long-time friends who have encouraged and cheered me on.

My twins – Daniel and Sovi – are the source of my joy and devotion. They inspire me to give my best every day in my work and in life.

Finally, my deepest gratitude to my incredible husband, Eric. It was Eric who propelled me to take this leap and to stay the course when things got hard. He is a wellspring of love, encouragement, friendship, and wisdom. This predoctoral internship year has displaced me from our home, and there are no words to adequately thank him for his sacrifice and endless patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 IS SOCIAL VALIDITY AN AFTERTHOUGHT IN SINGLE-CASE DESIGN STUDIES IN SCHOOL PSYCHOLOGY RESEARCH?

Wolf (1978) first introduced the concept of social validity to the field of applied behavior analysis. Since its introduction, the construct of social validity has been extended to school-based research and practices to demonstrate participants' attitudes toward new programs and interventions (Eckert & Hintze, 2000; Greer et al., 2012; Gresham & Lopez, 1996). Over the years, the term "social validity" has been referred interchangeably with *acceptance*, *satisfaction*, *cultural validity*, and other terminology (Carter, 2010). For the purpose of this study, social validity refers to consumers' perceptions surrounding the goals, procedures, and outcomes of a program or practice (Ledford & Gast, 2018; Schwartz & Baer, 1991). In their seminal papers, Kazdin (1977) and Wolf (1978) described the importance of assessing the degree of impact intervention efforts have on consumers and called for researchers to attend to this critical element of their work. Since that time, the evaluation of social validity has become an integral aspect of applied behavior analysis (Carr et al., 1999; Common & Lane, 2017). As such, Baer and colleagues (1987) posited that social validity should be incorporated as a secondary measure of effectiveness in behavioral methodology. In keeping with this position, social validity was included as one of the quality indicators for single-case design (SCD) research in special education (Horner et al., 2005). Further, comprehensive evaluation of social validity in SCD research requires diligent considerations of the social importance of outcomes, the feasibility and effectiveness of the interventions, and the maintenance of demonstrated effect (Horner et al., 2005).

The purpose of a social validity evaluation is to determine the acceptability of a program or treatment that encompasses the consideration of its goals, procedures, and outcomes. This component of intervention research is critical when attempting to implement research-based interventions on a broader scale (Cook et al., 2013). In order to enhance the viability of an intervention, researchers must anticipate aspects of the intervention that may not appeal to practitioners and consumers at large (Schwartz & Baer, 1991). To achieve this aim, social validity information should be solicited from a variety of stakeholders who will offer diverse perspectives about the intervention (Schwartz & Baer, 1991). Once collected, social validity data may be utilized strategically to inform current implementation procedures, program evaluation, and future program planning (Schwartz & Baer, 1991). Using social validity assessment in this way assists researchers in addressing potential limitations and shortcomings of an intervention procedures, materials, and goals (Schwartz & Baer, 1991; Snodgrass et al., 2018). As such, allowing social validity data to shape the development of an intervention may be a practical approach for promoting real-world application and sustainability (Schwartz & Baer, 1991).

Another potential benefit of assessing social validity is fostering consumers' buy-in of a program or practice. Soliciting consumers' feedback on proposed programs or newly developed interventions provides the intended audience a sense of shared control (Schwartz & Baer, 1991). Such practice not only promotes greater community buy-in but also provides the opportunity for researchers to bridge the research-to-practice gap (Snodgrass et al., 2018).

Social Validity as a Total Construct

Wolf (1978) recommended that the social validity construct be assessed along three dimensions: goals, procedures, and outcomes. Thus, the *total construct of social validity* refers

to a framework that encompasses all three facets (Gast & Ledford, 2018; Leko, 2014; Snodgrass et al., 2018). Regarding the social validation of intervention goals, researchers and practitioners should consider whether the established goals of the proposed treatment are valued by the participants themselves and the people with whom they interact (Wolf, 1978). From an ethical standpoint, researchers are obligated to consider whether teaching participants a specific skill or behavior will result in any meaningful impact on their quality of life (Gast & Ledford, 2018). Needless to say, some behavioral goals (e.g., improved interpersonal skills) may have universal social significance, while support for other goals (e.g., teaching sexual health concepts to students with severe disabilities) may be less clear and warrant reflective considerations as part of the process (Courtade et al., 2012). Such reflective practices can inform ethical decision-marking concerning the relevance of behavioral goals and the impetus behind intervention research and applications. Nonetheless, it behooves researchers to consider whether participants are likely to benefit from attaining particular goals or if the goals simply serve to answer the researcher's questions of interest (Gast & Ledford, 2018).

The second facet of social validity pertains to the appropriateness of treatment or program procedures (Wolf, 1978). This concept is sometimes referred as feasibility, defined as the potential for consumers and stakeholders to incorporate an intervention or program into their existing set of tools and practices (Caldarella et al., n.d.; Greer et al., 2012). Feasibility, in this instance, may be viewed favorably if barriers such as cost, time, efficiency, and training associated with the intervention are perceived to be manageable (Caldarella et al., n.d.; Greer et al., 2012). Equally important considerations are the administration time, compatibility within environmental contexts, and alignment with ethical standards (Glover & Albers, 2007; Ledford & Gast, 2018). As indicated across multiple research studies, procedures deemed socially

acceptable are more likely to be implemented with fidelity (Perpletchikova & Kazdin, 2005).

Finally, the third dimension of social validity relates to the social significance of intervention outcomes (Wolf, 1978). This aspect speaks to consumers' perceptions regarding the effectiveness of an intervention and their satisfaction with its results (Ledford & Gast, 2018). In other words, researchers and practitioners may profit from learning whether participants perceive the intervention as helpful and whether the outcomes achieved are impactful and meaningful to them (Ledford & Gast, 2018). To this end, early research has demonstrated a positive relationship between consumer satisfaction and measures of perceived effectiveness (Braukmann et al., 1976). To further support this finding, Eckert et al. (2017) found that students were likely to rate an intervention acceptable if they perceived the intervention as successful in improving their academic skills.

Social Validity Assessment Methods

Beginning with Kazdin (1977) and Wolf (1978), concerted efforts to diversify methods used in social validity assessment have been explored by single-case researchers (Carr et al., 1999; Common & Lane, 2017; Ledford & Gast, 2018). Although social validity can be measured in a variety of ways, five common approaches have been proposed: (a) subjective evaluation, (b) normative comparison, (c) maintenance, (d) blind ratings, and (e) participant preference (Common & Lane, 2017; Ledford & Gast, 2018; and Snodgrass et al., 2018).

A subjective evaluation of social validity involves garnering individuals' perceptions of particular dimensions of the goals, procedures, and outcomes of the independent variables in single-case designs (Ledford & Gast, 2018; Snodgrass et al., 2018). Kazdin (1977) showcased the importance of using subjective evaluation to validate treatment efficacy as well as to critique

the qualitative aspects of behavioral outcomes. Traditionally, subjective measures include rating scales, surveys, and questionnaires completed by intervention recipients or other parties (Finn & Sladeczek, 2001; Ledford & Gast, 2028). One potential strategy for assessing social validity is conducting semi-structured interviews with direct consumers (e.g., students receiving the intervention), indirect consumers (e.g., teachers, parents), and other stakeholders (e.g., administrators) (Common & Lane, 2017; Ledford & Gast, 2018).   Despite research demonstrating that interventions that were perceived favorably by students were more likely to increase their self-efficacy and motivation (Schunk,1996; Stipek, 1996), soliciting students' opinions of acceptability is less common than assessing the opinions of indirect consumers and other stakeholders in school-based, single-case research (Hurley, 2012; Linton, 1998).

A second approach for evaluating social validity is the use of normative comparisons. This process involves direct observations of the participants' targeted behavior to those of a reference group whose behaviors is considered "typical" or acceptable (Ledford & Gast, 2018). This method can be beneficial in determining the social importance of intervention goals and gauging whether participants have attained acceptable levels of the targeted behaviors (Ledford & Gast, 2018). Nonetheless, normative comparisons are not without limitations in that the reference group and the environment may not necessarily reflect optimal levels of performance across a wider set of contexts. For example, a normative criterion on reading achievement from a low-performing school may not align with the expected achievement standards at a state or national level for particular normative groups (Common & Lane, 2017).

Maintenance is a third approach to assessment that refers to the continued adoption and implementation of a program or practice following the completion of a research study (Kennedy,

2005). Understanding aspects of maintenance may provide further insight into developing successful interventions with sustained appeal (Kennedy, 2002). To this end, Kennedy (2002) called for "the integrative use of subjective evaluation and normative comparison" (p. 599) in measuring the extent to which behavior changes are sustained over time. Data on intervention maintenance can arguably strengthen researchers' claims of the social validity of an intervention (Kennedy, 2002).

One social validity method considered less subject to bias is the use of "blind" (third-party) ratings (Ledford & Gast, 2018). Individuals who are unaware of treatment conditions under investigation may be recruited to objectively determine intervention effectiveness, the social importance of these effects, and the social acceptability of implementation procedures (Common & Lane, 2017; Ledford & Gast, 2018).

Ultimately, participant (i.e., consumer) preference assessment may be the most relevant measure in determining direct consumers' perceptions of acceptability regarding treatment procedures (Ledford & Gast, 2018). Based on the existing literature regarding social validity assessment practices, preference assessment typically occurs via rating scales and after the completion of the intervention (Silva et al., 2019; Snodgrass et al., 2018). An important consideration is to solicit participants' feedback during intervention implementation in order to make modifications and refine procedures to best respond to participants' needs in real time (Hanley, 2010; Schwartz & Baer, 1991). In this manner, the information from preference assessments may directly inform changes in intervention development and support sustainable implementation thereafter (Schwartz & Baer, 1991).

Establishing Rigor in Social Validity Assessment

Since the original assertion of social validity's importance, the prevalence of social validity assessments in intervention research has increased significantly (Silva et al., 2019; Snodgrass et al., 2018; Villareal et al., 2015). In response to this trend, experts within the field have explored multiple ways to strengthen the scientific rigor of measuring this complex construct (Leko, 2014; Snodgrass et al., 2018). For example, Schwartz and Baer (1991) first called for specific changes in the existing practice of social validity assessment. They concluded that the majority of social validity assessment tools did not fully meet technical criteria, and subsequently appealed for increased psychometric rigor and attention to validity and reliability in social validity assessment (Schwartz & Baer, 1991). Today, a broad array of methods and rating scales for assessing perceived acceptability by specific target groups (e.g., children, parents, teachers) are easily accessible (Common & Lane, 2014; Finn & Sladeczek, 2001). Despite the plethora of methods and tools, social validity is typically measured via the administration of self-report questionnaires comprised of general ratings of acceptability (Finn & Sladeczek, 2001; Silva et al., 2019; Villarreal et al., 2015). In fact, Silva and colleagues (2019) reported that over 98% of the reviewed school psychology intervention studies used a self-report measure. Further, they found that studies were as likely to include researcher-developed measures as they were to use formally validated social validity measures (Silva et al., 2019). Similarly, in a separate review by Snodgrass and colleagues (2018), the majority of questionnaires used in the single-case studies reviewed were not validated instruments. Despite the call for increased attention to validity and reliability in assessment instruments, Villarreal et al.'s (2019) findings suggested that over 59% of the studies did not include psychometric data for social validity assessments.

In 2005, Horner and colleagues stipulated social validity as one of the seven qualitative indicators of single-case design (SCD) studies. They recommended that consideration of social

validity include the social importance of the dependent variable as well as the magnitude of change in that dependent variable. Further, they contended researchers should evaluate the practicality and cost effectiveness of implementation procedures within the context of their SCD research. Such specifications echo Wolf's (1978) guidelines in the social validation of goals, procedures, and outcomes as the total construct (Snodgrass et al., 2018). More recently, a synthesis of social validity in SCD research reported that of the studies that included a social validity assessment, only 24.3% conducted evaluations of total construct of social validity as defined by Wolf (Snodgrass et al., 2018).

Schwartz and Baer (1991) asserted that social validation assessments ought to be conducted prospectively, during intervention, and following the completion of intervention procedures. This position aligns with the concept of social validation as a dynamic process that occurs throughout the intervention process. Considering that intervention incompatibility is a common barrier to implementation, utilizing social validity information to shape intervention procedures during the process may be more beneficial than reflecting on it at the closing phase of the SCD studies (Long et al., 2016). Further, assessing participants' perceptions of intervention goals is especially beneficial when conducted before and during the delivery of intervention. When done in such a way, researchers and practitioners may potentially avert subjecting participants to interventions with goals of little value or meaning to them (Snodgrass at el., 2018). Despite such compelling reasons for conducting social validity evaluations across the intervention process, the majority of studies under review measured social validity only once and typically during the post-intervention phase (Silva et al., 2019; Snodgrass et al., 2018).

Barriers to Social Validity Assessment

In light of efforts to improve the existing practice of social validity assessment, substantial challenges continue to persist. First, social validity is a multifaceted construct subject to the influence of various factors within any given context (Carr et al., 1999; Snodgrass et al., 2018). For example, Schwartz and Baer (1991) asserted that multiple stakeholders should participate in the assessment of social validity for a particular intervention. Although the diversity in perspectives is preferred in many instances, having discordant perceptions can add another layer of complexity that may muddle social validity interpretation. For example, a social validity assessment of one intervention may yield dissimilar sets of feedback due to the unique nature of stakeholders' roles and their experiences in relation to the intervention. Hence, direct consumers may perceive the goals, procedures, and outcomes differently than members of the extended community (Ledford & Gast, 2018). In such case, how do researchers integrate divergent findings from social validity assessments with visual analyses of dependent variables when making conclusions about intervention effectiveness?

Another reason for the low prevalence of rigorous social validity assessment may be the absence of standardized methodologies for evaluating this construct (Snodgrass et al., 2018). Although Horner and colleagues (2005) included social validity as a quality indicator for SCD research, there exist no qualitative indicators that speak to the necessary technical characteristics of assessment tools, the timing of social validity assessment activities (i.e., before, during, and after intervention implementation), the evaluation of the total construct, and the application of social validity data in determining intervention effectiveness (Snodgrass et al., 2018). Without systematic, standardized guidelines, it is easy to see how social validity evaluations may be conducted in rudimentary ways or as an afterthought in SCD experiments (Kennedy, 2002; Snodgrass et al., 2018)

Despite an array of assessment options for measuring social validity (e.g., qualitative inquiry, normative comparison), the majority of SCD studies rely on self-report rating scales or surveys (Silva et al., 2019; Villarreal et al., 2015). This type of data collection method may have its utility; however, self-report surveys and rating scales tend to be restricted in the quality and type of information they offer (Leko, 2014; Snodgrass et al., 2018). Further, the developer of close-ended assessment tools becomes the arbiter of what social validity information is considered relevant for consideration. In essence, the use of surveys or rating scales may potentially eliminate the opportunity to fully capture participants' experience and perspectives about a particular intervention since most surveys contain dichotomous (i.e., yes/no) or ordinal (i.e., Likert-type scales) items (Leko, 2014). As such, potential information deemed important by participants may be underrepresented or excluded altogether.

Finally, the paucity of social validity rigor may be attributed to logistical challenges such as the additional cost and time required to conduct thorough assessments (Leko, 2014). Further, publishing conventions may dictate page limits, often precluding a comprehensive report of both visual analysis and social validity evaluation in the same manuscript (Carr et al.,1999).

The Benefits of SCD

Single-case design methodology is highly valued in the educational setting for many reasons. One advantage is the relative ease in execution as SCD studies do not require the use of control groups or randomization of participants (Ledford & Gast, 2018; Riley-Tillman & Burns, 2009). Due to the nature of schools, it is not always feasible to randomly assign students into each condition (Johnson & Christensen, 2017). Further, the use of a control group necessitates withholding interventions from students which raises the question of ethical concerns, especially when students in the control group also exhibit similar levels of difficulties as those in the

intervention groups (Riley-Tillman & Burns, 2009). As such, SCD methodology eliminates this ethical dilemma given that study participants act as their own experimental control (Ledford & Gast, 2018; Radley et al., 2020). In addition, SCD methodology allows for detecting differences in dependent variables across participants as well as within participants at the individual level (Horner et al., 2005; Ledford & Gast, 2018; Riley-Tillman & Burns, 2009). By analyzing findings at this level, researchers are able to generalize findings to specific individuals rather than across populations (Ledford & Gast, 2018). Within the context of monitoring a student's response to an intervention, an evidence-based intervention may not necessarily prove effective for every child (Burns et al., 2017). Thus, understanding intervention effectiveness for specific students warrants the type of investigation that focuses on intra-participant behaviors and responses (Horner et al., 2005; Ledford & Gast, 2018; Riley-Tilman & Burns, 2009). This aspect of SCD is especially ideal for educational researchers as many school-based practitioners work with students at the one-on-one or small-group level (e.g., school psychologists, special educators, counselors, and speech pathologists) (Radley et al., 2020; Riley-Tillman & Burns, 2009).

Trends and the State of Social Validity Assessment in School Psychology Research

Within the school psychology literature, single-case methodology has gained popularity for the reasons described earlier (Radley et al., 2020). For example, SCDs were commonly found in intervention studies published in school psychology journals from 2010 to 2014 (Villareal et al., 2017); 56.7% of studies by prominent school psychology intervention researchers used single-case or related designs (Villareal & Umana, 2017). In another systematic synthesis of publications in 13 journals focused on school psychology, Radley and colleagues (2020) found that between 1968 to 1994, the proportion of SCD articles did not exceed 20%

(Radley et al., 2020). However, a sharp rise in the proportion of SCD articles occurred in 1995, when over 30% of SCD articles reviewed were published (Radley et al., 2020).

Since Kazdin's and Wolf's introduction of the social validity construct, a major shift in how researchers view the importance of social validity assessment has transpired across behavioral and academic intervention lines of research (Finney, 1991). Given the rise in SCD studies in school psychology intervention research, it is commonly accepted that consumers' perceptions can have a significant influence on the evaluation of an intervention. Within school psychology, scholars have conducted systematic reviews to establish the state of social validity assessments. Roach and colleagues (2009) conducted a content analysis of articles from four major school psychology journals from 2002 to 2007 to determine the inclusion of students' perspectives in published work. Despite Wolf's (1978) assertation that direct consumers are the only legitimate individuals who can speak about an intervention, Roach and colleagues (2009) indicated that only 16.2% of the studies solicited students' opinions of acceptability. Even less common was assessing the opinions of students with disabilities in that only 2.7% of the reviewed publications reported on students' perceptions of acceptability (Roach et al., 2009).

In 2015, Villarreal and colleagues investigated the prevalence of reported acceptability data in six prominent school psychology journals from 2005 to 2014. In this review, the intervention studies comprised of case studies, experimental, quasi-experimental, and SCD studies. Of the 243 intervention studies, 30.5% included quantitative acceptability data that were provided mainly by teachers or other school personnel. The authors also found that more researcher-developed or adapted versions of a formal social validity instrument were utilized than any other type of assessment methods, while in a disconcerting 63.8% of the studies social validity was not monitored nor assessed in any way (Villarreal et al., 2015).

Silva et al. (2019) reviewed published works within the intervention literature across five school psychology journals between 2005 and 2017. Both group and single-case designs accounted for 268 intervention studies. Their purpose was to provide a current view of social validity assessment and its inclusion in school psychology research (Silva et al., 2019). Consistent with previous reviews, slightly more than one third of the intervention studies ($n =$ 108) included social validity data, derived primarily from self-report measures (Silva et al., 2019). Of these 108 studies, 68.5% of the assessments occurred post intervention (Silva et al., 2019). Moreover, approximately 74.1% of the studies solicited teachers' input on acceptability even though teachers were the implementers in less than half of the studies (Silva et al., 2019). In contrast to reviews by Roach et al. (2009) and Villarreal et al. (2015), Silva and colleagues found more than half of the studies garnered students' perceptions of acceptability. Still, assuming that interventions studies were targeted for student populations, it is disheartening that opportunities for students to express their views as direct consumers remained lower than that of teachers or other school personnel.

In the schools, adults often make decisions that impact students in multiple ways; hence, inviting children and adolescents to communicate their views through social validity evaluations is a meaningful way through which students' voices can be heard (Roach et al., 2009). Nonetheless, the current literature seems to suggest that social validity data continue to be underreported (Roach et al., 2009; Silva et al., 2019; Villarreal et al., 2015) and that social validity remains an understudied construct (Snodgrass et al., 2018).

Purpose of Systematic Review

Findings from Villarreal et al., 2015 and Silva et al., 2019 provided a general understanding of acceptability assessment practice in school psychology intervention research.

However, these findings did not elucidate our understanding of the state of social validity assessments in SCD studies specifically. To this end, we looked to the synthesis by Snodgrass et al., (2018) for a comprehensive view of social validity assessment practices in SCD research published in six highly-ranked special education journals from 2005 to 2016. Further, Snodgrass and colleagues' review systematically investigated the scientific rigor of the procedures used to evaluate social validity assessments across a variety of topics and populations (Snodgrass et al., 2018). As such, their systematic review provided a compelling framework for evaluating practices in social validity assessments, and our current study was loosely based on the general structure for conducting content analysis.

Although attending to the social validity of interventions and involving students in the essential components of intervention planning are considered professional responsibilities of school-based practitioners (Burns et al., 2017; NASP, 2020), the current literature suggests there is more to be done if we are interested in advancing the rigor of social validity assessment practices. As such, the current study was intended to contribute to the existing social validity literature by extending on previous published studies (Silva et al., 2019; Villarreal et al., 2015). The primary aim was to examine the practice and application of social validity assessments specifically in SCD studies within the field of school psychology. Employing adapted procedures from the Snodgrass et al. (2018) analysis, this review systematically explored the scientific rigor of social validity measurement methods, analyses, and data integration in published SCD studies across six school psychology journals, a perspective not undertaken in prior studies. Since reviewed publications from Silva et al., 2019 and Villarreal et al., 2015 were from school psychology journals, it was pertinent that articles selected for this content analysis were within the school psychology literature to allow for comparisons of findings. Further, to

maintain the scope of this analysis within a manageable size, it was determined that only articles

from school psychology journals would be selected.

Research Questions

The current systematic review seeks to answer the following research questions:

1. To what extent is social validity evaluated in school psychology SCD studies?

2. What are the characteristics of social validity assessments within SCD studies?

3. How frequently do researchers evaluate total construct social validity (i.e., goals, procedures, and outcomes) compared to partial construct (i.e., fewer than three dimensions)?

4. To what extent are social validity findings integrated with visual analyses to form conclusions about interventions?

Method

Six prominent school psychology journals were selected for inclusion in this systematic

review: *School Psychology Review*, *School Psychology Quarterly*, *Journal of School Psychology*,

Psychology in the Schools, *School Psychology International*, and *Journal of Applied School*

Psychology. All volumes and issues of the selected journals published from 2016 through 2020

were included. It was determined that the search period would start with journals from 2016

because the combined time periods in previous reviews (Silva et al., 2019; Villarreal et al., 2015)

extended from 2005 through 2017, and the intent of the current review was to minimize any

overlap in search periods. The researchers retrieved electronic copies of the journal publications

via university library databases.

Phase 1

For the first phase of the analysis (Figure 1.1), all issues of the six journals were hand-

searched by the primary author. A total of 1,371 articles were reviewed. The aim of this initial part of Phase 1 was to identify SCD studies. In general, What Works Clearinghouse (WWC) identified SCDs by the following features: (a) when an individual case was the unit of intervention and the unit of analysis. Cases might consist of a single participant or a cluster of participants (e.g., a school), (b) within the design, the case acted as its own control for purposes of comparison, and (c) the dependent variable was measured repeatedly within and across different conditions or levels of the independent variable, referred as phases (e.g., baseline phase, intervention phase) (Kratochwill et al., 2010). Search criteria for the identification of SCD studies were based on WWC standards (Kratochwill et al., 2010). The author reviewed titles and abstracts to look for key terms such as single-case design, single-subject design, multiple baseline, multiple probe, reversal design, multi-treatment design, alternating treatments, adapted alternating treatments, and changing criterion. Further, single-case design studies characteristically employ a small number of participants; thus, this indicator was also used to identify SCD studies. The search was not limited to a particular number of participants. All articles that included a single-case design were selected. Of the 1,371 articles, 80 were identified as SCD studies.

To address Research Questions 1 and 2, the 80 SCD articles were analyzed to identify those reporting a social validity assessment of any kind. Key terms such as social validity, social validation, acceptability, satisfaction, satisfaction survey, questionnaire, rating scale, interview, focused group, participant perceptions, and perspectives were used as potential indicators of a social validity assessment. Additionally, the author examined the methods and results to determine whether social validity assessments were conducted. Articles that only included anecdotal descriptions of social importance without social validity assessment data were

excluded. Anecdotal descriptions were those in which researchers presented informal

information without the use of any assessment methods nor direct evaluation of social validity.

From this process, 61 SCD publications were identified to include social validity assessments.

Because many of the publications reported on multiple assessments within a study, the analysis

resulted in 89 social validity assessments.

Next, an in-depth content review determined the characteristics of the 89 assessments

with regard to method, type of measure, respondents, temporal context (i.e., preintervention,

during intervention, postintervention), and type of reported data.

Phase 2

To answer Research Question 3, the second phase (Figure 1.1) involved differentiating

studies that conducted an evaluation of the total construct of social validity from those assessing

a partial construct. A total construct social validity assessment was defined as an effort to

evaluate all three dimensions of social validity: goals, procedures, and outcomes. To qualify as a

total construct assessment, the study needed to address all three factors in its assessment method.

A partial construct social validity assessment was identified when fewer than three dimensions

were assessed.

Phase 3

An in-depth analysis in the final phase (Figure 1.1) addressed Research Question 4.

Social validity data were coded for inclusion of data analysis methods. In this phase, data

analysis was defined as a description of how qualitative and/or quantitative social validity data

were obtained and what methods were used in the analysis. Finally, social validity results were

coded on how they were represented in relation to the visual analysis findings when making

claims about intervention effectiveness in the discussion. The degree of representation was

coded as: (a) No reference to social validity results, (b) Separate reference of social validity results from visual analysis results, (c) Vague reference in which data sets were mentioned without discussion of how they were related, and (d) Integrated reference of both social validity and visual analysis results.

Interrater Agreement

A graduate-level researcher trained in analysis procedures for all phases, assisted in establishing interrater agreement (IRA). IRA was calculated by dividing the number of agreements by the sum of agreements and disagreements, multiplied by 100.

The primary and assistant researchers independently coded 20.0% ($n = 275$) of the 1,371 articles, selected at random and achieved an IRA of 100% in the identification part of Phase 1. Thereafter, the researchers independently coded 25.0% ($n = 20$) of 80 SCD articles with social validity assessments. An IRA of 95% was achieved for this part of Phase 1.

In Phase 2, the two researchers independently coded 29.2% ($n = 26$) of the 89 social validity assessments and achieved an IRA of 96.15%.

In Phase 3, 29.2% ($n = 26$) of 89 assessments were coded independently, and an IRA of 92.31% was derived.

Results

The review evaluated the application of social validity assessment in single-case design studies across six school psychology journals in three phases. In Phase 1, we determined the extent to which social validity was measured by reporting on the number of SCD articles with social validity assessment and the number of total assessments from these publications. A second purpose of Phase 1 was to report on the characteristics of social validity assessments. Phase 2 was devoted to determining the types of constructs measured across all assessments.

Finally, we determined the extent to which social validity findings were integrated with visual analyses in Phase 3.

Phase 1

We screened a total of 1,371 articles across the six journals (Table 1.1). From this initial screening, 80 articles were single-case design studies. Of these 80 SCD studies, 61 (76.25%) included some form of social validity assessment. *Psychology in the Schools* (PITS) published the highest number of articles ($n = 471$) overall, and it accounted for the majority of SCD articles with social validity assessments ($n = 20$). However, *Journal of Applied School Psychology* (JASP) had the highest percentage of SCD studies with social validity assessments (88.89%) as 16 of their 18 SCD articles included some form of social validity evaluation.

Single-case research publications distributed across the five-year period were also reviewed (Figure 1.2), with the largest number published in 2016 and 2018 ($n = 19$ each year). However, there appeared to be a decline from 2019 ($n = 13$) to 2020 ($n = 11$). When we examined the trend for inclusion of social validity within these publications, there appeared to be a steady decline across the five-year span. However, of the SCD studies, the percentage of those with social validity assessments increased from 57.89% (2018), to 92.31% (2019), to 100.00% (2020).

Across the six journals, 61 articles yielded a total of 89 social validity assessments. Due to variability in assessment methods, measures, temporal context, and types of respondents within and across publications, it was necessary to present results at the assessment level. Many of the articles included more than one assessment, and the assessments were not necessarily administered to the same level of depth as others within the same article. For this reason, it was necessary to analyze at the assessment level. Additionally, some articles included only multiple

rating scales, while others included rating scales and interviews in the assessments. Thus, if results were presented at the article level, we would be excluding a large quantity of relevant information.

Social validity assessment characteristics are presented in Table 1.2. When evaluating social validity, researchers utilized self-report measures most frequently ($n = 86$; 96.6%). Two of the assessments were completed via interviews (2.25%), and one was administered via an immediate verbal question-and-answer format (1.12%). The majority of the self-report measures were published and previously validated ($n = 55$; 61.80%), followed by researcher-developed ($n = 31$; 34.83%), and other formats ($n = 3$; 3.37%). Teacher acceptability was reported most frequently ($n = 51$, 57.30%), followed by students ($n = 29$; 32.58%), and parents ($n = 9$; 10.11%). More than half of the assessments ($n = 47$; 52.81%) occurred after intervention completion or at the conclusion of the study. For 29 of the assessments (32.58%), researchers did not report on the temporal context of the evaluation. When results were reported, the majority ($n = 68$; 76.40%) included both quantitative and qualitative data, followed by instances where only qualitative data ($n = 15$; 16.85%) was provided.

Phase 2

In Phase 2 of the review, we determined to what extent SCD articles that included social validity assessment ($n = 61$; 76.25% of SCD articles) evaluated the total construct (i.e., goals, procedures, and outcomes) compared to the partial construct (i.e., fewer than three dimensions). Figure 1.3 presents types of social validity construct assessed by journal. Of the 61 articles, 35 were found to consider partial construct with JASP having the most articles ($n = 11$) in this group, followed by PITS ($n = 10$), and SPR ($n = 9$). With regard to total construct, 13 articles

were identified with PITS having the most articles ($n = 7$) in this group, followed by JASP and

SP(Q) ($n = 2$ each). A total of 13 publications evaluated for both partial and total social validity

constructs; SPR having the most articles ($n = 4$) in this group, followed by JASP and PIT ($n = 3$

each).

Additionally, Phase 2 analysis included a breakdown of types of methods used to assess

these aspects of social validity. Of the 89 social validity assessments used across the 61 articles,

58 measured the partial social validity construct, while 31 were considered total-construct

measures. Researchers explicitly stated their intent to evaluate total construct in only four

instances. Figure 1.4 displays the breakdown for categories of constructs organized by

assessment measure. Among the 55 published and validated measures used to evaluate social

validity, 26 included rating scale items that queried on all three social validity dimensions (i.e.,

goal, procedure, outcome), and 29 included items that only queried on two or fewer dimensions.

The majority of published and validated measures of the total construct ($n = 8$) were reported in

PITS, followed by JASP and SP(Q) ($n = 6$ each). Most of the published and validated measures

of the partial construct ($n = 12$) were included in JASP publications, followed by SPR ($n = 8$).

Of the 31 researcher-developed measures used in the articles we reviewed, five included items

that assessed all three dimensions, and 26 assessed fewer than three dimensions. PITS had the

majority of articles that included researcher-developed assessments with three total construct and

11 partial construct measures that were developed the authors. As for assessments administered

via the interviewing method, we found that none of the three assessments included questions that

queried all three dimensions of social

validity.

Phase 3

In the final phase of the review, we determined how social validity results were represented in relation to experimental findings in the discussion. The degree of representation was coded as 1) Integrated – both sets of social validity and visual analysis data were discussed in relation to one another and how they were directly connected, 2) Referenced – one set of data was referenced with regard to the other without discussion of how they were connected, 3) Separate – social validity findings were represented separately from visual analysis results, and 4) No reference – social validity findings were not represented in discussion. Figure 1.5 displays Phase 3 results. Of the 89 assessments in the reviewed publications, the discussion of 11 measures (12.36%) was fully integrated with other collected data, 21 measures (23.60%) were referenced in discussion with other results, 44 measures (49.44%) were discussed separately from other results, and 13 (14.61%) not referenced in the discussion. In our review, SPR published the majority of the integrated data representations ($n = 4$). Further, JASP and SP(Q) had the highest number of referenced representations ($n = 6$ each). The majority of the separate data representations ($n = 15$) were published in PITS. Finally, PITS and SPR published the highest number of reports ($n = 5$ each) in which social validity data were not represented in the discussion.

Discussion

The purpose of this review was to systematically examine the scientific rigor of social validity measurement methods, analyses, and data representation in relation to experimental findings within SCD studies published in six school psychology journals between 2016 and 2020. This study contributes to the current body of social validity literature in several ways: (a) determining the extent to which social validity was evaluated in school psychology SCD research, (b) examining the characteristics of social validity assessments in SCD studies, (c)

comparing the frequency of total versus partial construct assessments of social validity, and (d)

evaluating the extent to which social validity results were integrated with visual analysis findings

in discussions.

Social Validity in School Psychology SCD Studies

This systematic search yielded 80 articles with single-case methodology. Of these

articles, 61 had some form of social validity assessments. Compared to the Silva et al., (2019)

study in which 40.30% of all intervention studies from 2005 to 2017 reported acceptability, this

review found that over 76.25% of SCD publications from 2016 to 2020 included social validity

assessments. This finding represented a notable increase in the frequency of evaluating for social

validity among single-case researchers. Moreover, researchers in several publications assessed

multiple respondents and/or utilized different social validity assessment methods within the same

study. For this reason, a total of 89 assessments were used across the 61 publications. The

majority used different types of self-report rating scales (e.g., BIRS and CIRP) to assess different

individuals' perceptions rather than assessing social validity through multiple modes (e.g., rating

scales and interviews). In fact, only two publications described utilizing a combination of rating

scales and interviews in their evaluations. One possible explanation for the paucity in

multimethod assessments might be due to logistical barriers such as limited funding and the

extended time required to conduct interviews or implement other assessment modes (e.g.,

normative comparisons).

In addition to investigating the extent to which social validity was evaluated, we were

interested in examining the trend of single-case methodology and inclusion of social validity in

school psychology single-case research. Between 2016 and 2020, there appeared to be a general

decline in the number of SCD studies among the six school psychology journals, and due to this

decrease, the number of social validity assessments in SCD publications also fell. This finding

contrasted with earlier reviews (Radley et al., 2020; Villarreal et al., 2017) in which single-case

methodology was shown to have gained popularity within school psychology research. When

comparing the prevalence of SCD publications among the six journals, *Journal of Applied*

School Psychology (JASP) published the highest percentage of SCD articles (20.22%), and

School Psychology International (SPI) published the lowest percentage of SCD articles (1.23%).

Since JASP aimed to publish scholarly material that contributed to applied school psychology

practice, its high percentage of SCD studies was expected since applied researchers may be more

likely to utilize single-case methodology (Ledford and Gast, 2018). SPI, on the other hand,

sought to promote research studies that addressed key issues and developments in school

psychology world-wide. As such, single-case methodology might have little utility when

conducting research with this broader scope and purpose. Despite the decline in number of SCD

studies, the practice of assessing for social validity appeared promising in that the percentages of

SCD studies with social validity increased during the five-year span. As seen in Figure 1.2, it

was evident that a sharp rise occurred in social validity assessment from 2018 (57.89%) to 2020

(100.00%).

Assessment Characteristics

Consistent with previous reviews (Silva et al., 2019; Villarreal et al., 2015), 96.63% of

assessments were administered via self-report rating scales. This finding suggests that

researchers' overreliance on self-report rating scales has not changed since the early 2000s

(Villarreal et al., 2015). Rating scales and surveys appeal to researchers for multiple reasons.

One advantage may be that the administration of rating scales requires less researcher and

participant time, which ultimately minimizes investigative time. Further, they are easy to learn

and convenient to administer. Finally, survey results can be obtained and analyzed relatively

quickly compared to qualitative data. Despite the feasibility of rating scales, these instruments

are not necessarily the most useful tool for evaluating social validity since they tend to be

restricted in the quantity and quality of information they produce (Leko, 2014; Snodgrass et al.,

2018). Because the majority of rating scales contain close-ended or Likert-type items, it is not

possible for researchers to fully capture participants' experiences and perspectives about

interventions of interest. As such, the overreliance of rating scales may potentially undermine

the purpose of social validity assessments since information deemed important by participants

may be omitted altogether. As such, we support Leko's (2014) call for incorporating qualitative

inquiry as a part of a comprehensive social validity evaluation.

Published and previously validated rating scales accounted for 61.80% of the measures in

our review, while researcher-developed measures comprised 34.83%. These findings suggest an

increase in the use of published, validated instruments, a departure from earlier reviews in which

researchers relied on more of their own instruments (Villarreal et al., 2015) or were just as likely

to use researcher-developed as published, validated instruments (Silva et al., 2019).

Consistent with findings from Silva et al. (2019) and Villarreal et al. (2015) teachers were

asked to report on social validity in the majority of the assessments (57.30%). Students

comprised more than one third of the respondents. Although this current analysis revealed that

student social validity was more frequently collected than in other reviews (Roach et al., 2009;

Villarreal et al., 2015), the percentage of student respondents (32.58%) paled in comparison to

the 59.26% from the Silva et al. (2019) analysis. It is not unexpected that student perceptions

were less frequently sought than those of teachers for several reasons. Evaluating teachers'

perceptions may be more feasible and less time intensive than soliciting perspectives from

groups of students. Additionally, other barriers such as limited language among younger students and some students with disabilities, limited understanding of assessment items/questions, and limited availability of developmentally appropriate instruments for children may deter researchers. Nevertheless, given that most intervention studies within the school psychology literature were designed to address students' academic or behavioral problems, it is especially important that student perspectives are considered. Schunk (1996) and Stipek (1996) demonstrated that interventions that were perceived favorably by students were more likely to increase their self-efficacy and motivation, a direct correlation to mastery and competence of skills. To this end, it is imperative that researchers and practitioners consider students' perspectives as they may shed light on factors that are critical to intervention development, effectiveness, and implementation.

Regarding the timing of social validity assessments, the analysis revealed that the majority (52.81%) of assessment occurred after intervention completion or at the conclusion of the study. This result is consistent with those from prior analyses (Silva et al., 2019). Given that direct consumers' perceptions of acceptability related to interventions are critical in directing changes in intervention development and promoting sustainable adoption of them thereafter, researchers and practitioners may consider conducting assessments throughout the intervention process. By assessing for acceptability prospectively, during intervention, and following completion of intervention procedures, researchers and practitioners may avoid subjecting participants to interventions with unacceptable goals and procedures.

Finally, it was found that over three-quarters of the assessments yielded quantitative data accompanied by qualitative descriptors to further elucidate the meaning and context of findings. This was especially helpful for result interpretation since ratings scale items varied in their

ordinal

values.

Prevalence of Total versus Partial Construct

We examined the extent to which SCD researchers assessed for total versus partial social validity construct. Total construct was coded when an assessment included a social validity instrument or method that contained inquiry items assessing for all three dimensions identified by Wolf (1978): goals, outcomes, and procedures. Likewise, partial construct was identified when there were inquiry items that assessed for two or fewer of these dimensions. Interestingly, the majority of researchers did not clearly state their intent to assess for any of these three factors of social validity. In fact, the language used to describe their assessments generally was focused on the concept of acceptability. Several authors referenced Wolf (1978) and his assessment specification for social validation of goals, procedures, and outcomes but consequently proceeded to evaluate with instruments that queried only one or two of the factors. Of the 89 assessments in the studies we reviewed, only four included explicit statements surrounding total construct evaluations paired with corresponding assessment procedures.

Our analysis of types of constructs by journal revealed that 35 of the 61 publications contained some form of partial construct assessments, 13 had total construct assessments, and 13 evaluated for both constructs. Our analysis indicated that the majority of the 89 assessments considered only partial constructs ($n = 58$). Further, the most frequently used social validity assessment was some form of published, validated rating scale ($n = 29$). It is not surprising that researchers relied on readily accessible instruments to measure social validity out of expediency. The dearth of available total-construct instruments and methods, lack of guidance and standards for conducting social validity assessments, and the additional cost and time to conduct

comprehensive evaluations may have contributed to these findings. Social validity was highlighted as one of the qualitative indicators of single-case research (Horner et al., 2005), but our review suggested social validity assessment practices in school psychology SCD research are not yet comprehensive and conducted with purposeful intention.

Data Integration

Although Horner and colleagues (2005) asserted that social validity assessment is integral to high-quality single-case research (SCR), to our knowledge, journal editors and peer reviewers have not required researchers to account for social validity findings when determining intervention effect. In the final phase of our review, we explored how findings from social validity assessments were combined with visual analysis results to draw conclusions about intervention effectiveness.

For 11 of the assessments (12.36%) reported in reviewed studies, authors integrated assessment results and discussed social validity and experimental findings in relation to one another. For almost half of the assessments reported, the authors ($n = 44$; 49.44%) represented their experimental and social validity findings separately in the discussion. For an additional 21 assessments (23.60%), the authors vaguely referenced social validity results with experimental results without discussing their relationship to one another. Finally, there were 13 social validity assessments (14.61%) for which the authors did not represent or refer to the resulting data in their discussion at all. As such, representation of social validity findings in these studies were confined strictly to the results. In summary, the majority of the authors uniformly relied on experimental results when making claims about intervention effectiveness.

Our analysis appeared to suggest that researchers generally value social validity as a required component of SCR, but they have yet to utilize these results in conjunction with

experimental effects to inform their conclusions about intervention effectiveness. This begs the question of why social validity assessments were conducted in the first place if their findings were ultimately disregarded. To remedy this persistent problem regarding social validity assessment, single-case researchers may consider utilizing mixed-methods approaches (Greene, 2007). Multiple methods may be ideal for enhancing the rigor of social validity assessment practice as it allows for more than one way to evaluate an intervention. For example, in a single mixed-method study, researchers can investigate the experimental effects of an intervention concurrently with its social validity. By employing both experimental and qualitative methods within a single-case design, researchers have the opportunity to demonstrate functional relation combined with a rigorous assessment of social validity in pursuit of identifying effective interventions.

Limitations and Future Directions

Although findings from this review extended the body of research on the practice of social validity assessment in school psychology SCD studies, this study was not without limitations. First, previous content analyses in school psychology have measured similar variables to those in this study; however, there are currently no standard guidelines for coding journal content. For this reason, it is expected that operationalization of variables might vary across research teams, thus comparisons to earlier content reviews should be interpreted with caution. Second, to maintain the scope of this review at a manageable size, the study included publications from only six school psychology journals. Since we did not examine other journals in related fields (e.g., special education, applied behavior analysis), we likely excluded a large quantity of SCD studies with social validity assessments. Third, several studies incorporated maintenance phases or generalization probes as potential indicators of social validity (Kennedy,

2002). However, these features within the SCD studies were excluded since they alone did not ensure social validity (Kennedy, 2002). Fourth, although we explored the level of integration of experimental and social validity findings when determining intervention effects, we did not examine the correlation between efficacy and social validity. Fifth, two major types of interventions (e.g., academic, social-emotional/behavioral) were noted, but we did not incorporate this component in the analyses by disaggregating results by intervention type or focus.

Future research could explore whether the practice of social validity assessments differed for different types of interventions. Other potential ways to expand on this study may include examining the state of social validity assessments for practices implemented at different grade levels. For example, it would benefit the research community to ascertain whether social validity assessment characteristics vary at the elementary, middle, and high-school level.

Implications for Research and Practice

This content analysis revealed a few encouraging aspects of the current state of social validity assessment. Nevertheless, the majority of our findings suggested that there was still a lot to be done to advance the rigor of social validity practice. Given the importance of social validity to successful implementation and viability of research-based interventions, researchers must attend to several elements. Since the early 2000s, researchers have relied heavily on self-report rating scales as their instrument of choice. To fully capture direct consumers' experiences and perspectives about an intervention, other methodological options (i.e., qualitative inquiry, observations) should be utilized to allow for a more comprehensive and meaningful social validity evaluation. One potential solution is the use of mixed-methods evaluation in order to ensure an in-depth social validity assessment.

Further, since most school-based interventions are developed to address student deficits, students are (in essence) the direct consumers. Although it is highly recommended that multiple stakeholders assess the social validity of an intervention (Schwarts & Baer, 1991), direct consumers are truly the ones to determine its acceptability (Wolf, 1978). To this end, practitioners and researchers have the professional responsibility (NASP, 2010) to provide opportunities for students to communicate their views about school-based practices that have the potential to directly impact them.

Additionally, when conducting social validation, goals, procedures, and outcomes should be at the forefront of any assessment. Although some instruments may not contain items that address all three dimensions, it is incumbent upon researchers to supplement them with additional tools or methods so that the total construct can be assessed to the fullest extent possible. Two final aspects to consider are the timing of assessments and the consideration of social validity findings when concluding about intervention effectiveness. In this respect, researchers should examine their purpose for conducting a social validity assessment. If the intent is to shape the development, implementation, and effectiveness of an intervention, then assessments should be conducted throughout the intervention process. The results from these assessments should be included as an integral part of the data set to drive researcher decision-making.

Conclusion

Since the early 2000s (Villarreal et al., 2015), the practice of social validity assessment has remained generally unchanged. It is not enough to conduct social validation because it is a required indicator of SCR (Horner et al., 2005). As intervention researchers, we collectively have the responsibility to thoroughly evaluate all aspects of an intervention, and to do so, we

must craft our studies to include thoughtful and purposeful efforts to assess social validity. To this end, researchers have the opportunity to advance social validation to the next level of rigor and to end the practice of relegating social validation to an afterthought in intervention research.

References

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*, 91-97.

Brauckmann, C. J., Kirigin, K. A., & Wolf, M. M. (1976, September 3-7). *Achievement place: The researcher's perspectives* [Paper presentation]. American Psychological Association 84th Annual Convention, Washington, DC, United States.

Burns, M. K., Riley-Tillman, T. C., Rathvon, N. (2017). *Effective school interventions: Evidence-based strategies for improving student outcomes* (3rd ed.). The Guildford Press.

Carr, J. E., Austin, J. L., Britton, L. N., Kellum, K. K., Bailey, J. S. (1999). An assessment of social validity trends in applied behavior analysis. *Behavioral Interventions*, *14*, 223-231.

Carter, S. L. (2010). Background of social validity. In S. L. Carter, *The social validity manual* (pp. 1-20). Elsevier.

Common, E. A., & Lane, K. L. (2017). Social validity assessment. In J. K. Luiselli (Ed.), *Applied behavior analysis advanced guidebook*: *A manual for professional practice*. (pp.73-92). Academic Press.

Cook, B. G., Cook, L., & Landrum, T. J. (2013). Moving research into practice: Can we make dissemination stick? *Exceptional Children*, *79*, 163-180.

Courtade, G., Spooner, F., Browder, D., & Jimenez, B. (2012). Seven reasons to promote standards-based instruction for students with severe disabilities: A reply to Ayres, Low ery, Douglas, & Sievers (2011). *Education and Training in Autism and Developmental Disabilities*, *47*, 3-13.

Eckert, T. L., Hier, B. O., Hamsho, N. F., & Malandrino, R. D. (2017). Assessing children's

    perceptions of academic interventions: The Kids Intervention Profile. *School Psychology*

    Quarterly, *32*(2), 268-281.

Eckert, T. L., & Hintze, J.M. (2000). Behavioral conceptions and applications of acceptability:

    Issues related to service delivery and research methodology. *School Psychology*

    Quarterly, *15*, 123-148.

Finn, C. A., & Sladeczek, I. E. (2001). Assessing the social validity of behavior interventions:

    A review of treatment acceptability measures. *School Psychology Quarterly*, *16*, 176-206.

Finney, J. W. (1991). On further development of the concept of social validity. *Journal of*

    Applied Behavior Analysis, *24*, 245-249.

Greene, J. C. (2007). *Mixed methods in social inquiry*. Jossey-Bass.

Greer, F. W., Wilson, B. S., & DiStefano, C., & Liu, J. (2012). Considering social validity in the

    context of emotional and behavioral screening. *NASP School Psychology Forum:*

    Research in Practice, *6*(4), 148-159.

Gresham, F. M., & Lopez, M.F. (1996). Social validation: A unifying construct for school-based

    consultation research and practice. *School Psychology Quarterly, 11*, 204-227.

Hanley, G. P. (2010). Toward effective and preferred programming: A case for the objective

    measurement of social validity with recipients of behavior-change programs. *Behavior*

    Analysis in Practice, *3*(1), 13-21.

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. https://doi.org/10.1177/001440290507100203

Hurley, J. J. (2012). Social validity assessment in social competence interventions for pre-school children: A review. *Topics in Early Childhood Special Education*, *32*, 164-174.

Johnson, R. B. & Christensen, L. (2017*). Educational research: Quantitative, qualitative, and mixed approaches*. Sage.

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification*, 1(4), 427-451.

Kennedy, C. H. (2002). The maintenance of behavior change as an indicator of social validity. *Behavior Modification*, *26*(5), 594-604.

Kennedy, C. H. (2005). *Single-case designs for educational research*. Allyn & Bacon.

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

Ledford, J. R., & Gast, D. L. (Eds.). (2018). *Single case research methodology: Applications in special education and behavioral sciences* (3rd ed.). Routledge.

Leko, M. M. (2014). The Value of Qualitative Methods in Social Validity Research. *Remedial and Special Education*, *35*(5), 275–286. https://doi.org/10.1177/0741932514524002

Linton, S. (1998). *Claiming disability: Knowledge and identity*. New York University Press.

Long, A. C. J., Sanetti, L. M. H., Collier-Meek, M. A., Gallucci, J., Altschaefl, M., Kratochwill, T. R. (2009). An exploratory investigation of teachers' intervention planning and per ceived implementation barriers. *Journal of School Psychology*, 55, 1-26.

National Association of School Psychologists. (2020). *NASP practice model: Improving outcomes for students and schools*.

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*, *12*, 365-383.

Radley, K. C., Dart, E. H., Fischer, A. J., & Collins, T. A. (2020). Publication trends for single-case methodology in school psychology: A systematic review. *Psychology in the Schools*, 57, 683-698.

Riley-Tillman, T. C., & Burns, M. K. (2009). *Evaluating educational intervention: Single-case design for measuring response to intervention*. The Guildford Press.

Roach, A. T., Wixson, C. S., Talapatra, D., & LaSalle, T. P. (2009). Missing voices in school psychology research: A review of the literature 2002-2007. *The School Psychologist*, *63*, 5-10.

Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33,* 359–382.

Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis*, *24*, 189-204.

Silva, M. R., Collier-Meek, M. A., Codding, R. S., & DeFouw, E. R. (2019). Acceptability assessment of school psychology interventions from 2005-2017. *Psychology in the Schools*, *57*, 62-77.

Snodgrass, M. R., Chung, M. Y., Meadan, H., & Halle, J. W. (2018). Social validity in single-case research: A systematic literature review of prevalence and application. *Research in Developmental Disabilities*, *74*, 160-173.

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

Villarreal, V., Castro, M. J., Umana, I., & Sullivan, J. R. (2017). Characteristics of intervention research in school psychology journals: 2010-2014. *Psychology in the Schools*, *54*, 548-559.

Villarreal, V., Ponce, C., & Gutierrez, H. (2015). Treatment acceptability of interventions published in six school psychology journals. *School Psychology International*, 36(3), 322-332.

Villarreal, V., & Umana, I. (2017). Intervention research productivity from 2005-2014: Faculty and university representation in school psychology journals. *Psychology in the Schools*, *54*, 1094-1105.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, *11*(2), 203–214.

Table 1.1 *Articles Searched and Included in Review*

Journal (2016-2020)	All Articles	Single-case Design (SCD) Articles	SCD Articles w/ Social Validity Assessment	Total Social Validity Assessments
JASP	89	18	16 (88.89%)	21
JSP	227	6	5 (83.33%)	7
PITS	471	28	20 (71.43%)	27
SPI	177	2	1 (50.00%)	1
SP(Q)	259	9	5 (55.56%)	11
SPR	148	17	14 (82.35%)	22
TOTAL	1,371	80	61 (76.25%)	89

Note. Percentages were derived from calculating the number of SCD articles with social validity assessments out of SCD articles. JASP = *Journal of Applied School Psychology*, JSP = *Journal of School Psychology*, PITS = *Psychology in the Schools*, SPI = *School Psychology International*, SP(Q) = *School Psychology (Quarterly)*, SPR = *School Psychology Review*

Table 1.2 *Characteristics Across Social Validity Assessments (n=89)*

Assessment Characteristics	*n*	%
Method		
Self-report	86	96.63
Interview	2	2.25
Other	1	1.12
Type of Measure		
Published/Validated	55	61.80
Researcher-developed	31	34.83
Other (e.g., interview, verbal Q & A)	3	3.37
Respondent		
Teacher	51	57.30
Student	29	32.58
Parent	9	10.11
Timing		
Postintervention	47	52.81
Preintervention	0	0.00
Pre & Post	6	6.74
Other (e.g., after each phase, weekly)	7	7.87
Not reported	29	32.58
Type of Reported Data		
Qualitative & Quantitative	68	76.40
Qualitative	15	16.85
Quantitative	5	5.62
Not reported	1	1.12
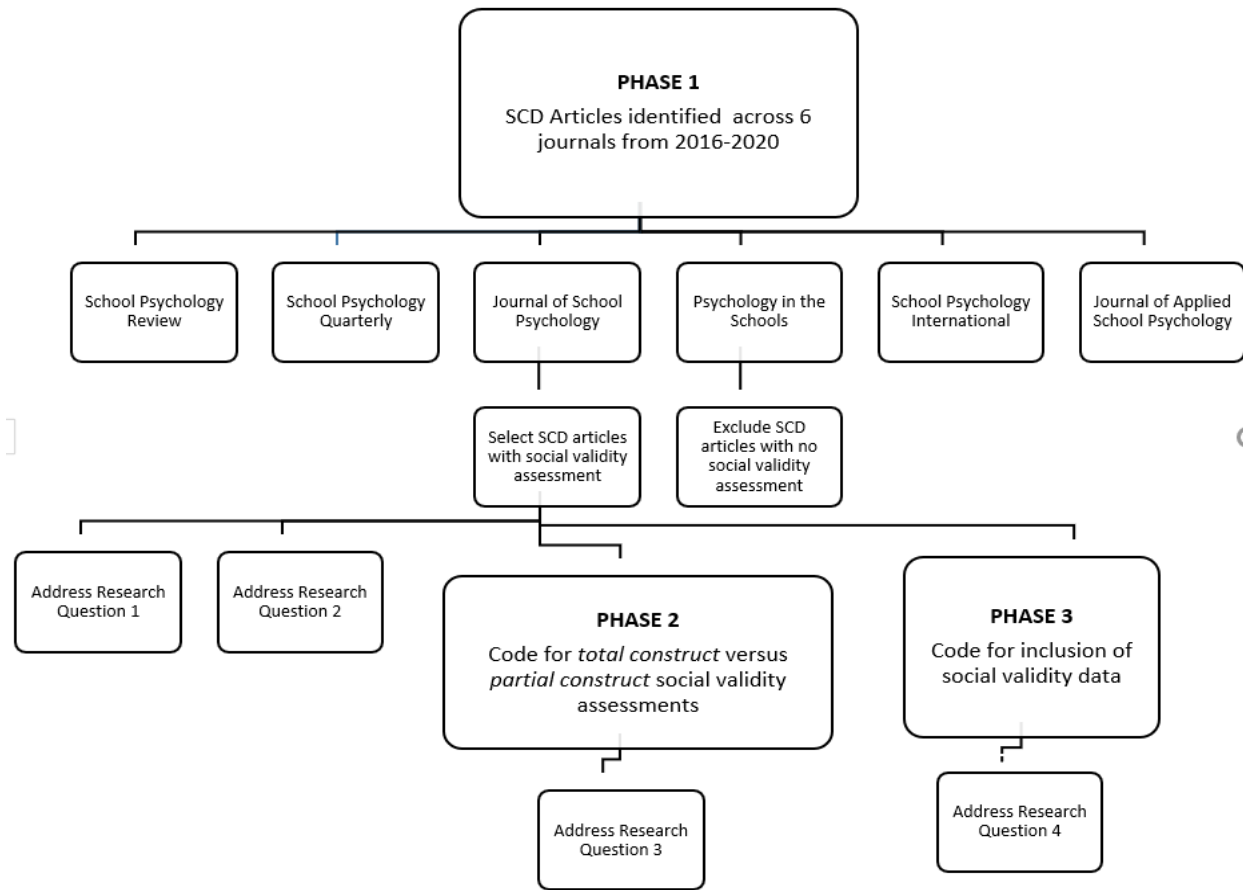
Figure 1.1 *Systematic Review Phases 1-3*
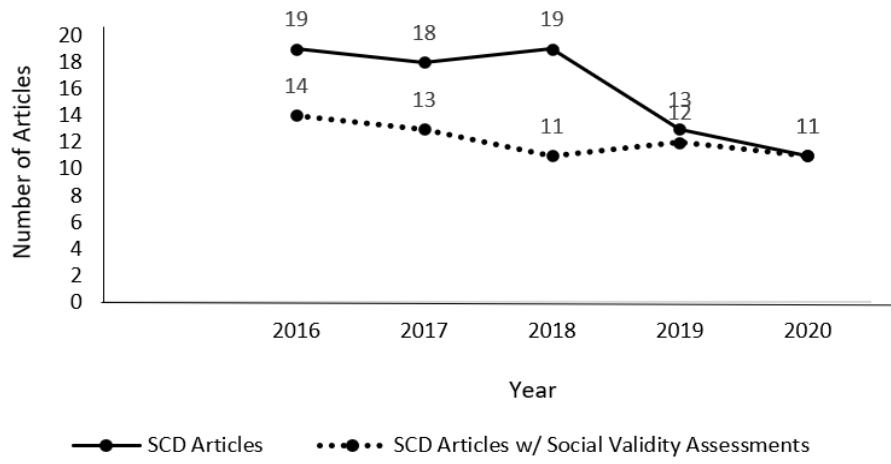
Figure 1.2 *Social Validity Inclusion in Single-case Publications by Year*
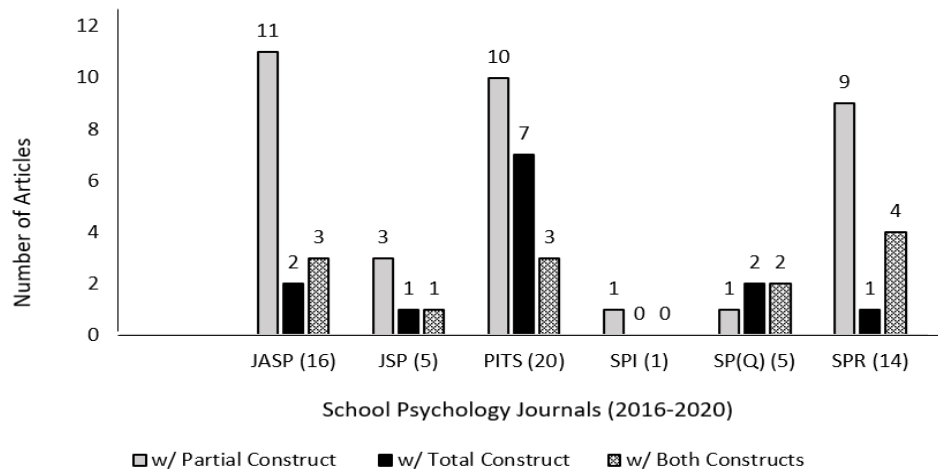
Note. The numbers in parentheses next to journals represent the total articles with social

validity assessment. JASP = *Journal of Applied School Psychology*, JSP = *Journal of*

School Psychology, PITS = *Psychology in the Schools*, SPI = *School Psychology*

International, SP(Q) = *School Psychology (Quarterly)*, SPR = *School Psychology Review*

Figure 1.3 *Types of Social Validity Constructs Assessed by Journal*

	JASP	JSP	PITS	SPI	SP(Q)	SPR
□ Interview-Partial	0	0	2	0	0	1
□ RD-Partial	3	1	11	1	2	8
▨ RD-Total	0	1	3	0	0	1
▨ PV-Partial	12	3	3	0	3	8
▥ PV-Total	6	2	8	0	6	4
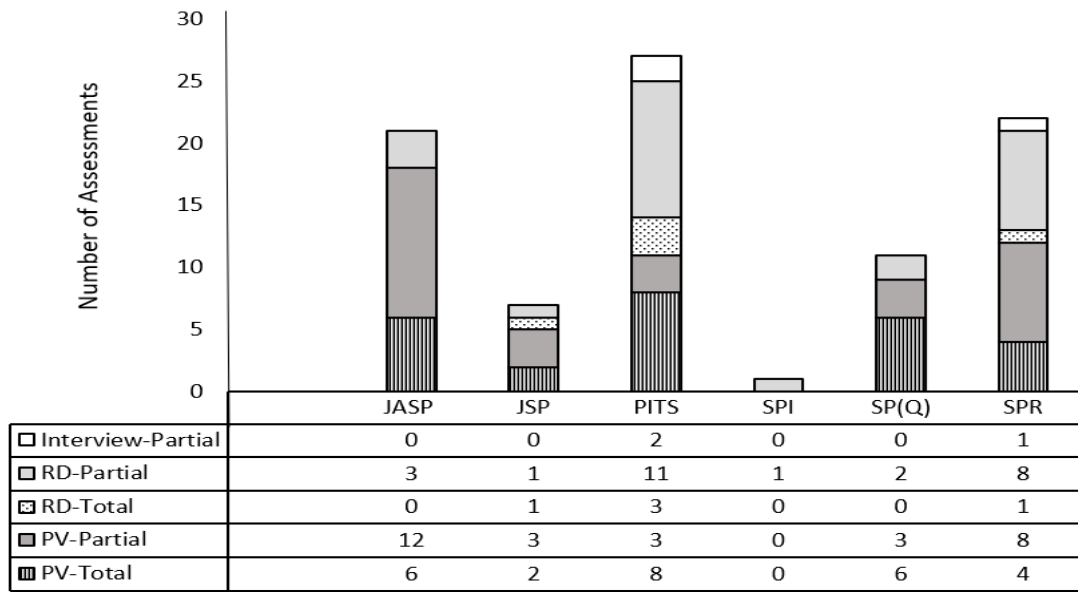
Note. RD = Researcher-developed, PV = Published/validated, Partial = Partial construct,

Total = Total construct

Figure 1.4 *Social Validity Constructs by Assessment Method*

	JASP	JSP	PITS	SPI	SP(Q)	SPR
☐ Not referenced	1	0	5	1	1	5
☐ Separate	12	4	15	0	2	11
◼ Referenced	6	3	4	0	6	2
■ Integrated	2	0	3	0	2	4
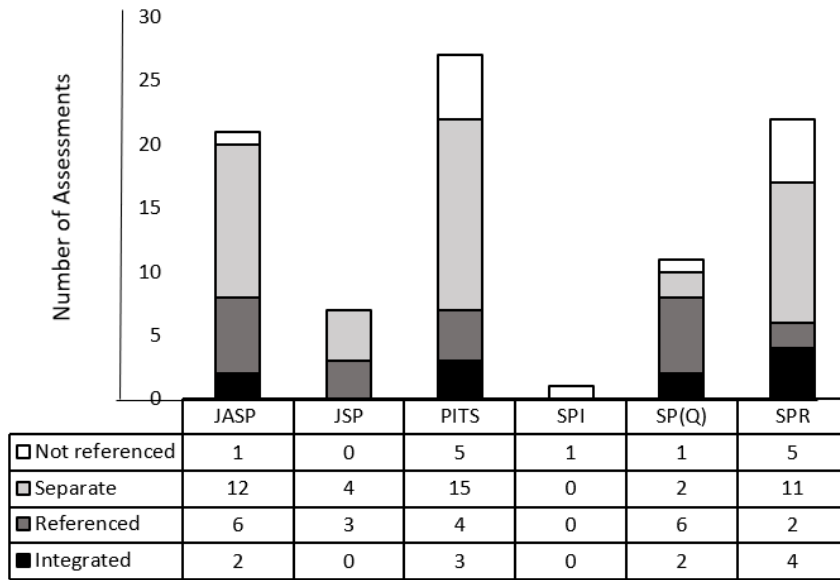
Figure 1.5 *Social Validity Data in Relation to Visual Analysis*

2 AN INVESTIGATION OF TWO INTERVENTION PROCEDURES AND THE SOCIAL VALIDITY OF THE "ASK-READ-TELL" READING STRATEGY

Over four decades ago, Durkin's (1978-1979) called attention to the absence of reading comprehension instruction in the schools and caused quite a sensation amongst the education establishment. Following many hours of classroom observations, she documented that teachers devoted a large portion of reading class time assessing students on what they read rather than instructing students on how to derive meaning from text. Further, Durkin (1978-1979) found that comprehension strategies were merely mentioned in passing and explicit instruction on the practice and application of strategies never occurred.   Subsequently, Durkin (1981) examined reading curriculum teachers' manuals and discovered that they offered little guidance for teaching comprehension and mainly focused on training students to answer low-level, literal questions with a single correct answer. The Durkin studies (1978-1979; 1981) highlighted a need for explicit instruction on comprehension strategies and instigated a dramatic shift in how the comprehension process was viewed.

Perhaps one of the most influential publications on reading instruction is the National Reading Panel's report on early reading instruction (National Reading Panel, 2000). The National Reading Panel's (2000) meta-analysis of research-based instructional methods for reading identified the five pillars of effective reading instruction: phonemic awareness, phonics, fluency, vocabulary, and comprehension. In the area of reading comprehension, the Panel regarded vocabulary development and systematic explicit instruction of comprehension strategies as essential for enhancing reading comprehension (National Reading Panel, 2000). The Panel's findings also revealed that reading comprehension of text is best facilitated by instructing

students an array of systematic strategies to assist in the recall of information, question generation, and summarizing of information (National Reading Panel, 2000).

Most researchers agree the fundamental goal of reading is to achieve meaning from text. Sideridis and colleagues (2006) defined effective reading skill as a combination of the ability to (1) decode words with fluency and (2) gain meaning through text. Hence, it is hardly surprising that successful readers frequently monitor their understanding and employ active metacognitive strategies in order to extract meaning from what they read (Anderson, 1992; Brown et al., 1995; Pressley & Wharton-McDonald, 1997). Even more revealing were studies that examined differences in reading behaviors between skilled and poor readers. For example, Paris and Myers (1981) found that poor readers did not actively monitor or correct reading errors while they read. Additionally, poor readers focused more on pronunciation of words within passages rather than their meaning and forgot more information on delayed tests when compared to skilled readers (Paris & Myers, 1981). This study's findings also suggested that struggling readers lacked the skills to effectively monitor their comprehension of text and to employ "fix-up skills" when necessary (Paris & Myers, 1981).

Pressley and Wharton-McDonald (1997) asserted that skilled readers are strategic readers. In contrast to poor readers, skilled readers attend to word meanings as exhibited by their ability to recognize more nonsense phrases embedded within passages. They also underlined anomalous information or unfamiliar words more consistently than their counterparts (Paris & Myers, 1981). Other active strategies employed by skilled readers include asking questions while they read, re-reading challenging parts of passages, taking notes while reading, and using the dictionary to look up unfamiliar words (Paris & Myers, 1981).

National statistics suggest a troubling trend in reading achievement across grade levels. According to the National Assessment of Educational Progress (NAEP, 2019), the average reading score for fourth-grade students in 2019 represented a decline from performance in 2017; a similar decline was observed in the performance of eighth-grade students. The 2019 achievement-level results indicate lower percentages of fourth- and eighth-grade students performed at or above the NAEP Proficient level in reading in 2019 compared to 2017. Only 35% of fourth graders and 34% of eighth graders performed at or above NAEP Proficient level in 2019. When compared to the 2017 NAEP data, these findings reflect a small but notable decline for both grade levels. These results suggest that reading instruction in the nation's schools may not effectively prepare students for reading success in the classroom and beyond.

In most academic contexts, the majority of learning occurs through reading, but students with reading difficulties often lack the skills to successfully comprehend text-focused classroom instruction (Brown et al., 1995; Mason et al., 2006; Paris & Myers, 1981). In fact, elementary students who read below proficiency levels are at risk of having limited access to content in core subject areas and experiencing reading difficulties throughout their school career (Lane, 2014). Moreover, research has demonstrated that reading ability affects attainment in all subjects (Mulcahy et al., 2001). For example, students with low comprehension skills exhibit difficulties with solving word problems in math and accessing curriculum content in science and social studies (Mulcahy et al., 2001).

Further, pervasive reading difficulties may set students on trajectories associated with a variety of negative consequences, including grade retention, school dropout, and diminished job opportunities (Lyon, 2001). Other associated risks with low literacy skills include future unemployment, poor earning potential, risky health behaviors, criminality, incarceration, low

psychological self-concept, and low life satisfaction (Mulcahy et al., 2001). For multiple reasons, teaching students effective reading skills and strategies is critical in preventing associated long-term adverse effects of reading difficulties.

The Prevalence and Challenges of Expository Text in Upper Elementary Education

During the early elementary grades, formal reading instruction typically places an emphasis on students' development of phonemic awareness, decoding skills, and reading fluency (Armbruster et al., 2001; Lane, 2014). In fact, educators and reading experts generally describe this as the "learning to read" phase (Armbruster et al., 2001; Durkin, 1978-1979; Lane, 2014). Beginning around the third to fourth grade, the instructional focus shifts from teaching students to read to expecting students to independently utilize reading skills in order to learn (Allington & Johnston, 2002). In other words, third- and fourth-graders are expected to transition to the "reading to learn" phase (Allington & Johnston 2002; Durkin, 1978-1979).

In addition to this shift in reading expectations for upper elementary students, a shift in text genre is also evident around the time students enter the fourth grade. Whereas students in the early grades often work with narrative texts (e.g., storybooks), upper elementary students frequently access content materials via expository or informational texts (O'Connor et al., 2017). In fact, most reading in upper elementary grades involves expository content (O'Connor et al., 2017). Hence, the ability to comprehend is key to students' academic success in subject areas such as science and social studies (Gersten et al., 2001; Mason et al., 2006; Zimmermann & Reed, 2020). Not only do students acquire science concepts, technological information, and social science principles mainly by reading expository materials, they are also required to demonstrate reading comprehension by analyzing and synthesizing complex concepts from

various expository sources such as speeches, technical writings, periodicals, and research articles (Zimmermann & Reed, 2020).

Compared to narrative texts, expository texts pose a myriad of challenges for readers. First, expository passages can be daunting as they are characteristically dense with unfamiliar information and content-specific vocabulary; and processing these elements can be cognitively taxing (Gajria et al., 2007). Another reason may be the length of many expository passages. Informational texts are typically longer than narrative texts. Stories, on the other hand, are interspersed with conversational prompts and dialogue to provide the audience with context and clues related to how the plot is unfolding (Bereiter & Scardamalia, 1987). Additionally, common informational text structures such as cause-effect, compare-contrast, and problem-solution organizations are more abstract than story plot structures (Gersten et al., 2001; Zimmermann & Reed, 2020). Some expository passages may integrate a mix of text structures, thus creating even more difficulties for struggling readers (Zimmermann & Reed, 2020). Not surprisingly, research suggested that processing expository content often poses more barriers for readers than narrative materials (Gersten et al., 2001). In light of these findings, it is not surprising that students with learning disabilities experience greater difficulty in comprehending informational texts than narrative texts or literary writings (Denton et al., 2015).

Because fourth- and fifth-grade students are expected to utilize independent reading skills to access content materials, upper elementary grade educators presume that these students have mastered reading comprehension fundamentals (Edmonds et al., 2009; Mahdavi & Tensfeldt, 2013; Mason et al., 2006). Operating under this assumption, teachers may not provide explicit instruction in strategies that support effective reading comprehension (Berkeley & Larsen, 2018; Durkin, 1978-1979; Edmonds et al., 2009; Gersten et al., 2001; Mahdavi & Tensfeldt, 2013;

Mason et al., 2006). Further, due to increasing accountability in meeting state standards and passing end-of-year assessments, teachers often focus on teaching content rather than instructing students on how to read for understanding (Edmonds et al., 2009; Pressley, 2002). In fact, reading experts identified fourth- and fifth-grade end-of-year assessments as being similar to assessments of secondary students in regard to accountability expectations (Berkeley & Larsen, 2018). For example, older elementary students are expected to work with a broad range of texts and face substantial pressure to meet the demands of rigorous curricular content (Armbruster et al., 2001; Berkeley & Larsen, 2018; Gersten et al., 2001).

In a climate where approximately 80% of students with learning disabilities have significant reading difficulties (Society for Neuroscience, 2020), it is imperative that educators recognize the importance of providing direct comprehension instruction. While educators spend a great deal of effort teaching students to read in the early grades, it is disheartening that we continue to observe the "fourth-grade slump" in reading achievement (Hirsch, 2003). Over three decades ago, Chall and colleagues' (1990) documented the decline in students' reading scores between the third and fourth-grade and this diminished performance continued as student moved into higher grades. As such, upper elementary educators require knowledge of best practices in reading comprehensive instruction and intervention to prevent students from falling farther behind.

Factors Impacting Reading Comprehension among Students with Learning Disabilities

More than half of students with learning disabilities spend over 80% of their school day in the general education classrooms (Cortiella & Horowitz, 2014), where they are required to meet the same rigorous, state-mandated standards as their peers without disabilities (Magiera & Zigmond, 2005). As students progress through school, they are increasingly expected to access

curriculum materials through independent reading (Scruggs, Mastropieri, Berkeley, & Graetz, 2010). Nonetheless, students with LD are often ill-prepared to meet these expectations for a variety of reasons (Berkeley et al., 2010; Berkeley & Larsen, 2018; Gersten, et al., 2001; Mahdavi & Tensfedt, 2013).

From the start, children who experience difficulties with basic reading skills are more likely to experience cascading patterns in reading difficulties (Stanovich, 1986). This phenomenon is often referred as the "Matthew Effect" (Stanovich, 1986). Stanovich observed that early learners who struggle to read tend to engage with fewer texts, participate in fewer strategic reading practices, and learn fewer vocabulary words. As such, these students fall increasingly further behind more-skilled readers, who decode more easily, read with proficiency, repeatedly engage in strategic reading habits, and subsequently developed larger vocabularies and wider content knowledge (Stanovich, 1986).

In addition, studies suggest that students' academic self-efficacy may be associated with their reading achievement. Lee and Jonson-Reid (2016) indicated that children formulate self-beliefs related to school at an early age, and they documented a significant association between academic self-efficacy and students' reading ability (Lee & Jonson-Reid, 2016). Moreover, many students with LD may not perceive a direct connection between their behaviors and learning outcomes (Pintrich & Schunk, 1996), and they often mistakenly interpret initial unsuccessful attempts at a learning task as an indicator of their ability level (Baird et al., 2009).

Impaired motivation and low task persistence are other important factors that hinder reading comprehension in students with LD. Students who have consecutive failing experiences sometimes perceive failure and success differently than their peers (Walker, 2003). As such, students who repeatedly experience unsuccessful efforts to overcome academic difficulties can

become less motivated to persevere at learning tasks (Gersten et al., 2001). In the context of reading, Stanovich (1986) found that students with reading difficulties were likely to gravitate toward environments that minimize engagement with reading tasks and to avoid recreational reading outside of the classroom. Hence, with fewer opportunities to practice reading skills, these students continued to struggle, fell farther behind their peers, and eventually developed the belief that persistent effort does not lead to positive academic outcomes (Pintrich & Schunk, 1996).

Another factor related to comprehension difficulty among students with LD is inefficient cognitive processing. Researchers suspect that inefficiency rather than deficiency most characterizes the problems experienced by students with disabilities (Gersten et al., 2001). Although students with LD may have the cognitive tools to process information effectively, experts suspect that breakdowns occur during metacognitive processing (Gersten et al., 2001). In other words, during strategic processing of information, students with LD experience more difficulty directing cognitive activities in a deliberate and reflective way (Gersten et al., 2001). For example, while reading, students may not recognize the need for active monitoring of comprehension (Berkeley & Larsen, 2018; Gersten et al., 2001; Zimmermann & Reed, 2020). Hence, they may not intuitively know to revisit challenging passages, re-read, or adjust reading speed to enhance comprehension (Berkeley & Larsen, 2018; Firat, 2019; Gersten et al., 2001; Zimmermann & Reed, 2020). Another contributing factor may be the absence of strategy instruction in many classrooms, suggesting some students are not equipped with tools for deciphering challenging text (Gersten et al., 2001; Zimmerman & Reed, 2020). For students who were taught strategies, it may be difficult for some to know when and which strategies to apply, while other students may rely on simpler, but less efficient procedures (Mason et al., 2006). Conversely, some strategies may appear complex and daunting, deterring students from

employing them altogether (Gersten et al., 2001; Zimmerman & Reed, 2020). Lastly, although studies demonstrated the benefit of comprehension strategy application for students with LD, researchers also found that these students often require additional time and repeated guided practice to fully grasp new procedures (Antoniou & Souvignier, 2007). Without appropriate support and practice opportunities, students with LD may not utilize taught strategies independently.

Another factor essential to the comprehension of text relates to students' knowledge of text organization and structure. In comparison to narrative texts, students with LD have been shown to exhibit more difficulty understanding and making inferences from expository texts (Denton et al., 2015; Gersten et al., 2001; Mahdavi & Tensfeldt, 2013; Mason et al., 2006; Zimmermann & Reed, 2020). The comprehension difficulties of students with LD may be explained, in part, by specific challenges presented in expository text. For example, expository writing often contains content-specific vocabulary and places a high demand on students' general content knowledge (Gajria, Jitendra, Sood, & Sacks, 2007). Reading comprehension depends on the reader's familiarity with vocabulary used to communicate domain-specific concepts (Gersten et al., 2001). Students with disabilities may work from a limited vocabulary base, bringing less of background content knowledge to their reading tasks (Gersten et al., 2001). Their comprehension suffers as a result of their vocabulary deficits (Gersten et al., 2001; Mason et al., 2016). Further, students with reading disabilities have been found to not have a sense of how informational texts are organized (Gersten et al., 2001; Mason et al., 2006). For this reason, they typically approach expository text with no particular plan of action (Meyer, Brandt, & Bluth, 1980), processing information in a random, disconnected fashion. Skilled readers, on the other hand, break large pieces of information into "chunks", maintain conscious control of their

reading process, and make internal connections among ideas as they read (Gersten et al., Mason et al., 2006; Zimmermann & Reed, 2020).

Another source of comprehension difficulty for students with LD is the assimilation of erroneous background knowledge during reading tasks. Multiple researchers have indicated activating prior knowledge and connecting it to text-based information enhances reading comprehension (Berkeley & Larsen, 2018; Firat, 2019; Gersten et al., 2001; Paris & Myers, 1981; Pressley & Wharton-McDonald, 1997). In other words, comprehension is the interaction between novel information and previously acquired knowledge. As Idol-Maestas (1985) pointed out, existing schema serve as the knowledge base from which readers draw upon while reading. As they encounter novel information in texts, readers interpret and assimilate new concepts into their existing knowledge base in order to achieve meaning (Idol-Maestas, 1985). Students with LD may utilize this technique; but for other students, this strategy can actually exacerbate comprehension difficulties. Williams (1993) found that poor readers sometimes inserted inaccurate or irrelevant background knowledge into their texts. This insertion of misinformation led to inaccurate concept representation and interpretation of texts (Gersten et al., 2001; Williams, 1993). Moreover, these faulty connections of new information with irrelevant background knowledge caused further confusion, leading to additional breakdowns of comprehension. As a result, students experienced difficulties with formulating predictions, making inferences, identifying accurate story themes, and summarizing (Williams, 1993).

Monitoring meaning during reading is hardly a new concept and has been established as an essential form of metacognition for successful comprehension of text. Since the seminal Durkin studies (1978-1979; 1981), experts in the field have viewed reading as a goal-directed activity. Within the context of reading, the aim is to derive meaning from text. Such activity

requires the learner to activate his or her cognitive awareness and to maintain certain behaviors in order to attain personal goals (Berkeley & Larsen, 2018). These processes are referred as self-regulation or self-monitoring (Berkeley & Larsen, 2018; Paris & Myers, 1981). Zimmerman (2008) conceptualized self-regulation as an integration of motivational, behavioral, and metacognitive processes. Further, effective self-regulation with regard to reading occurs in three sequential phases: before (forethought phase), during (performance phase), and after (self-reflection phase) (Zimmerman, 2008). The self-regulated reader is described as goal-directed, strategic in reading behaviors, self-aware, reflective, and flexible in his or her processing of information (Zimmerman (2008). Conversely, students with LD often are unaware that they should actively monitor their comprehension, thus they do not recognize the need to repair comprehension problems (Gersten et al., 2001). In fact, Paris and Myers (1981) investigated differences in comprehension monitoring between skilled and struggling readers. They asked students to read stories containing nonsense words and phrases. As students read, Paris and Myers measured comprehension monitoring by tallying the number of times students spontaneously hesitated, repeated, and self-corrected. Based on these observations, they posited that struggling readers did not actively monitor incomprehensible information nor correct nonsense words and phrases (Paris & Myers, 1981). This absence or inability to self-evaluate during reading tasks consequently resulted in poor comprehension for students with reading disabilities (Berkeley & Larsen, 2018).

Comprehension Instruction, Strategies, and Tools

Afflerbach and colleagues (2008) described reading strategies as "deliberate, goal-directed attempts to control and modify the reader's efforts to decode text, understand words, and construct meanings of text" (p.368). The use of strategies represents intentional effort to achieve

a goal, whether it be comprehension of a book chapter, deciphering a piece of poetry, or understanding instructions for assembling a bookcase (Afflerbach et al., 2008). Previously called transactional strategies, metacognitive strategies are used to enhance readers' engagement with their texts (Anderson, 1992; Brown et al., 1995; Pressley & Wharton-McDonald, 1997). Pressley and Wharton-McDonald (1997) asserted that proficient readers consciously employ strategies while reading and take an active role in the construction of meaning from text. When reading for comprehension, skilled readers actively apply strategies such as previewing text, activating prior knowledge, using context clues, self-monitoring, and summarizing in order to enhance their comprehension (Berkeley et al., 2010; Berkeley & Larsen, 2018; Mason et al., 2006; McCallum et al., 2011; Pressley & Wharton-McDonald, 1997).

Direct and Explicit Instruction

There are multiple methods for improving students' reading comprehension. Previous research reviews have identified critical elements for effective reading comprehension instruction for students with disabilities (Gersten et al., 2001; Berkeley et al., 2010; Shelton et al., 2019). One established method of improving comprehension is direct strategy instruction (Gersten et al., 2001; Berkeley et al., 2010; Shelton et al., 2019). Following the Durkin studies, reading comprehension research repeatedly demonstrated the benefits of direct and explicit instruction of strategies (Berkeley & Larsen, 2018; Gersten, et al., 2001; Mahdavi & Tensfedt, 2013). In their seminal study, Wong and Wilson (1984) demonstrated that students with LD could overcome ineffective reading habits when they were taught to read strategically and monitor for comprehension. With explicit instruction, students were able to organize pieces of information into categories around subtopics, increased their metacognitive awareness, and engaged actively in practices to improve their understanding of text (Wong & Wilson, 1984). In

their research exploring students' limited access to content materials due to poor reading skills, Brown and colleagues (1995) found that second graders who learned strategies acquired more content from daily lessons compared to their counterparts. Specifically, for upper elementary students, explicit instruction with a focus on teaching readers how to approach expository texts has proven to be beneficial (Armbruster et al., 2001; Berkeley & Larsen, 2018; Gersten et al., 2001; Lane, 2014; Mason et al., 2006).

Purposeful and explicit instruction is vital to supporting students with reading difficulties and can have a positive impact on the reading behaviors and engagement of students. For example, Anderson (1992) demonstrated that 6th through 11th graders who were given reading strategy instruction were more willing to read challenging texts and to attempt to understand complicated materials. Further, students provided with explicit instruction became more active participants in their reading experience (Mahdavi & Tensfeldt, 2013) and more involved in deriving meaning from what they read (Anderson, 1992).

Alnahdi (2015) demonstrated that explicit, systematic instruction led to improved reading skills in individuals with disabilities. In fact, when compared to other groups of students, students with LD benefitted more from strategy instruction and made the greatest gains in improving their comprehension performance (Mahdavi & Tensfeldt, 2013). Finally, explicit instruction provides a systematic framework for guiding students in how and when to use strategies, enhancing their ability to self-monitor, apply tools to help them problem-solve, and persist through difficult texts (Gersten etal., 2001).

Multiple-Strategy Instruction

Although individual comprehension strategies can be helpful, efficacious application of strategies involves coordinating and combining several strategies to maximize reading

comprehension. Multiple-strategy instruction trains students to adjust strategies and use them flexibly as needed to enhance comprehension (Armbruster, 2001). Pressley and colleagues (1992) noted the remarkable shift away from teaching one strategy at a time as there is no one strategy that will provide for the reading needs of all children (Gersten et al., 2001; Mahdavi & Tensfeldt, 2013). Thus, multiple instructional elements are necessary to support comprehension for readers of diverse needs and reading ability levels (Gersten et al., 2001). Further, no one strategy can be effectively applied for processing various genres and text organizations (Gersten et al., 2001; Mahdavi & Tensfeldt, 2013; Mason et al., 2006).

Several multiple-strategy packages have been shown to yield improved comprehension of text across grade levels (Gersten et al., 2001). Reciprocal teaching (Palinscar and Brown, 1984) is a well-established multiple-strategy intervention that combines four strategies: asking questions, summarizing, clarifying any confusing word or sentences, and predicting what might occur next. In this intervention, teachers and students collaborate in helping students learn the four strategies and apply them flexibly as they are needed in reading narrative and expository texts. Another multi-approach strategy, POSSE (Englert & Mariage, 1991), uses a graphic organizer in combination with peer-mediated instruction and a variety of metacognitive strategies focusing on expository text structures. When practiced and applied over a sustained period of time, students using POSSE demonstrated significant generalizable effects when engaging with novel texts (Gersten et al., 2001). Overall, the use of multiple-strategy interventions has been found to not only improve the immediate processing of text, but also to generalize to other reading tasks outside of the classroom environment (Gersten et al., 2001).

The Use of Strategies Before, During, and After Reading

Effective comprehension of text involves the coordination of highly complex and well-developed skills during each distinct stage of a reading task. Specifically, researchers describe three crucial phases of the reading process that, when applied systematically, appear to improve students' ability to construct meaning from text (Edmond et al., 201; Mason et al., 2016; McCallum et al., 2011; Pressley & Afflerbach, 1995). Prior to reading, students are encouraged to have clear reading goals, activate prior knowledge, think about the title or topic, and decide what they want to learn about the topic (Idol-Maestas, 1985; Mason et al., 2006; McCallum et al., 2011; Pressley & Afflerbach, 1995; Ridge & Skinner, 2011). Next, while reading a text, students should monitor their understanding by thinking about reading speed, connecting textual information to their own knowledge, using "fix-up" strategies such as re-reading, slowing down their reading rate, and using other strategies to clarify understanding of the text (Mason et al., 2006; McCallum et al., 2011; Pressley & Afflerbach, 1995). Finally, during the post-reading phase, readers continue thinking about the passage, identify main ideas, summarize key information either via writing or peer discussion, answer pre-reading questions, and encode details into long-term memory (Mason et al., 2006; McCallum et al., 2011; Pressley & Afflerbach,1995).

Self-Monitoring and "Fix-Up" Strategies

An essential component of reading with understanding is the ability to reflect on a task and evaluate how well it is being executed. Skilled readers are proficient at monitoring their understanding and use metacognitive strategies to think about and engage purposely with the text. In fact, they know when they understand what they read and when they do not (Armbruster, 2001; Berkeley & Larsen, 2018; Paris & Myers, 1981). Once they detect difficulties, self-regulated readers employ "fix-up" strategies to resolve problems in comprehension (Armbruster,

2001; McCallum et al., 2011). "Fix-up" strategies include identifying where the difficulty occurs, knowing what the problem is, adjusting reading rate to fit the text demands, re-reading trouble spots in the passage, looking within the text for clues, and increasing focus during reading (Armbruster, 2001, McCallum et al., 2011; Pressley et al., 1997). Research has established the role of metacognition as a key factor in students' achievement (Berkeley & Larsen, 2018). Although students with disabilities may experience comprehension difficulties, they can be trained to engage in strategic behaviors and utilize metacognitive skills to enhance their comprehension of written materials as they progress through the upper grades (Armbruster, 2001; Berkeley & Larsen, 2018; Lane, 2014)

Self-Questioning Behaviors

Another important reading comprehension strategy is self-questioning. When engaging in self-questioning, students stop periodically to ask and answer questions related to the text (Mahdavi & Tensfeldt, 2013; Taylor et al., 2002). By generating questions, readers become more cognizant of whether they can answer the questions and if they can make meaning from what they are reading (Armbruster et al., 2001). There is evidence to suggest that self-questioning is a form of active responding to text (Taylor et al., 2002). This strategy has been shown to aid in monitoring one's reading thus resulting in greater comprehension of written materials (Taylor et al., 2002). More specifically, self-questioning has been shown to be effective for improving the reading comprehension of elementary school students (Chan, 1991). McGee and Johnson (2003) found that elementary students who were taught to ask questions and make predictions based on the context clues in passages significantly increased their comprehension scores. In fact, the less proficient readers in this study achieved greater gains than the more skilled readers (McGee & Johnson, 2003).

Common components of the self-questioning procedure include direct instruction, guided and independent practice, and corrective feedback (Mahdavi & Tensfeldt, 2013; Taylor et al., 2002). Although this procedure requires students to generate their own questions, a modified self-questioning procedure in which students are provided with several general questions and prompts may be more appropriate for students with disabilities (Taylor et al., 2002). In Taylor and colleagues' study (2002), students were trained to ask and answer 10 generic questions printed on a laminated card (e.g., Who is the main character? Where does the story take place) at two predetermined points within the reading passage and at the end of the passage. At each stopping point (marked by a star) in the passage, students were required to answer all 10 of the self-questions and record their responses. Students could re-answer or change their answers at each stopping point. Prior to taking a comprehension test, students listened to their recorded responses. Results indicated that students attained significantly greater comprehension in the self-questioning condition than in the no intervention condition (Taylor et al., 2002).

Text Discussion and "Thinking Aloud"

The Institute for Education Sciences What Works Clearinghouse's (IES-WWC) recommendations for improving comprehension in elementary students is to engage in discussion-based interactions about the meaning of text (Shanahan et al., 2010). Studies have demonstrated the positive effects using discussion as a comprehension practice with older students (Shanahan et al., 2010). In fact, the IES Reading Panel suggested utilizing activities that encourage students to "argue for or against points raised in the discussion, resolve ambiguities in the text, and draw conclusions or inferences about the text" (Shanahan et al., 2010, p. 23). As such, teachers have a critical role in asking questions that instigate high-level thinking and facilitating discussion about the meaning of text.

Related to the IES Panel's recommendation of interactive engagement during reading tasks, peer interaction has been shown to have a positive impact on comprehension as students discuss their ideas and assist one another in answering questions related to readings (Mahdavi & Tensfeldt, 2013). Further, collaborative work with peers is critical for facilitating "thinking aloud" about text being read. Thinking aloud can be integrated as part of the self-questioning procedure or during interactive dialogue with peers (Gersten et al., 2001; Vaughn et al., 2000). Moreover, Vaughn and colleagues (2000) posited that interventions with a "thinking aloud" component are critical in producing optimal outcomes for students' learning and understanding of text. In addition to improved comprehension, Fox and Wilkinson (1997) demonstrated the beneficial effect of student participation in peer-led discussion groups (i.e., children literacy clubs) on engagement and increased enjoyment of reading content.

Procedural Organizers

Visual organizers have been shown to aid students in processing informational text as well as orient students to employ strategies as they read. The review of reading comprehension strategies by Sencibaugh (2007) found that students with LD were successful in increasing their comprehension performance when visually related strategies were used. In fact, Idol-Maestas (1985), a prominent figure in reading research, developed an "advance organizer" to guide students through the steps of various comprehension strategies during reading. Her research was based on a review of 135 studies concluding that the use of procedural organizers had a positive effect on comprehension (Idol-Maestas, 1985). She formulated the strategy steps into an acronym--TELLS--comprised of the following steps: (T) Title – what is the title and does it give clues about the story; (E) Examine and skim for clues; (L) Look for important words; (L) Look for hard words; (S) Setting – decide on setting and whether the story is fact or fiction (Idol-

Maestas, 1985). The aim of this step-by-step organizer was to encourage readers to attend to the text, activate existing background knowledge, and look for clues within the context of the passage. As a result, all participants in Idol-Maestas's (1985) single-case design study attained increased comprehension scores during the TELLS intervention. Likewise, results from the Ridge and Skinner multiple-baseline study (2011) suggested improved comprehension levels and rates on expository passages when the TELLS procedure was employed across all three students.

TWA, a multiple-strategy comprehension tool for expository reading, is another procedural organizer that has been demonstrated to enhance reading comprehension throughout the reading process (Baker et al., 2002). TWA (Think before reading, think While reading, think After reading) incorporates empirically-based comprehension strategies into nine steps (Baker et al., 2002). Using the TWA step-by-step guide, students are prompted to activate background knowledge, and think about the author's purpose, what they already know about the topic, and what they want to learn prior to reading (Firat, 2019; Mason et al., 2006). As students read, they are encouraged to attend to their reading rate, connect new information to acquired knowledge, and reread difficult parts of the text (Firat, 2019; Mason et al., 2006). Once students complete the passage, they are required to think about the main idea, summarize, and decide what they have learned from the reading (Firat, 2019; Mason et al., 2006). Findings from a study by Mason and colleagues (2004) suggested that introduction of TWA instruction within the Self-Regulated Strategy Development (SRSD; Graham & Harris, 2003) framework resulted in improved comprehension for students with and without LD. Additionally, a single-case design study in which three sixth-grade students with LD were taught to use the TWA organizer supported Mason's (2004) findings of improved expository text comprehension (Firat, 2019).

The "ART" of Reading Comprehension

The Ask-Read-Tell (ART) strategy is comprised of several evidence-based components (McCallum et al., 2011). Similar to the TELLS (Idol-Maestas, 1985) and the TWA (Baker et al., 2002) interventions, ART is a multicomponent procedural framework consisting of step-by-step guidance for readers to follow as they engage with expository texts. Additionally, ART procedures encourage the reader to apply comprehension strategies before, during, and after the reading task. Specifically, metacognitive features in the ART package include (a) self-questioning prior to reading, (b) active self-monitoring of comprehension and the application of strategies to "repair" comprehension difficulties while reading, and (c) an optional peer discussion of text and "thinking aloud" session following the completion of the reading task (McCallum et al., 2011).

The initial research on ART consisted of a group-design study to investigate the intervention's impact on reading comprehension with high school students (n =115) in a summer enrichment program for at-risk youth (McCallum et al., 2011). The study was conducted over a two-week period and was held on a university campus. Participants were 10th and 11th grade students, and 15-18 years old. They attended the enrichment program for one hour each day over the period of 8 consecutive weekdays. Students were assigned to one of three groups: ART (Ask, Read, Tell), ART+PD (Ask, Read, Tell + Peer Discussion), or a control group. Students in both the ART or the ART+PD group received instruction on the procedure as they read fourth-grade level passages consisted of 400 words. The control group was simply asked to read the passages. All three groups were required to answer 10 multiple-choice questions once they completed their reading.

McCallum et al., (2011) observed higher comprehension scores among students receiving the ART+PD condition; however, they were not able to achieve significant differences between

students receiving the ART intervention (without PD) and students in the control group

(McCallum et al., 2011). There were several reasons that might account for the overall outcome.

First, the literature suggests that students with reading disabilities typically have difficulties in

learning self-regulatory strategies (Swanson & Siegel, 2001). Second, the brief treatment time of

two weeks may not have been sufficient to allow students to learn and master the ART strategy

(McCallum et al., 2011). Research suggests that students who are at-risk or diagnosed with LD

require interventions of increased intensity; one of the ways to increase intensity is by

lengthening the duration of any given intervention (Fuchs et al., 2014). Furthermore, the self-

regulated strategy development (SRSD) model developed by Harris and colleagues (2008)

suggests a sufficient time investment is needed to adequately teach academic strategies and self-

regulated skills. Specifically, the process of developing and mastering an academic strategy

begins with teacher modeling and guidance, progressing to students' independent application and

practice, and eventually resulting in generalization of the strategy (Harris et al., 2008). Due to

the short intervention timeframe, it is possible that participants in the McCallum et al. study

(2011) did not have sufficient support and practice to achieve competency, and for this reason,

did not benefit from the intervention. A third limitation of McCallum and colleagues (2011)

study relates to generalizability. Because participants were recruited from a two-week summer

remediation camp, the setting was potentially different from the classroom environment

(McCallum et al., 2011). A fourth limitation relates to the integrity of strategy implementation.

The study design did not include procedures to provide feedback and reinforcement to students,

thus it is possible that extrinsically motivated students did not feel compelled to read and follow

the ART procedures with fidelity, which might have negatively impacted the overall outcome

(McCallum et al., 2011). Finally, students participating in the ART intervention were not

provided with opportunities to discuss their needs and share their perspectives on the benefits

and challenges of using the ART strategy. Gathering students' perspectives may assist

researchers to refine the application of the ART strategy in authentic classroom settings.

The Social Validity Framework

Social validity is generally defined as the level of acceptance for variables related to a

procedure or program (Carter, 2010). From its initial conceptualization and application in the

field of applied behavior analysis (Wolf, 1978), social validity has been extended to school-

based research and practices to demonstrate participants' attitudes toward new programs and

interventions (Eckert & Hintze, 2000; Greer et al., 2012; Gresham & Lopez, 1996). Since that

time, the importance of social validity has gained increased relevance within the context of

schools in part due to its link to intervention fidelity and effectiveness (Eckert et al., 2017;

Horner et al., 2005; Leko, 2014).

Wolf's (1978) seminal work articulated social validity as having three defining

components: social significance (i.e., desirability deemed by society), social appropriateness

(i.e., acceptability deemed by participants or consumers), and social importance (i.e., satisfaction

deemed by participants or consumers). Building on an integrated definition of social validity

based on current research, this study focused on the acceptability aspect of social validity,

operationally defined as the perceived benefit, relevance, and importance of a program or

intervention (Greer et al., 2012). Specific to school-based practices, the existing body of

literature has demonstrated intervention acceptability as a critical factor influencing

implementation fidelity, outcome expectancy, and treatment effectiveness (Eckert et al., 2017;

Miltenberger, 1990; Waas & Anderson, 1991). For these reasons, school psychologists (and

other educators) generally seek to garner acceptability among teachers, students, and other

stakeholders in regard to assessment tools, curriculum, interventions, and instructional methods (Eckert et al., 2017; Glover & Albers, 2007; Nastasi & Truscott, 2000; Rowe, 2012).

Paradoxically, although student achievement continues to be the central focus in the field of education and the demand for effective, evidence-based interventions continues to rise, there appears to be relatively few social validity studies examining student acceptability of evidence-based academic interventions, specifically with regard to reading interventions (Chanese, 2007; Eckert et al., 2017; Rowe, 2012). In fact, the body of current social validity research reflects a proclivity towards investigation of acceptability of behavioral rather than academic interventions (Eckert & Hintz, 2000; Eckert et al., 2017).

The Press to Explore Students' Perceptions of Academic Interventions in the Schools

Within the multi-tiered systems of support (MTSS), there is an increased emphasis placed on student participation and engagement with the learning process (Rohrbeck et al., 2003). Since the provision of academic interventions includes students as direct consumers, intervention effectiveness and sustainability generally would be enhanced if students' views regarding academic interventions were taken into consideration (Mautone et al., 2009). Schunk (1996) and Stipek (1996) demonstrated that interventions that were perceived favorably by students were more likely to increase their self-efficacy and motivation, which in turn resulted in the mastery and competence of academic skills. To this end, it is imperative that researchers and practitioners consider the acceptability of academic interventions and strategies from students' perspectives (Eckert et al., 2017; Rowe, 2012; Waas & Anderson, 1991). In fact, Wolf (1978) posited that only consumers themselves can truly determine whether a program or intervention is helpful and socially important.

Involving students in the selection and implementation of academic interventions by way

of assessing their perceptions may be beneficial in several ways. First, monitoring treatment acceptability is essential for ethical and practical reasons. Educational professionals have an ethical obligation to develop and utilize interventions that are responsive to students' unique needs (Jacob & Hartshorne, 2007). As such, student perceptions of an intervention may inform decisions related to the selection, delivery, and sustainability of services. Second, including the perspectives of students may result in nuanced knowledge that informs the enhancement of a strategy's applicability and usefulness. Third, despite the importance of educators' perspectives regarding intervention utility and feasibility, the sole reliance on adult perceptions may potentially limit the breadth of understanding regarding treatment acceptability. For instance, teachers' perceptions of interventions represent only the views of individuals who are on the delivery side of the implementation process. It is also necessary to consider the lived experiences of those on the receiving end by exploring their perceptions and feelings regarding intervention goals and implementation procedures. Thus, to effectively tailor interventions in meaningful ways that respond to students' needs and preferences, it is especially critical to consider the views from the consumer's stance. Finally, obtaining input from students may raise potential concerns or barriers that were previously not considered by intervention developers and researchers (Finifter et al., 2005). When providing consultative services, students' feedback may inform school psychologists' decision making as they collaborate with school personnel on the selection and delivery of academic interventions (Arra & Bahr, 2005).

Purpose of Study

Although the ART strategy is recommended as a reading intervention for students with deficits in reading comprehension (https://www.interventioncentral.org/response-to-intervention), there appears to be limited empirical evidence to support its efficacy in improving

in students' academic performance (McCallum et al., 2011). Based on findings from existing

literature, students with poor comprehension skills appear to benefit from using multiple

strategies that encourage them to actively interact with the text throughout the reading process,

monitor their understanding, ask questions, think aloud, engage in discussion about their reading,

and utilize a procedural organizer to cue them of strategic steps (Baker et al., 2002; Firat, 2019;

Idol-Maestas, 1985; Mason et al., 2006; National Reading Panel, 2000; Ridge & Skinner,

2011).

The primary goal of the current study was to extend research on the ART procedure by

investigating its effects on the reading comprehension of upper elementary school students with

learning disabilities. The study replicated certain aspects of the intervention procedure by

McCallum and colleagues (2011). Similar to the McCallum et al., (2011) study, three conditions

were used as the independent variable in this study: 1) a Control condition in which students did

not use any prescribed strategies while reading, 2) an ART condition in which students followed

the three-step cognitive strategy, and 3) an ART + PD (Peer Discussion) condition in which

students adhered to the 3-step ART sequence followed by a discussion of text with a peer. The

dependent variables were students' reading comprehension level (%C) and rate (%C/M) as

measured by students' comprehension performances on 400-word expository passages.

The second purpose of the study was to address several limitations in the McCallum et

al., (2011) study through modifications to that study's design. The current study included the

following enhancements: 1) Extending the study period from 2 weeks to approximately 10

weeks, 2) Utilizing a single-case design to achieve a clear understanding of individual

differences within participants and across participants over time, 3) Providing participants with a

four-day ART and ART+PD training after the baseline phase and prior to the intervention phase,

using the flexible self-regulated strategy development (SRSD) framework (Harris et al., 2008) to ensure accurate application of the ART and ART+PD procedures, and 4) Reinforcing participants on a weekly basis for active participation in the learning and application of ART in order to maintain a consistent level of motivation and engagement.

The final purpose of this investigation was to garner participating students' perspectives on their experiences with learning and applying the ART procedure as well as their perceptions of the intervention goals, implementation procedures, and outcomes.

Phase 1: Research Questions

Phase 1 of the current study aimed to investigate the following research questions:

1. Does ART lead to an increase in reading comprehension level and rate on expository text for upper elementary students with a learning disability in reading comprehension?

2. Does the addition of a peer discussion component to the ART strategy lead to greater gains in comprehension level and rate?

Phase 1: Alternating Treatments

Participants

The Special Education Coordinator nominated fourth-grade students who potentially fit the inclusion criteria of the study. The inclusion criteria were as followed: (a) Received special education under the Individuals with Disabilities Education Act (IDEA) category of Specific Learning Disability, (b) Demonstrated difficulties with reading comprehension (based on teacher report and direct assessments of reading achievement), (c) Received specialized instruction in reading comprehension in a resource classroom, (d) Received special education services in a small group of three to four students for at least one period a day, and (e) Received at least 80%

of their instruction in the general education setting. Other inclusion criteria included English as students' primary language; enrollment in the fourth grade; normal or corrected hearing and vision; and intellectual functioning within the average range. Using these criteria, the primary researcher screened the list of seven nominees, and five students were selected for recruitment. The primary researcher contacted parents/caregivers of these students via email with information about the study. Parents/caregivers of three students responded and indicated interest in the study. All three nominees participated in the study after obtained parent consent and student assent were obtained.

Bryson

Bryson was an eleven-year old, multi-racial boy in fourth grade. Bryson previously attended school at a nearby district and moved to his current school at the beginning of his fourth-grade year. Since second-grade, Bryson has received special education under the SLD category for reading. At the time of the study, he was receiving 45 minutes of daily specialized instruction in reading in a small-group setting. His reading comprehension performance on the Winter Administration of Renaissance STAR Reading benchmark was at a 3.1 grade-equivalency.

Charlotte

Charlotte was a ten-year old, White girl in fourth grade. In first grade, Charlotte was diagnosed with ADHD by her pediatrician. Charlotte was prescribed Guanfacine and Dyanavel XR daily. Charlotte's mom reported that the medications had helped Charlotte with attention and focus. Since the second grade, Charlotte participated in the Early Intervention Program for reading comprehension in a small group twice a week. Despite having received interventions, she was still experiencing difficulty with comprehending content materials. By the beginning of

her fourth grade year, Charlotte was evaluated and found eligible for special education under the SLD category for reading. At the time of the study, Charlotte was receiving 45 minutes of specialized instruction in reading per day for five days a week in a small-group setting. Her reading comprehension performance on the Winter Administration of Renaissance STAR Reading benchmark was at a 3.4 grade-equivalency.

Cole

Cole was a ten-year old White boy in fourth grade. He had been attending his current school since second grade during which he received RTI interventions for basic reading skills, reading fluency, and reading comprehension twice a week. Intervention data indicated that he was making progress in basic reading skills and slow progress in reading fluency. However, his reading comprehension was significantly below grade-level. At the beginning of third grade, Cole was evaluated and found eligible for special education under the SLD category for reading. At the time of the study, he was receiving 45 minutes of specialized instruction each day in a small-group setting. His reading comprehension performance on the Winter Administration of Renaissance STAR Reading benchmark was at a 2.9 grade-equivalency.

Discussion Partners

Three additional fourth-grade students served as discussion partners during the discussion component of the ART+PD intervention. Students serving as discussion partners for the ART+PD condition were recruited from outside of the participants' special education classrooms. The lead fourth-grade teacher nominated eight students who potentially fit the inclusion criteria for the study. The selection criteria for discussion partners included English as students' primary language; enrollment in the fourth grade; fulltime placement in general education; at or above grade-level reading achievement; and regular school attendance. Other inclusion criteria were as

followed: (a) appropriate interactions with peers on a regular basis, (b) frequent compliance with teacher instruction, (c) and the ability to sustain attention and participation for the entire length of the ART+PD intervention. The primary researcher met with the homeroom teachers of nominated students to verify participation eligibility prior to emailing students' parents/ caregivers with study information. Three nominees participated in the study after parent consent and student assent were obtained.

Kylie

Kylie was a nine-year old, White girl in the fourth grade. She participated in the Gifted and Talented program once a week. She had been attending the current school since kindergarten and described herself as an avid reader. Her reading comprehension performance on the Winter Administration of Renaissance STAR Reading benchmark was at a 6.3 grade-equivalency.

Nate

Nate was a nine-year old, Hispanic boy in the fourth grade. He had been enrolled at the current school since first grade. Nate enjoyed learning about science and playing computer games. His reading comprehension performance on the Winter Administration of Renaissance STAR Reading benchmark was at a 6.2 grade-equivalency.

Rhea

Rhea was a ten-year old, White girl in the fourth grade. She participated in the Gifted and Talented program once a week and had been enrolled at the current school since pre-Kindergarten. She indicated reading interests in Greek mythology and mystery series. Her reading comprehension performance on the Winter Administration of Renaissance STAR Reading benchmark was at a 6.9 grade-

equivalency.

Setting

Participants and discussion partners were from an elementary school in a Southeastern

public school district. The selected school had approximately 490 students spread across grades

Pre-Kindergarten through fifth. The school was located in a peri-urban community that was

surrounded by farms as well as a mix of newly developed and established neighborhoods. The

school demographics consisted of over 80% White, almost 11% Black, 4.5% Hispanic, 2.9%

Multi-race, 0.2% Asian/Pacific Islander, and 0.4% American Indian/Alaskan Native.

The investigation was conducted in one of the school's resource rooms where training on

the ART procedure, implementation of data collection procedures, and administration of

comprehension tests took place. The resource room was slightly smaller than a typical

classroom that could accommodate up to eight students.

Experimental Design

An alternating treatments design (Barlow & Hayes, 1979; Ledford & Gast, 2018) was

used to compare the effects of the no intervention (Baseline), ART, and ART+PD on each

student's reading comprehension level (%C) and reading comprehension rate (%C/M). In the

alternating treatments design (ATD), participants were exposed to an experimental condition in

which two or more interventions were rapidly alternated from one session to the next (Barlow &

Hayes, 1979; Leford & Gast, 2018). Specific to this study, participants received ART and

ART+PD interventions in an alternating fashion from one session to the next over the course of

multiple weeks. Following the establishment of a stable baseline data path, a training phase of

the ART and ART+PD was implemented for four consecutive sessions. Thereafter, an

intervention phase began with intervention data collected on two sessions per week for a

minimum of five consecutive weeks. Since two intervention conditions were alternatively

administered during the week for multiple consecutive weeks, controlling for order effect was

taken into consideration. As such, counterbalancing the implementation order of the two

interventions across weeks was used.   To minimize the possibility of carryover and sequencing

effects, a random alternation of conditions was implemented with no condition repeating until all

conditions had been conducted (Ledford & Gast, 2018). Additionally, the intervention sessions

were separated by at least one full day to minimize the effects of multi-treatment interference

(Ledford & Gast, 2018).

Dependent Variables

For a thorough assessment of reading comprehension, Skinner (1998) recommended that

both comprehension level and rate be considered. For example, two students who achieved the

same number of correctly answered comprehension questions may be perceived as having

similar reading skills. However, knowing that Student A took less time than Student B to read

the same passage may better differentiate the reading skill of each student (Skinner, 1998).

Thus, to capture both the reading time and the comprehension level of participants in this study,

comprehension level (%C) and comprehension rate (%C/M) were measured as dependent

variables. The reading comprehension level (%C) was defined as the percentage of

comprehension questions answered correctly for a given passage. The %C was calculated by

totaling the number of multiple-choice questions answered correctly out of 10 and multiplying

by 10. The reading comprehension rate (%C/M) was defined as the percentage of the passage

comprehended for each minute spent reading (Ridge & Skinner, 2011; Skinner, 1998). The

%C/M was calculated by multiplying the %C value by 60 and dividing by the number of seconds

required to read the passage.

A critical requirement for an ATD is the reversibility of the dependent variables (Ledford & Gast, 2018). Within the reading comprehension literature, findings suggest that during strategic processing of information, students with LD experience more difficulty directing cognitive activities in a deliberate and reflective way (Gersten et al., 2001). For example, while reading, students may not recognize the need for active monitoring of comprehension (Berkeley & Larsen, 2018; Gersten et al., 2001; Zimmermann & Reed, 2020). Thus, they may not intuitively know to apply strategies such as self-questioning, re-reading challenging parts of text, or adjusting reading speed to enhance comprehension (Berkeley & Larsen, 2018; Firat, 2019; Gersten et al., 2001; Zimmermann & Reed, 2020).

Although students in the current study were taught reading comprehension strategies, they may not have intuitively known when and which strategies to apply when working on their own (Mason et al., 2006). For this reason, the ART and ART+PD procedures were designed to systematically guide students in the use of comprehension strategies as they engaged in reading tasks. Thus, it was expected that the dependent variables would be influenced by the application of the intervention procedures. However, once the interventions were removed, participants were expected to revert back to performance levels that might be comparable to those in the baseline phase. Finally, although multiple previous studies demonstrated the benefit of comprehension strategy application for students with LD, researchers also reported that these students may require additional time and repeated guided practice to fully grasp new procedures (Antoniou & Souvignier, 2007). Without sustained support and practice opportunities, students with LD may not utilize taught strategies independently (Harris et al., 2008).

Materials

During the baseline, training, and experimental sessions, students read passages from the

Timed Reading Series (Spargo, 1989). The expository passages covered a wide range of

nonfiction topics. Each passage had 400 words and was accompanied by ten multiple-choice

comprehension questions with three possible answer choices. Passages were randomly assigned

to conditions and sessions.

In addition to reading passages, participants were provided with the ART Student

Worksheet or the ART+PD Student Worksheet based on which intervention condition they were

assigned to for that session (McCallum et al., 2011). Participants in the ART session received

the ART Student Worksheet with directions for the three procedural steps (Ask-Read-Tell) and

fix-up strategies only. Participants in the ART+PD session were provided with a slightly

different version of the ART Student Worksheet that included all the steps described earlier with

the additional step of a discussion over the reading with peers. The first worksheet was referred

as the ART form and the second as the ART+PD form. Both versions of the form contained

explicit steps of the strategy and guided students to follow the three steps (Ask-Read-Tell) and

"fix-up" strategies (McCallum et al., 2011). The "fix-up" strategies included: 1) Reread the

paragraph, 2) Use a slow reading rate, 3) Pay full attention to reading, 4) Underline unfamiliar

words, and 5) Use context clues to help figure out unknown words. The purpose of the Student

Worksheet was to serve as a visual reminder of the steps as well as a procedural organizer to

encourage active adherence to the procedure.

Preference Assessment

Since the study spanned over an entire quarter of the school year, it was essential that

participants remained motivated and received recognition for their participation, engagement,

and academic growth throughout the research time period. Another purpose for building in

reinforcement was to encourage participants to utilize effective strategies and adhere to the ART

or ART+PD procedures with fidelity. After student assent was obtained, the primary researcher

met with all six participants to identify items, privileges, social activities considered to be

meaningful and motivating among the participants. Through collaboration with participants'

special education and homeroom teachers, the primary researcher determined which rewards

were feasible. All reinforcers were of little monetary value (i.e., edibles, Dollar Store items,

board game sessions with the primary researcher, online educational games, and homework

passes). Reinforcers were made available to participants via a token economy. Each participant

had a reward chart in his or her intervention folder. At the end of every session, participants

could earn a star for exhibiting active participation, engagement, and effort. On the last day of

the week, participants had the opportunity to trade in their stars for desirable items.   In an effort

to avoid confounding the data, reinforcement was implemented across baseline, training, and

intervention phases.

Social Validity Assessment

The Kids Intervention Profile (KIP; Eckert et al., 2017) is an eight-item rating scale that

measures students' perceptions of academic intervention acceptability and its perceived impact

on their skills. The KIP consisted of a 5-point anchored scale which ranged from 1 (*not at all*) to

5 (*very, very much*). Boxes of increasing size corresponded with each point on the scale

representing the relative strength of the responses (e.g., bigger box = higher acceptance). Test-

retest reliability for the KIP is 0.70 and validity coefficients range from 0.79 to 0.80 (Eckert et

al., 2017). For this study, the KIP was slightly modified to reflect wordings for the

ART/ART+PD intervention (Appendix A). The advantages of using the KIP include its

suitability for measuring elementary-aged students' perceptions and its flexibility for evaluating

any academic intervention. Thus, it was unlikely that these modifications impacted the technical adequacy of the measure.

The KIP was administered to collect participants' and discussion partners' quantitative acceptability data at three different timepoints: after the ART and ART+PD training phase, after four intervention sessions (2 ART and 2 ART+PD), and at the conclusion of the intervention phase. Refer to Appendix B for the KIP administration script.

Procedures

This study was conducted approximately over 10 weeks, the equivalent of a quarter in a typical school year. The study commenced upon approval of Georgia State University IRB. The primary researcher delivered training of the ART and ART+PD strategies and collected baseline and intervention data. A graduate-level researcher and one special education teacher were recruited to assist in procedural fidelity and in establishing interrater agreement for the study.

The principal researcher scheduled 30-minute sessions on two days per week for the ART and ART+PD implementation. Sessions were scheduled to fit in with participants' and discussion partners' typical school day. Sessions were canceled if participants were absent.

Determining Instructional Level

During one-on-one sessions, each participant was asked to read three passages from the Timed Reading Series (Spargo, 1989), starting with his or her current grade level. Word correct per minute (WCPM) was recorded for each passage, and median scores were used to evaluate performance at each grade level. According to Shapiro (2004), instructional level is determined when the reader is able to read the highest level at which he/she reads between 70 to 100 words correct per minute with fewer than seven errors. After each set of readings, the passage level

selections were adjusted based on participants' reading performance until instructional level criteria were met. If a participant performed at frustration level, three additional probes were provided at the next lower level until instructional level criteria were met. Passages at participants' instructional level were used during baseline, training, and intervention phases.

Baseline

In the baseline phase, each participant was presented with a copy of a randomly selected reading passage at his/her instructional level. The researcher read from a script (Appendix C) to maintain standardized administration of baseline procedures. Participants were instructed to read aloud at their normal reading pace. As participants read, the researcher kept time and documented the total time required for participants to complete reading each passage. The total time in minutes was converted to seconds to be used for subsequent calculation of the comprehension rate (%C/M). After completing the passage, participants were asked to answer ten multiple-choice (MC) comprehension questions about what they had read. The passages were collected, and participants were not able to refer back when answering multiple-choice questions. The number of correctly answered MC questions was used to calculate the comprehension level (%C). To calculate the comprehension rate (%C/M), the %C value was multiplied by 60 then divided by the total number of seconds required for participants to read the passage. To establish interrater agreement, the special education teacher was provided with copies of completed ten MC comprehension questions to grade independently from the primary researcher to establish interrater agreement. For each participant, five baseline data points were collected for comprehension level and rate.

Training Phase

Training of the intervention procedures took place after baseline. This was done in an effort to avoid confounding the baseline data. During the training phase, the primary researcher trained all participants in the ART procedure across two 30-minute training sessions. The training procedures for the ART procedure are outlined in Appendix D. Thereafter, participants and discussion partners received training in the ART+PD for two 30-minute training sessions. Appendix E outlines the ART+PD training procedures.

ART

In the ART condition, participants were given a copy of a randomly selected passage at their instructional level. At the same time, they were given the ART Student Worksheet (Appendix F). Once all materials were distributed, the researcher oriented participants to the ART procedure by reviewing each step and instructed participants to use the Student Worksheet to help guide their reading. Next, the participants were directed to read aloud their passage and to complete the ART Student Worksheet as they read. Once participants finished reading their passage and completed their ART Student Worksheet, they were asked to answer ten MC comprehension questions. As participants read, the researcher kept time and documented the total time required for participants to complete reading each passage. As in the baseline phase, the number of correctly answered MC questions and total reading time were used to calculate comprehension level and rate respectively for each participant. The special education teacher graded a copied set of comprehension questions for interrater agreement purposes. A script for the ART implementation procedures is presented in Appendix G.

ART+PD

In the ART+PD condition, all procedures were identical to the ART condition. However, after participants completed the ART+PD Student Worksheet (Appendix H), they were paired up

with a randomly assigned partner to discuss what they learned from the passage, share their questions from Step 1 and the answers to their questions (Step 3). They were also asked to talk about any interesting information from the reading. The primary researcher facilitated the discussion portion of the intervention. A script for the ART+PD implementation procedures is presented in Appendix I.

Interobserver Agreement

Interobserver agreement (IOA) was calculated on 31.03% of reading comprehension question sets, and IOA was also calculated for 31.03% of the timed reading sessions. IOA was calculated by dividing the number of agreements by the sum of agreements and disagreements and multiplying by 100 (Ledford & Gast, 2018). An IOA of 94.44% for comprehension question scoring and an IOA of 88.89% for the reading times were achieved.

Procedural Fidelity

All sessions were video-taped. To ensure procedural fidelity, all sessions were implemented according to the checklist outlining required procedures specific to each intervention condition (i.e., ART, ART+PD) (Appendix J). The primary researcher implemented each component on the checklist and checked off each intervention component as it was completed. Procedural integrity was monitored through both self-observation (i.e., by the primary researcher) and through ratings from a second observer. A graduate-level researcher served as the second observer and viewed 40.00% of the taped Baseline sessions, 37.50% of the taped ART sessions, and 47.06% of the taped ART+PD sessions. The second observer marked off on a procedural checklist while observing each session. Procedural integrity was calculated by dividing the number of steps completed correctly by the total number of steps required of

each condition during a given session and multiplied by 100 to obtain a percentage. For all sessions, procedural integrity equaled 100%.

Data Analysis

Visual Analysis

Systematic visual analysis procedures are key in evaluating characteristics of data patterns and determining the presence of experimental control (Ledford & Gast, 2018). An advantage of visual analysis is the ability to detect intervention effectiveness over time, across conditions, and across multiple participants (Ledford & Gast, 2018). Hence, intervention effects were evaluated using visual analysis of data characteristics within and across conditions that included trend, level, and variability (Ledford & Gast, 2018). Specifically, comprehension level (%C) and comprehension rate (%C/M) were directly compared between baseline and implementation of the ART and the ART+PD interventions. Specific to ATD, the degree of differentiation between data paths was evaluated to determine whether there was a consistent difference in level between adjacent data points from different conditions (Ledford & Gast, 2018). Finally, visual analysis was used to draw conclusions about the presence of functional relations. (Ledford & Gast, 2018).

Nonoverlap of All Pairs (NAP; Parker & Vannest, 2009)

In addition to visual analysis, an effect size analysis using Nonoverlap of All Pairs (NAP; Parker & Vannest, 2009) was calculated to further support interpretations from the visual analyses. NAP is a nonparametric effect size used to measure the effect of ART and ART+PD on each dependent variable. Within NAP, all data points during the intervention were compared with all baseline data points for overlap to provide a valid effect size (Parker & Vannest, 2009). Interpretations of NAP results are as followed: weak effects: 0-.65; medium effects: .66-.92;

large effects: .93-1.0 (Parker & Vannest, 2009). In the current study, NAP was calculated by comparing baseline to intervention phase data points for both comprehension level and rate for each participant.

Phase 1: Results

Table 2.1 lists the mean percentages for each dependent variables in the baseline and intervention phases along with the effect sizes for each intervention for all three participants. NAP was calculated to compare each individual datum in baseline to each datum in each intervention phase. For this study, the baseline was used as the primary comparison condition to each intervention condition (e.g., ART or ART+PD). Overall, participants demonstrated an increase on comprehension level and rate over baseline for both ART and ART+PD procedures. For Bryson and Cole, ART+PD proved more effective than ART in increasing comprehension level and rate, whereas ART was more effective in raising Charlotte's level and rate. Furthermore, NAP calculations indicated that Bryson had greater ART+PD effect sizes for level and rate than Cole.

Bryson

Comprehension Level (%C)

Immediately following the implementation of the ART intervention, Bryson's %C increased from 30.0% (last baseline data point) to 70.0% (first ART intervention data point) (Figure 2.1). There was also an increase in the mean %C from baseline phase ($M = 54.4\%$) to intervention phase ($M = 74.0\%$), demonstrating a level change between the two phases. The decreasing baseline trend was a clear contrast to the flat trend of the ART phase. Based on visual analysis of the ART procedure, there appeared to be an increase in Bryson's comprehension level. Finally, baseline data appeared to have moderate variability (range: 30.0-

70.0%) when compared to the more stable ART phase (range: 70.0-90.0%). Visual analysis indicated experimental control since clear divergence between baseline and ART intervention phases with regard to level, trend, and variability were established.

Figure 2.1 reveals an immediate increase in Bryson's %C from the last baseline data point of 30% to the first intervention data point of 70% after ART+PD was introduced. A rise in level was evident when comparing the means of baseline (M = 54.0%) and ART+PD phases (M = 81.67%). The decreasing baseline trend was a distinct contrast from the increasing ART+PD trend. Because clear separations between the levels and trends among the phases were evident, experimental control appeared to have been established. Regarding data stability, both baseline (range: 30.0-70.0%) and ART+PD (range: 70.0-100.0%) phases were similarly moderate.

Statistical analyses were used to further support visual analyses. In this instance, the percentage of nonoverlapping data for ART+PD (97.0%) was greater than ART (92.0%), suggesting that ART+PD was slightly more effective in raising Bryson's comprehension level. Moreover, ART+PD had a large effect (Parker & Vannest, 2009) on increasing Bryson's comprehension level.

Comprehension Rate (%C/M)

In Figure 2.2, Bryson's %C/M increased from 4.38% (last baseline data point) to 8.09% (first intervention data point) following the application of ART. Further, the ART %C/M mean (M = 10.65%) was higher than the baseline %C/M mean (M = 8.86%), an indication that ART was effective in raising his comprehension rate. The increasing trend in the ART phase diverged from the baseline's decreasing trend. Regarding data stability, the baseline data were moderately variable (range: 4.38-13.46%), while the ART phase showed fairly stable %C/M scores (range: 7.30-17.76%) with one outlier during the third ART session when %C/M increased to 17.76%.

Overall, the visual analysis showed that experimental control was established based on demonstrated separations between baseline and intervention phases in level, trend, and variability.

Following the application of ART+PD, %C/M level increased from 4.38% (last baseline data point) to 7.61% (first intervention data point) as shown in Figure 2.2. There was also an increase in the mean %C/M from baseline phase (M = 8.86%) to the ART+PD phase (M = 12.33%). The decreasing trend in baseline was a distinct contrast to the increasing trend during the intervention phase. Although variability among baseline data points was moderate (range: 4.38-13.46%), ART+PD data points (range: 7.61-15.79%) appeared to be of low variability. Finally, experimental control was evident based on visual distinctions between baseline and intervention phases in level, trend, and variability.

Based on NAP calculations, ART+PD (0.80) was shown to be moderately effective in raising Bryson's comprehension rate, whereas ART (0.68) was shown to be less effective (Parker & Vannest, 2009).

Charlotte

Comprehension Level (%C)

Charlotte experienced a slight decrease in level from the last baseline data point of 60.0% to the first ART intervention data point of 50.0%. On the day of the first ART session, Charlotte reportedly did not take her prescribed ADHD medication which might have contributed to the decline in her comprehension performance. She experienced significant difficulty attending to her reading, skipped multiple rows of text, and frequently lost her place during her read-aloud portion of the ART procedure. Nonetheless, an increase in the mean %C was evident when comparing ART phase (M = 83.3%) to the baseline phase (M = 74.0%). The decreasing baseline

trend was a marked contrast to the ART increasing trend. Compared to the moderate variability

in the baseline phase (range: 60-90%), the data points were more stable across the ART phase

(range: 80.0-100.0%), excluding one outlier (50%) during the first session of the ART phase.

Overall, the graphs in Figure 2.3 suggested evidence of experimental control as clear separations

between level, trend, and variability across the baseline and ART phases were demonstrated.

As shown in Figure 2.3, the %C level increased from the last baseline data point of 60%

to the first data point of 70% following the application of the ART+PD intervention.

Additionally, the ART+PD mean %C (M = 76%) was slightly higher than that of the baseline (M

= 74%). The baseline phase's decreasing trend was in contrast to the increasing trend of the

ART+PD phase. In terms of variability, the baseline data (range: 60.0-90.0%) were less stable

when compared to those in the intervention phase (range: 70.0-80.0%). Data suggested that

distinct separations between baseline and ART+PD phases in level, trend, and variability were

observed.

Based on NAP calculations, ART (0.72) was moderately effective in raising Charlotte's

comprehension level when compared to ART+PD (0.52), which demonstrate a weak effect size

(Parker & Vannest, 2009).

Comprehension Rate (%C/M)

As presented in Figure 2.4, Charlotte's %C/M increased slightly from 10.62% (last

baseline data point) to 11.11% (first intervention data point) following the application of ART.

Additionally, the ART %C/M mean (M = 19.80%) was higher than the baseline %C/M mean (M

= 16.20%), suggesting that ART was effective in increasing Charlotte's comprehension rate. A

divergence in trend was evident when comparing the ART increasing trend to the baseline

decreasing trend. During the first ART session, Charlotte's performance was most likely

impacted due to a change her ADHD medication routine. As such, the intervention phase was

characterized by variable %C/M scores (range: 11.11-25.86%) whereas baseline scores (range:

10.62-21.62%) were less variable. Based on the visual analysis of the graphs across the baseline

and ART phases, level, trend, and variability were visually distinct from one

another.

Following the implementation of the ART+PD, Charlotte's %C/M increased from

10.62% (last baseline data point) to 13.64% (first ART+PD data point). However, the ART+PD

%C/M mean ($M = 15.00\%$) was lower than that of the baseline ($M = 16.20\%$), signifying a

decline in comprehension rate during the ART+PD phase. The decreasing trend during baseline

was in contrast to the gradually accelerating trend during the ART+PD phase with one outlier

during the 3rd ART+PD session when %C/M decreased to 7.28%.   The variability in baseline

data was moderate (range: 10.62-21.62%). However, the ART+PD variability (range: 7.28-

18.90%) was relatively high. Figure 2.4 presents graphs across the phases.

The NAP calculation for ART (0.70) indicated a moderate effect size compared to the

small ART+PD effect size (0.52) in raising Charlotte's comprehension rate.

Cole

Comprehension Level (%C)

Figure 2.5 displays Cole's comprehension levels across various phases. Immediately fol-

lowing the implementation of the ART intervention, Cole experienced an increase in %C from

60.0% (last baseline data point) to 100.0% (first ART intervention data point). There was also

an increase in the mean %C from the baseline phase ($M = 78.0\%$) to the ART phase ($M =$

82.0%)., demonstrating a level change between the two phases. The trends during the baseline

and the ART phases were both decreasing. Finally, data in both baseline (range: 60.0-90.0%)

and ART (range: 70.0-100.0%) phases appeared to have moderate variability. With the exception of the separation in levels, it is unclear whether trend and variability were distinctly differentiated between the baseline and ART phases.

Cole had an immediate increase in %C from the last baseline data point of 60.0% to the first intervention data point of 90.0% after ART+PD was introduced. There was also a level increase when comparing the baseline mean ($M = 78.0\%$) to the ART+PD mean ($M = 93.3\%$). The decreasing trend in the baseline phase was in contrast to the flat trend in the ART+PD phase. Compared to the stable set of ART+PD data points (range from 90.0-100.0%), baseline data were more variable (range: 60.0-90.0%). Because clear separations between the levels, trends, and variability between the phases were evident, experimental control appeared to have been established (Figure 2.5).

For Cole's comprehension level (%C), the ART+PD effect size (0.87) was greater than the ART effect size (0.58), suggesting that ART+PD was noticeably more effective in raising his comprehension level. Moreover, ART+PD also had a medium effect on increasing Cole's comprehension level.

Comprehension Rate (%C/M)

Cole's comprehension rates across various phases are presented in Figure 2.6. After having received the ART intervention, Cole increased his %C/M from 10.65% (last baseline data point) to 16.39% (first ART data point). However, his mean %C/M decreased from the baseline phase ($M = 13.03\%$) to ART phase ($M = 12.51\%$). The decreasing trend was demonstrated in both baseline and ART phases. Finally, variability in both phases appear to be low (Baseline range: 10.65-16.36%; ART range: 10.77-16.39%). Overall, there were not distinct separations in level, trend, and variability between baseline and intervention phases.

Cole experienced a slight increase in %C/M after having received the ART+PD intervention; his %C/M increased from 10.65% (last baseline data point) to 11.44% (first ART+PD data point). There was also a small increase in the mean %C/M from baseline phase ($M = 13.03\%$) to the ART+PD phase ($M = 14.59\%$), suggesting a level change between the two phases. Further, the baseline and ART+PD trends diverged in opposite directions, whereas the fluctuation among data points in both phases was similarly low (Baseline range: 10.65-16.36%; ART+PD range: 11.44-18.29%).

NAP calculations supported ART+PD (0.67) as the more effective intervention in raising Cole's comprehension rate when compared to ART (0.40). Based on Parker & Vannest's tentative ranges, (2009), ART+PD had a moderate effect on Cole's comprehensive rate.

Social Validity

Three fourth-grade participants and three fourth-grade discussion partners were administered the KIP to obtain an estimate of intervention acceptability at three different timepoints: after the training phase, after four intervention sessions (2 ART and 2 ART+PD), and at the conclusion of the intervention phase. Table 2.2 presents means and standard deviations for KIP items across the three timepoints.

In general, ratings across the three timepoints for the majority of the KIP items remained consistent with the exception of Items 4 and 7. During the first ($M_1 = 3.00$, $SD_1 = 1.41$) and second ($M_2 = 3.50$, $SD_2 = 1.38$) KIP administrations, students reported that they wished they could work more on the intervention to improve their reading comprehension for some of the time. However, by the end of the study, students indicated that they wished they could work more to improve their comprehension a lot of the time ($M_3 = 4.00$, $SD_3 = 1.41$). Furthermore,

students reported that their reading comprehension improved a lot (M_1 = 4.67, SD_1 = 0.82) after

completing the intervention training but later indicated that their comprehension only improved

some (M_2 = 3.83, SD_2 = 1.17). At the end of the study, students concluded that their reading

comprehension had improved a lot (M_3 = 4.50, SD_3 = 0.55).

Based on ratings of other KIP items, students reported that they enjoyed learning

comprehension skills each week (M_1 = 4.33, SD_1 = 0.52; M_2 = 4.67, SD_2 = 0.52; M_3 = 4.67, SD_3 =

0.82) and liked being taught how to read in order to improve their reading comprehension (M_1 =

4.17, SD_1 = 0.75; M_2 = 4.17, SD_2 = 1.17; M_3 = 4.50, SD_3 = 0.84). At no time point, did they

indicate they did not want to learn comprehension skills (M = 1.00, SD = 0.00 for all timepoints).

Further, students reported they very much liked receiving feedback on their reading time and

their performance on the MC comprehension worksheet (M_1 = 4.00, SD_1 = 1.15; M_2 = 4.83, SD_2

= 0.41; M_{3} = 4.83, SD_3 = 0.41) and they indicated that having this feedback was very helpful (M_1

= 4.67, SD_1 = 0.52; M_2 = 4.67, SD_2 = 0.52; M_3 = 4.50, SD_3 = 0.55). Finally, students reported

that their reading comprehension has not been negatively impacted by the intervention (M_1 =

1.17, SD_1 = 0.41; M_2 = 1.00, SD_2 = 0.00; M_3 = 1.17, SD_3 =

0.41).

Phase 2: Qualitative Study

In the current school district, the ART strategy is recommended as a reading

comprehension intervention for students receiving MTSS services. Nonetheless, there appears to

be a paucity of research associated with students' acceptability and perceptions of their

experiences as consumers of academic interventions. The collection and reporting of social

validity data on the ART strategy were intended to provide a more in-depth understanding of

students' perspectives, feelings, and experiences during the acquisition and application phase of

the ART strategy in their classrooms. The present study contributed to the current literature on social validity of school-based intervention practices in two primary ways. To the best of the researcher's knowledge, this was the first study to explore the acceptability of the reading strategy Ask-Read-Tell (ART) among upper elementary students with LD. Second, the study examined the total construct of social validity within the context of an elementary school setting. In addition to obtaining acceptability data from the KIP, the researcher gathered social validity data via semi-structured interviews.

Traditionally, social validity is studied using questionnaires or rating scales (Finn & Sladeczek, 2001). Despite their usefulness, quantitative rating scales and surveys can be restricted in the amount and type of information they provide (Leko, 2014). Due to the nature of these instruments, potential information deemed important by participants may be underrepresented or excluded altogether. Furthermore, most surveys contain dichotomous (i.e., yes/no) or ordinal (i.e., Likert-type scales) items and may not fully capture the complex array of experiences from participants' perspectives (Leko, 2014).

Given that a large number of social validity studies are steeped in the quantitative tradition, the purpose and nature of the qualitative component of this study were to allow for holistic examinations of the perspectives of school-aged children with regard to intervention objectives, procedures, and skill outcomes (Creswell & Creswell, 2018; Crotty, 2015; Leko, 2014). Further, this methodology supported situating participants' voices at the forefront of the study (Creswell & Creswell, 2018; Crotty, 2015; Leko, 2014). Specific to this study, data collected through interviews from students who had learned and applied the reading comprehension strategy were expected to provide important themes of social validity that may be overlooked in studies that rely solely on quantitative procedures (Creswell & Creswell, 2018).

Using qualitative methods to "capture unintended impacts….and illuminate dimensions of desired outcomes that are difficult to quantify" (Patton, 2002, p. 152), allowed the researcher to develop a richer understanding of social validity as a total construct.

Phase 2: Research Question

Phase 2 of the current study investigated the following research question:

1. What are participants' perceptions of social validity about the ART intervention regarding procedures, effectiveness, and relevance for supporting their reading performance?

Participants

All participants and discussion partners from Phase 1 of the study were selected for in-person interviews during Phase 2.

Setting

All interviews were conducted at the research site described in Phase 1.

Data Source

After obtaining assent from participants and discussion partners, the researcher conducted semi-structured interviews with each student in person. Interview questions are presented in Appendix K. Interviews were conducted in a private space to maintain confidentiality. Additionally, interviewees were assured that their responses would not be linked to them, and their identity would be protected during data collection and reporting of the study findings. Students were asked questions about their thoughts, perceptions, and experiences related to the acquisition and application of the ART intervention. The interviews lasted between 15 to 30

minutes. All interviews were audio recorded and transcribed verbatim by Otter.ai, an online transcription service.

Data Analysis

Consensual Qualitative Research

Following the transcription of all interviews, a bracketing process occurred to identify and address biases and assumptions about the study's focus (Hays & Singh, 2012). Next, the coders immersed in the transcripts prior to starting the data analysis. Consensual qualitative research (CQR) methodology was applied to analyze transcribed content and build consensus systematically regarding themes grounded in the data. The CQR process involved incorporating multiple perspectives, consensus among team members, and the continual return to the raw data during the coding process (Hill & Knox, 2021).

The data analysis team consisted of the primary researcher and another doctoral candidate, serving as coders and one Ph.D. faculty member, serving as the auditor. The auditor reviewed and supervised the coding and thematic development. Through the consensus-building process, the coders primarily determined the broad domains, summarized domain data into concise terms, and conducted cross-analyses to develop common patterns in the findings. Before the coding commenced, the coders received CQR training consisted of readings in the professional literature and practice sessions with multiple interview data sets.

Domains

To begin the domain coding, the coders created a list of domains derived from the interview questions. Next, each coder independently reviewed all six transcripts and made recommendations for modifications of domains. For example, indistinct domains were combined, or new domains were created to represent unexpected information. The coders came

together and worked to achieve consensus on the list of domains. The list of domains was sent to the auditor for review. Per the auditor's suggestion, the domains were further combined to achieve a final list of seven domains.

Next, the coders worked as a team to code the transcribed data into domains. The coders took turns reading blocks of data out loud until the topic changed. Thereafter, each coder made suggestions about which domains were most appropriate and voiced rationales for suggestions. The coders discussed differences of opinion and arrived at a consensus regarding the most appropriate domains for the data. On several occasions, some data were coded into more than one domain. NVivo 12 was used to organize data by domains.

Core Ideas

The next coding phase involved distilling down the interviewees' words into concise, clear units of ideas that were as close to the data as possible. To derive at core ideas for each chunk of text, the coders worked together to eliminate repetitions and nonrelevant aspects of interview responses. Similar to the domain development process, the coders took turns reading blocks of responses aloud and stating core ideas that best captured the essence of what interviewees said in fewer words and with more clarity. Differences in opinions and rationales were discussed until a consensus was reached. Core ideas for each domain were organized in Excel spreadsheets and were forwarded to the auditor for review. Based on the auditor's review and feedback, some domains were combined, resulting in five primary domains.

Cross-Analysis

In the cross-analysis, the coders worked together to achieve a higher level of abstraction and to identify thematic similarities (categories) and dissimilarities within domains across interviews. The coders rotated reading the core ideas out loud and proposed themes and patterns

that emerged from them. Examining the themes and patterns, the coders created categories. Categories were grouped based on thematic similarities and further refined into subcategories when there were subtypes within categories. The cross-checking and consensus-building process was used as previously described by the coders and auditor.

Data Quality

The following steps were taken to ensure credibility and trustworthiness. During the interview process, member checks were conducted with interviewees to verify the researcher's understanding and interpretations (Hays & Singh, 2012). Further, the researcher engaged in peer debriefing following interviews (Hays & Singh, 2012). During data analysis, the researcher and the second coder worked to remain reflexive and looked for exceptions and contrary evidence within the interview data. Additionally, triangulation of data sources involved the inclusion of several student voices and comparison of themes from the qualitative inquiry to quantitative response patterns of the KIP measure (Hays & Singh, 2012).

Ethical Considerations

In the current study, ethical principles were at the forefront during its planning and implementation. To this end, steps were taken to ensure that informed consent was sought during which participants were explained in understandable language the nature of the study, any foreseeable risks and benefits, limits of confidentiality, and the voluntariness of participation (Hays & Singh, 2012). It was especially important that explicit explanations were provided to assure participants that their decision to participate or not participate would have no impact on their grades nor result in disciplinary action. Another ethical concern related to the potential coercive nature of qualitative inquiry and the researcher's role as an internal researcher (i.e., intern school psychologist) at this site. For this reason, the researcher revisited the consent

process with participants throughout the study and looked for subtle or gestural reluctance on the part of participants as potential indicators of withdrawal of consent.

Phase 2: Results

Five domains emerged from the CQR data analysis: Reading/Learning Barriers, Importance and Impact of Reading Skills, ART/ART+PD Generalization and Impact, Acceptability/Satisfaction, and Recommendations. Descriptions of domains and their related categories and subcategories along with participants' excerpts are presented here. Sentence fillers (i.e., "um", "like", "you know", and "kinda") from quotations were removed to maximize clarity for the reader.

Domain 1: Barriers to Reading and Learning

Representing this domain is the summary of interviewees' perceptions of the reasons why reading was difficult for them and/or their peers and how these difficulties negatively impacted learning. Several interviewees stated that decoding unfamiliar words was hard, as one student shared, "Reading some of the words makes it difficult." Some interviewees also reported that while reading, they often became distracted and frequently lost their place. In describing her tendency to become distracted while reading, one student noted, "Sometimes whenever I go to read, I look at a word. And then I look up the page or look at something else cuz maybe someone tried to talk to me, and I go to look back, and I'm on a different sentence that I wasn't on." Others cited that they could generally decode words but had trouble understanding what they read, especially when reading long passages. One student explained, because reading passages are "very long… some students don't understand what they read."

Due to reading difficulties, participants reported taking more time to complete schoolwork and experiencing more problems in classes with heavy content materials (i.e., ELA,

social studies, and science). Additionally, participants reported that reading difficulty negatively affected independent test-taking since misreading words on test questions might alter how they answered. As such, they reported feeling stressed and anxious when taking tests due to worries about responding incorrectly and getting poor grades. One student commented, "Because whenever you read a passage, and you have to do comprehension questions, and you don't understand it, you're probably going to get them wrong. And if you get them wrong, that's like a lot of pressure…. sometimes that can get to you." Another student shared, "(Students) could not understand the questions, and they get really overwhelmed or stressed out. They would probably think that they're not going to get a good grade."

The final barrier to learning was related to the classroom learning climate. Several interviewees indicated that they did not always feel supported and were hesitant to go to teachers for help. Specifically, while doing classwork or taking a test, some encountered reading obstacles but were afraid to speak out. In describing this dilemma, one student commented, "Some students don't necessarily understand what they read, but they're too afraid to speak out about it. And they try to ask their friends, and the teacher will say, you can't talk or do anything. So that doesn't necessarily help if they're too scared to go to the teacher."

Domain 2: Importance and Impact of Reading Skills

This domain summarized participants' perspectives on the importance of reading skills, the impact of reading skills on academic and life goals, and participants' attitude about reading. In general, participants believed reading skills were important for the following reasons: to access curriculum content, to better comprehend text, to achieve reading goals, to perform well on assignments and tests, to earn better grades, to lessen performance anxiety, to gain more independence when doing schoolwork, and to expand their knowledge. Further, participants

reported that having good readings skills would help them do well in college and in their future jobs. They also reported that adequate reading skills would enable them to improve their future in general. When responding to why reading skills were important, one student shared, "The more you understand what you're reading…. the skills will help you improve in the future." Another student stated, "I think if I get a job where I have to read a lot, or I get into college, and I have to read a whole bunch of textbooks, then it will most likely help me."

Regarding interviewees' attitudes about reading, one student stated that reading was frustrating for him. Conversely, others expressed that reading was pleasurable and provided them with an outlet. One noticeable change in attitude was related to interviewees' shift away from reading quickly, and instead, they recognized the importance of reading for meaning. Several interviewees shared that they used to perceive reading as a competition with peers and focused on reading for speed. However, after the ART intervention, they preferred to comprehend their materials thoroughly even if it took them longer. In sharing about her change in attitude, one student noted, "In third or second grade, I used to always, I don't know why I thought it was a competition, I always would race through the reading and trying to read too fast. But now I just take more time and try to understand it."

Domain 3: Generalization of ART/ART+PD and Impact

This domain summarized how interviewees applied ART/ART+PD strategies in settings beyond the intervention sessions and the impact of these strategies on their reading and learning. Regarding generalization of ART/ART+PD, interviewees reported using some aspects of the intervention when doing homework/classwork and when reading for English/language arts, science, and social studies. When reading for social studies, one student shared, "Sometimes in social studies when we're reading alone or doing the questions, it helps me with the harder words

that I'm less familiar with. And I use the skills sometimes to help me understand a word or a question better." Another student noted, "Right now it really helps me with language arts." One student said she used some of the fix-it strategies by underlining unfamiliar words and using context clues to help her figure out vocabulary definitions in class. She commented, "Strategies like context clues…. underlining words that you don't know; I think they help a lot…. cuz in some words that I've read, it's really helped me to actually understand the words." In addition to classwork, several interviewees indicated using ART strategies when they read for pleasure: "I have been using the strategies. Yeah, that's how I'm reading the chapter book."

Two interviewees shared that they had begun looking at chapter titles and mentally generated some questions from them. They also remembered to reread parts of their books that were unclear. One student said he remembered to slow down his reading when he had trouble understanding the text, and this fix-it strategy had helped him to remember more of what he read. In describing his reading, one student related, "I like rereading, and slowing down is better because sometimes when I want to go really fast through a book, I hardly understand anything. And I'm just gonna go back over that. And when I go over it slowly, then I can understand better."

Concerning the impact of ART/ART+PD on learning, interviewees reported that the intervention had helped in several ways. First, ART/ART+PD was perceived as helping some interviewees stay more focused while they read. Second, one student noticed improved performance on classwork because he was taking his time previewing and reading passages. He related that ART strategies helped him to be more thorough with his reading which ultimately helped him to answer comprehension questions. Finally, another student noticed increased independence when using the ART/ART+PD strategies during reading assignments because he

relied less on his parents for help. One student shared, "I need less help in reading from my parents," and another student concurred, "I can read the problems. People don't have to read for me to help me read it."

In addition to the positive impact on learning, ART/ART+PD appeared to confer various positive impacts on reading behaviors. For example, several interviewees stated that the intervention procedures prompted them to slow down their reading, to examine unfamiliar words more closely, to use context clues in defining new vocabulary words, and to direct their thinking about the passage content as they read. One interviewee shared that applying ART/ART+PD strategies helped her to take responsibility for her own comprehension and noted, "It helps me to know that if I don't understand something, I should be honest and go back and read it." Although the intervention was not designed to improve decoding skills, one interviewee reported that by slowing down his reading, he was able to take the time to decode challenging words. This interviewee stated, "If I don't know how to pronounce something, I'd be able to go back and try to read it, then I'll move on to the next sentence." Overall, interviewees reported acquiring a new approach to reading along with incremental improvement on their comprehension skills. An interviewee spoke about "slowing down and not just rushing through to try and finish quicker than everybody." One interviewee noticed a slight change in her reading level and commented, "I think it hasn't done a massive difference, but I feel like I'm not at the same level that I was before I was doing this. So yeah, I'm at a slightly higher level." Multiple participants also shared that they were able remember more of what they read.

Domain 4: Acceptability/Satisfaction

This domain focused on interviewees' perceptions of intervention procedures, outcomes, and reading goals. Additionally, it summarized factors that contributed to interviewees'

acceptability and general satisfaction with the ART and ART+PD interventions. Some interviewees were dissatisfied with Steps 1 and 2 for the following reasons: difficulty with generating questions based on passage titles, vagueness of passage titles, the task of physically marking paragraphs with +/- signs when monitoring one's comprehension was cumbersome, and the length of the passages. On the other hand, a few interviewees reported no negative aspects of the interventions and did not have a preferred step within the ART/ART+PD procedures. Participants also indicated that their reading comprehension had not been negatively impacted by the interventions.

While there was mild dissatisfaction about a few aspects of the interventions, the majority of interviewees stated feeling very satisfied (average of 9.7 out of 10, with 10 being extremely satisfied) since the interventions were helpful with improving overall reading comprehension, with classwork requiring reading, and in working towards a reading goal. One interviewee simply said, "I actually liked it. It actually helped me." Further, he stated feeling optimistic that the intervention might help him acquire reading skills that were closer to his grade-level and commented, "It helps me figure out how to read and maybe help me get up to a fifth-grade reading level, hopefully before fifth grade." Others believed learning the ART/ART+PD strategies would equip them with self-help skills that would lead to independent reading. Further, interviewees described the ART/ART+PD as fun, interesting, and enjoyable.

More specifically, some interviewees reported liking Step 1 in which they examined the title and generated two questions based on the title. They believed that Step 1 provided them with a general idea of the passage and a "hook' that piqued their interest about the passage. In sharing her perspective about Step 1, one student said, "I like it because it helps me…to get really interested and hooked in the passage. Makes it more interesting." Further, Step 1 was

cited as being helpful for interviewees in identifying the main idea of each passage.

Interviewees stated that the fix-it strategies in Step 2 helped them to slow down their reading, to monitor their understanding of text, and to reflect honestly about their own comprehension. In addition, the +/- markings aided them with visual tracking of each paragraph and served as reminders for them to go back and reread parts of passages. In his response to why he liked marking paragraphs, one interviewee said, "I think it does help me because if I don't understand something, the marks are right next to me to tell me."

Finally, participants indicated that the discussion component of the ART+PD enabled them to share and to hear others' opinions about the passages. In addition to the exchange of ideas, students said they gained new perspectives from engaging in discussions, enjoyed the chance to connect socially with others, and gave them the chance to help their peers. When sharing his opinions about the discussion part of ART+PD, an interviewee said,

> I think that's good because when you can talk with a partner. It helps you study, and it helps you get better at reading when you're talking with somebody else. They also understand when you understand, and you can see the other person's perspective about things. And you can think, hey, what if I look at it this way?

Interviewees also reported that the discussion generated increased their interest in the passages and helped them to answer the ten comprehension questions. "I like that (step) because it helps me with the questions. Also, it makes me interested in what they have to say. And sometimes their questions helped me get an answer right on the test after," one interviewee noted in explaining why she enjoyed the discussion of passages.

An unexpected factor that enhanced acceptability was related to interviewees' connection with the interventionist (primary researcher) and the learning environment. One interviewee

reported feeling calm when the interventionist worked with her and shared, "I feel very calm and not worry…. I have someone to help me by my side instead of just working by myself." Another student stated that he could trust the interventionist. He commented, "I like how you say at the beginning, we can be totally honest with each other, no one's gonna see this." Others described the learning environment as safe, confidential, and positive. In fact, being in a safe environment encouraged students to ask for help, as one interviewee indicated, "I like how you aren't one of those teachers who say, 'If you need help, then just ask me a question about the word' and then when I want to ask, it's like, "no, I can't help you.'"

Lastly, one student reported that earning reinforcers motivated her to do well and gave her something to look forward to at the end of the week. When describing her excitement about reinforcers, she noted, "I like how you keep us excited with the prizes on Fridays, and this keeps us working hard and trying our best."

Domain 5: Recommendations

This domain highlighted interviewees' recommendations regarding future implementation of the interventions and suggestions for improvement of ART/ART+PD.

To make ART/ART+PD procedures easier and better for fourth-graders, interviewees suggested reducing the number of questions in Step 1, reducing the number of multiple-choice questions, arranging these questions in the same sequential order as passage content, and shortening reading passages. One interviewee who experienced inattention and high distractibility suggested leaving off the physical +/- markings in Step 2 as they were visually distracting for her. She commented, "Honestly, I don't like the plus and minus coding. It makes it kind of harder because it makes me not focus. I would rather do it in my head… like I imagine it's there."

All interviewees expressed the desire to continue using ART/ART+PD for themselves and recommended that other students learn about the procedures. They stated that this reading intervention would help other students with their reading comprehension. One interviewee shared his perspective, "I would recommend it because if some people have difficulty understanding things, or they just can't remember or anything. I think just using this could help them understanding more."

Discussion

The Effects of ART and ART+PD

The primary goals of the study were to compare the effects of ART and ART+PD on the reading comprehension level and rate relative to the baseline and relative to one another in fourth-graders with learning disabilities.

When evaluated as a whole, both ART and ART+PD interventions were effective in raising participants' comprehension level and rate in comparison to the baseline. As such, these results differ from the McCallum et al., findings (2011) in that no significant differences were found between the control and the ART conditions. Nonetheless, results from this study are consistent with the body of research that supports the use of multiple strategies that include active engagement with the text throughout the reading process, monitoring one's understanding, asking questions, thinking aloud, engaging in discussion about the reading, and utilizing a procedural organizer with strategic steps for students with reading comprehension difficulties (Baker et al., 2002; Firat, 2019; Idol-Maestas, 1985; Mason et al., 2006; National Reading Panel, 2000; Ridge & Skinner, 2011). Moreover, visual analyses in the current study suggested that gains made on comprehension level and rate were generally maintained to some degree over time. Further, findings revealed that there was a clear differentiation between the effectiveness

of the ART and ART+PD conditions for each student.

In the cases of Bryson and Cole, evidence from visual and statistical analyses indicated that ART+PD was more effective than ART in raising their comprehension levels and rates. This finding is consistent with the McCallum et al., result (2011), revealing that the ART+PD produced significantly higher scores than the ART. Furthermore, results from the current study align with the current literature supporting the use of discussion-based interactions for improving comprehension in elementary students. In fact, positive effects were evident in studies in which students were given opportunities to participate in activities that encourage them to resolve ambiguities in text, think critically "aloud", vocalize opinions, and to formulate conclusions or inferences about the text (Shanahan et al., 2010).   Furthermore, peer interaction has been shown to have a positive impact on comprehension as students discuss their ideas and assist one another in answering questions related to the readings (Mahdavi & Tensfeldt, 2013).

Of the two students, Bryson responded more favorably than Cole to the ART+PD, and there may be possible reasons accounting for this difference. Although both boys exhibited the motivation to improve their comprehension skills, Cole's moods appeared more variable. There were sessions during which he seemed frustrated and/or withdrawn due to incidents that had occurred earlier in his school day. At other times, he stated feeling fatigued from staying up late the night before. As such, Cole's low moods and/or fatigue appeared to negatively impact his engagement during the discussion component of the ART+PD. Bryson, on the other hand, presented with an even disposition and put forth consistent effort in his discussions of the text. Because Bryson was fully engaged in the discussions on a consistent basis, he appeared to have profited more from this aspect of the ART+PD intervention. Additionally, at the time of the study, Cole's reading level (2.9 grade-equivalency) as measured by the Winter Renaissance

STAR Reading benchmark was slightly lower than that of Bryson (3.1 grade-equivalency). Based on this indicator, it was possible that Cole experienced more reading difficulty. Thus, it is reasonable to expect that Cole would require the intervention at a greater intensity level in order to demonstrate more significant increases in comprehension level and rate. Within the context of the study, each ART+PD session lasted for approximately 30 minutes and occurred only once a week. For this reason, it is easy to see that students with a reading disability may not demonstrate a sizable improvement in comprehension when they are provided with a less intensive intervention level.

For the majority of Charlotte's sessions, the ART intervention resulted in greater gains in comprehension level and rate when compared to the ART+PD. During the ART sessions, Charlotte read with greater focus and applied fix-it strategies more consistently. She appeared to have greater success in tracking her reading, thus resulting in less time to complete the passages. Moreover, she appeared to retain more of what she read since the absence of the discussion allowed for an immediate application of her knowledge when answering comprehension questions.

Charlotte's performance contrasted noticeably during the ART+PD sessions when compared to the ART sessions and may be explained in several ways. First, during the discussion, Charlotte appeared to be easily derailed to talking about unrelated topics. Often times, she required redirection to return to the passage topic and prompts from her discussion partners of what they had talked about. Of the three participants, Charlotte appeared to value the discussion for its social aspect and perceived it as an opportunity to socialize with her peers. For this reason, the time spend during text discussion was interspersed with on- and off-topic exchanges which distracted Charlotte from the passage content. Thus, she did not appear to

benefit comprehension-wise from the discussion component of ART+PD. Furthermore, because the ART+PD procedure was designed for the participant and her discussion partner to be in the same workspace, Charlotte was observed to be less attentive and more distracted by the presence of the other student than when she was receiving the ART intervention as a solo participant. During the read-aloud portion, Charlotte experienced more difficulty tracking her reading in the presence of another student. Often times, she would look up from her reading due to nearby sounds or stimuli. As a result, Charlotte would skip rows of text, lose her spot, or forget to apply fix-it strategies. Findings from the current study corroborates prior investigations showing variable responses and/or nonresponses to peer-mediated interventions (i.e., PALS) in children with inattention and high distractibility (Al Otaiba & Fuchs, 2006; Falk & Wehby, 2001). In fact, Al Otaiba and Fuchs (2006) found that the distractibility or behavioral problems of some students were so significant that much of the intervention time was lost to behavior management. Although peer-mediated instruction has been demonstrated to improve reading skills of students performing below grade level (McCullough et al., 2020), it is important that individual learning and social/emotional characteristics of students are considered when prescribing interventions.

Social Validity

An additional purpose of the study was to explore participating students' perspectives on their experiences with learning and applying ART/ART+PD as well as their perceptions of the intervention goals, implementation procedures, and outcomes. To fully evaluate students' perspectives on the social validity of the ART and ART+PD interventions in a meaningful way, multiple methods were utilized. Hence, the KIP, a self-report rating scale, and semi-structured interviews presented as complementary methods in providing a comprehensive basis for making claims about social validity.

Administered at different timepoints throughout intervention process, the KIP yielded quantitative results suggesting high acceptability of intervention procedures and high satisfaction with outcomes among participants. Additionally, semi-structured interviews were used to capture information deemed important by participants that might have been underrepresented or excluded from the prior quantitative method. As such, the qualitative inquiry resulted in a rich and complex body of testimonies showcasing students' perceptions of reading and learning barriers, perspectives on the importance of reading skills, the relationship of reading skills to future goals, attitude about reading, their application of the ART/ART+PD strategies, and the impact of ART/ART+PD on their reading and learning. Further, the total construct of social validity was assessed through inquiry of students' perceptions of intervention procedures, reading goals, and outcomes. Finally, student recommendations related to future implementation and suggestions for improvement of the ART/ART+PD were solicited.

In general, the qualitative findings substantiated quantitative KIP results in that students perceived the interventions to be enjoyable, effective in improving their understanding of text, beneficial when doing schoolwork, and relevant in helping them achieve reading goals. Furthermore, students shared insights about why reading was difficult for them, their classroom learning climate, and changes in their reading behaviors.

Integrating both sets of social validity data allowed for a deeper understanding of students' perspectives and experiences. Further, the mixed methodology supported the inclusion of student feedback throughout the intervention phase as well as garnered their suggestions for improving the utility of ART/ART+PD. By situating students as the primary respondents, the researcher hoped to underscore the importance of considering students' voices in determining the social validation of intervention effects.

Integrating Experimental and Social Validity Findings

As previously noted, data indicated high social validity with regard to ART/ART+PD interventions. Participating students generally reported liking the intervention procedures, valuing established intervention goals, and detecting improved comprehension of their texts. The positive changes in their comprehension performance as perceived by participants were supported by experimental findings. Based on visual analyses and effect size calculations, all three participants showed substantial increases in their comprehension rates and/or levels as a result of the ART/ART+PD. Although functional relations between independent and dependent variables were documented, integrating social validation data with experimental results further solidified claims of intervention effectiveness. In summary, interventions that were perceived favorably by consumers were more likely to increase their motivation and self-efficacy, a direct link to mastery and skill competency (Schunk, 1996; Stipek, 1996).

Limitations and Future Directions

Although the results of this study indicated that ART and ART+PD were effective in increasing reading comprehension level and/or rate, limitations to the interpretation of data should be considered. First, alternating treatment designs (ATD) are susceptible to carryover effects as a result of the rapid alterations. Participants who learned to use one academic strategy on a given day may carry over that knowledge to another academic strategy on the next day. Efforts were made to minimize this potential problem by randomizing the presentation of the ART and ART+PD conditions and ensuring that there was ample time between each session. Second, due to the nature of ATD, the effects of repeated and continuous use of each intervention could not be demonstrated due to the short duration between the two conditions. Third, participants with very specific profiles related to academic and cognitive functioning were

selected for this study, thus, these specifications may potentially limit generalizability. Finally, during the course of the study, participants were receiving ongoing specialized instruction in reading comprehension as part of their IDEA services, rendering it difficult to attribute the change in comprehension performance to exclusively the independent variables.

From an applied perspective, future studies may investigate whether changes in student reading comprehension can be replicated in classroom settings when applying ART/ART+PD. Another potential way to expand on this study may include examining performance differences in students with and without ADHD. As such, it may be beneficial to explore the suitability of ART+PD for students with significant levels of inattention and distractibility. Moreover, future research could investigate whether increased frequency of the interventions would improve students' comprehension skills to even greater extents than were demonstrated in the current study. Finally, since there are several evidence-based multiple-strategy tools for expository reading, future comparison studies may be considered to determine the most efficient and effective instructional strategies for enhancing students' comprehension.

Conclusion and Implications for Practice

In conclusion, the results of the current study suggest that using a multi-approach strategy such as ART, can lead to improvements in the reading comprehension performance of upper-elementary school students. Furthermore, adding a peer discussion component may enhance positive effects of ART for some students. Such positive effects were demonstrated within sessions and maintained to some extent over time.

All participants had a history of reading difficulties and were each found to have a specific learning disability in reading. As such, the current findings support using these interventions as a promising option when planning for students with specific learning needs. It is

also noteworthy to highlight students' testimonies of liking the intervention procedures, valuing the established goals of the interventions, and believing that their comprehension skills had improved as a result.

To this end, school psychologists consulting with educators and collaborating as members of problem-solving teams may consider recommending ART/ART+PD interventions to help improve expository reading comprehension. Moreover, educators implementing ART/ART+PD would be well served to incorporate curriculum materials used in daily classroom instruction. Potentially, this strategy may result in gains in general reading performance as well as in content area performance. During reading of text, students should be encouraged to actively monitor their comprehension by thinking about reading speed, connecting textual information to their own knowledge, and using explicit strategies to enhance their understanding. As shown in the current study, the additional peer discussion component led to greater gains for two of the participants. Thus, whether teachers choose to utilize the ART/ART+PD strategies as prescribed or to incorporate only certain aspects of the interventions in their classrooms, the practice of asking questions that instigate high-level thinking and facilitating discussion of text remains as an essential element in helping students achieve meaning from their reading. We hope that findings from this study will equip educators with additional tools to support upper-elementary students in accessing their content-area curriculum as text difficulty increases throughout the upper-elementary grades and beyond.

References

Afflerbach, P., Pearson, P., & Paris, S.G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher*, 61(5), 364-373.

Allington, R.L., & Johnston, P.H. (2002). *Reading to learn: Lessons from exemplary fourth-grade classrooms*. Guildford Press.

Alnahdi, G.H. (2015). Teaching reading for students with intellectual disabilities: A systematic review. *International Education Studies*, *8*, 79-87.

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities, 39*(5), 414-431.

Anderson, V. (1992). A teacher development project in transactional strategy instruction for teachers of severely reading disabled adolescents. *Teaching and Teacher Education*, *8*, 391-403.

Antoniou, F. & Souvignier, E. (2007). Strategy instruction in reading comprehension: An inter vention study for students with learning disabilities. *Learning Disabilities: A Contempo rary Journal*, *5*(1), 41-57.

Armbruster, B., Lehr, F., Osborn, J. (2001). Put reading first: The research building blocks for teaching children to read: Kindergarten through grade 3 (3rd ed.). Retrieved from https://www.nichd.nih.gov/sites/default/files/publications/pubs/Documents/PRFbook let.pdf

Arra, C. T., & Bahr, M. W. (2005). Teachers' and students' preferences for mathematics inter ventions: Implications for teacher acceptability in consultation. *Journal of Educational &*

Psychological Consultation, 16, 157–174

http://dx.doi.org/10.1207/s1532768xjepc1603_2

Baird, G.L., Scott, W. D., Dearing, E., & Hamill, S. K. (2009). Cognitive self-regulation in youth

with and without learning disabilities: Academic self-efficacy, theories of intelligence,

learning vs. performance goal preferences, and effort attributions. *Journal of Social and*

Clinical Psychology, *28*, 881-908.

Baker, S., Gertsen, R., & Scanlon, D. (2002). Procedural facilitators and cognitive strategies:

Tools for unraveling the mysteries of comprehension and the writing process, and for

providing meaningful access to the general curriculum. *Learning Disabilities Research*

and Practice, 17, 65-77.

Barlow, D. H., & Hayes, S. C. (1979). Alternating treatments design: One strategy for comparing

the effects of two treatments in a single subject. *Journal of Applied Behavior Analysis*,

12, 199-210.

Berkeley, S., & Larsen, A. (2018). Fostering self-regulation of students with learning disabilities:

Insights from 30 years of reading comprehension intervention research. *Learning Disabil*

ities Research & Practice, *33*(2), 75-86.

Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2010). Reading comprehension instruction

for students with learning disabilities, 1995-2006: A meta-analysis. *Remedial and*

Special Education, *31*(6), 423-436.

Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1995). A quasi-experimental validation

of transactional strategies instruction with low-achieving second grade readers. *Journal*

of Educational Psychology, *88*, 18-37.

Burns, M. K., Riley-Tillman, T.C., & Rathvon, N. (2017). *Effective school interventions* (3rd ed.). The Guildford Press.

Carter, S. L. (2010). Background of social validity. In S. L. Carter, *The social validity manual* (pp. 1-20). London, UK: Elsevier.

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind.* Harvard University Press.

Chan, L. K. S. (1991). Promoting strategy generalization through self-instructional training in students with reading disabilities. *Journal of Learning Disabilities*, *24*, 427-433.

Chanese, J. A. M. (2007). *The impact of design methodology on teachers' acceptability and usability ratings of repeated reading* [Unpublished doctoral dissertation]. University of Connecticut.

Cortiella, C., & Horowitz, S.H. (2014). *The state of learning disabilities: Facts, trends, and emerging issues*. National Center for Learning Disabilities.

Creswell, J. W., & Creswell, J. D. (2018). *Research design*. SAGE Publications Inc.

Crotty, M. (2015). *The foundations of social research*. SAGE Publications Inc.

Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P.A., …Carter, S. (2015). Text-processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text. *Reading Research Quarterly*, *50*, 393-416.

Durkin, D. (1978-1979). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 15*, 481-533.

Durkin, D. (1981). Reading comprehension instruction in five basal reading series. *Reading Research Quarterly*, *16*(4), 515-544.

Eckert, T. L., & Hintze, J.M. (2000). Behavioral conceptions and applications of acceptability: Issues related to service delivery and research methodology. *School Psychology Quarterly*, *15*, 123-148.

Eckert, T. L., Hier, B. O., Hamsho, N. F., & Malandrino, R. D. (2017). Assessing children's perceptions of academic interventions: The Kids Intervention Profile. *School Psychology Quarterly*, *32*(2), 268–281. https://doi.org/10.1037/spq0000200

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnaken berg, J. W. (2009). A synthesis of reading interventions and effects on reading compre hension outcomes for older struggling readers. *Review of Educational Research*, *79*(1), 262-300.

Englert, C. S., & Mariage, T. V. (1991). Making students partners in the comprehension process: Organizing the reading "POSSE." *Learning Disabilities Quarterly*, *14*, 123-138.

Falk, K. B., & Wehby, J. H. (2001). The effects of peer-assisted learning strategies on the beginning reading skills of young children with emotional or behavioral disorders. *Behavioral Disorders*, *26*(4), 344–359.

Finifter, D. H., Jensen, C. J., Wilson, C. E., & Koenig, B. L. (2005). A comprehensive, multitiered, targeted community needs assessment model: Methodology, dissemination, and implementation. *Family & Community Health*, *28*(4), 293–306. https://doi.org/10.1097/00003727-200510000-00003

Finn, C. A., & Sladeczek, I. E. (2001). Assessing the social validity of behavioral interventions: A review of treatment acceptability measures. *School Psychology Quarterly*, *16*(2), 176.

Firat, T. (2019). Effects of the TWA strategy instruction on reading comprehension of students with learning disabilities. *Educational Research Quarterly*, *43*(2), 24-54.

Fox, M., & Wilkinson, L. (1997). No longer travelers in a strange country. *Journal of Children's Literature*, *23*, 6-15.

Fuchs, D., Fuchs, L. S., & Vaughn, S. (2014). What is intensive instruction and why is it important? *Teaching Exceptional Children, 46* (4), 13-17.

Gajria, M., Jitendra, A. K., Sood, S., & Sacks, G. (2007). Improving comprehension of expository text in students with LD: A research synthesis. *Journal of Learning Disabilities*, *40*, 210-225.

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, *71*(2), 279-320.

Glover, T. & Albers, C. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*(2), 117-135.

Good, R. H., Baker, S. K., & Peyton, J. A. (2009). Making sense of nonsense word fluency: Determining adequate progress in early first-grade reading. *Reading & Writing Quar terly*, 25(1), 33-56.

Graham, S. & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 323-344). Guildford.

Greer, F. W., Wilson, B. S., & DiStefano, C., & Liu, J. (2012). Considering social validity in the context of emotional and behavioral screening. *NASP School Psychology Forum: Research in Practice*, *6*(4), 148-159.

Gresham, F. M., & Lopez, M.F. (1996). Social validation: A unifying construct for school-based consultation research and practice. *School Psychology Quarterly, 11*, 204-227.

Harris, K. R., Graham, S., Mason, L.H., & Friedlander, B. (2008). *Powerful writing strategies for all students*. Brookes.

Hays, D. G., & Singh, A. A. (2012). *Qualitative inquiry in clinical and educational settings*. Guilford Press.

Hill, C. E., & Knox, S. (2021). *Essentials of consensual qualitative research*. American Psychological Association.

Hirsch, Jr., E. D. (2003). The fourth-grade plunge. The cause. The cure. *American Educator*, *27*(1), 1-50.

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. https://doi.org/10.1177/001440290507100203

Idol-Maestas, L. (1985). Getting ready to read: Guided probing for poor comprehenders. *Learning Disability Quarterly*, *8*, 243-254.

Interventional Central. (n.d.). *How to improve reading comprehension with a cognitive strategy: Ask-Read-Tell.* https://www.interventioncentral.org/node/966325

Jacob, S., & Hartshorne, T. S. (2007). *Ethics and law for school psychologists, 5th ed* (pp. xv, 413). John Wiley & Sons Inc.

Johnson, R. B. & Christensen, L. (2017*). Educational research: Quantitative, qualitative, and mixed approaches*. Sage.

Lane, H. (2014). *Evidence-based reading instruction for grades K-5* (Document No. IC-12). Retrieved from http://ceedar.education.efl.edu/tools/innovation-configurations/

Ledford, J. R., & Gast, D. L. (Eds.). (2018). *Single case research methodology: Applications in special education and behavioral sciences* (3rd ed.). Routledge.

Lee, Y. L., & Jonson-Reid, M. (2016). The role of self-efficacy in reading achievement of young children in urban schools. *Child and Adolescent Social Work Journal*, *33*, 79-89.

Leko, M. M. (2014). The Value of Qualitative Methods in Social Validity Research. *Remedial and Special Education*, *35*(5), 275–286. https://doi.org/10.1177/0741932514524002

Lyon, G. R. (2001). *Measuring success: Using assessments and accountability to raise student achievement*. Statement to the Committee on Education and the Workforce, U.S. House of Representatives, Washington, DC. Retrieved from https://ia800201.us.ar chive.org/28/items/ERIC_ED465820/ERIC_ED465820.pdf

Magiera, K., & Zigmond, N. (2005). Co-teaching in middle school classrooms under routine conditions: Does the instructional experience differ for students with disabilities in co-taught and solo-taught classes? *Learning Disabilities Research & Practice*, *20*, 79-85.

Mahdavi, J. N., & Tensfeldt, L. (2013). Untangling reading comprehension strategy instruction: Assisting struggling readers in the primary grades. *Preventing School Failure*, *57*(2), 77-93.

Mason, L. H. (2004). Explicit self-regulated strategy development versus reciprocal questioning: Effect on expository reading comprehension among struggling readers. *Journal of Educa tional Psychology*, *96*, 283-296.

Mason, L. H., Meadan, H., Hedin, L., & Corso, L. (2006). Self-regulated strategy development for expository text comprehension. *Teaching Exceptional Children*, March/April, 47-52.

Mautone, J. A., DuPaul, G. J., Jitendra, A. K., Tresco, K. E., Junod, R. V., & Volpe, R. J. (2009). The relationship between treatment integrity and acceptability of reading interventions for children with attention-deficit/hyperactivity disorder. *Psychology in the Schools, 46,* 919–931. http://dx.doi.org/10.1002/pits.20434

McCallum, R. S., Krohn, K. R., Skinner, C. H., Hilton-Prillhart, A., Hopkins, M. Waller, S., & Polite, F. (2011). Improving reading comprehension of at-risk high-school students: The art of reading program. *Psychology in the Schools*, *48*(1), 78-86.

McCullough, S. N., Wright, J., & Sutherland, K. S. (2020). Peer-mediated instruction. In R. T. Boon, M. D. Burke, & L. Bowman-Perrott (Eds.), *Literacy instruction for students with emotional and behavioral disorders* (pp. 93-109). Information Age Publishing, Inc.

McGee. A., & Johnson, H. (2003). The effect of inference training on skilled and less skilled comprehenders. *Educational Psychology*, *23*(1), 49-59.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, *16*, 72-103.

Miltenberger, R. G. (1990). Assessment of Treatment Acceptability: A Review of the Literature. *Topics in Early Childhood Special Education*, *10*(3), 24–38. https://doi.org/10.1177/027112149001000304

Mulcahy, E., Bernardes, E., & Baars, S. (2001). The relationship between reading age, education, and life outcomes. *LKMco,* 1-13.

Nastasi, B. K., & Truscott, S. D. (2000). Acceptability research in school psychology: Current trends and future directions. *School Psychology Quarterly*, *15*(2), 117-122.

National Assessment of Educational Progress. (2019). *See how the U.S. students performed in reading at grades 4 and 8.* Retrieved from https://www.nationsreportcard.gov/high lights/reading/2019/

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.*

Retrieved from https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Docu
ments/report.pdf

O'Connor, R. E., Beach, K. D., Sanchez, V., Bocian, K.M., Roberts, S., & Chan, O. (2017). Building better bridges: Teaching adolescents who are poor readers in eighth grade to comprehend history text. *Learning Disability Quarterly*, *40*, 174-186.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognitive and Instruction*, *1*, 117-175.

Paris, S. G., & Myers, M. (1981). Comprehension monitoring, memory, and study strategies of good and poor readers. *Journal of Reading Behavior*, *13*(1), 5-22.

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, *40*, 357-367.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). SAGE Publications Ltd.

Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education: Theory, research, and applica tions*. Prentiss Hall Merrill.

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. Kamil, P. Mosenthal, P. Pearson, & R. Barr (Eds.), *Handbook of reading research* (3rd ed., pp. 545-562). Lawrence Erlbaum.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Erlbaum.

Pressley, M., Harris, K. R., & Marks, M. B. (1992). But good strategy instructors are constructivists! *Educational Psychology Review*, *4*, 3-31.

Pressley, M., & Wharton-McDonald, R. (1997). Skilled comprehension and its development through instruction. *School Psychology Review*, *26*(3), 448-467.

Ridge, A. D., & Skinner, C. H. (2011). Using the TELLS prereading procedure to enhance com prehension levels and rates in secondary students. *Psychology in the Schools*, *48*(1), 46-58.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95,* 240 –257. http://dx.doi.org/10.1037/0022-0663.95.2.240

Rowe, S. S. (2012). Reading intervention and assessment acceptability: A literature review. *School Psychology Forum*, *6*(3), 77-88.

Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33,* 359–382.

Scruggs, T. E., Mastropieri, M.A., Berkeley, S., & Graetz, J. E. (2010). Do special education interventions improve learning of secondary content? A meta-analysis. *Remedial & Special Education*, *31*, 437-449.

Sencibaugh, J. M. (2007). Meta-analysis of reading comprehension interventions for students with learning disabilities: Strategies and implications. *Reading Improvement*, *44*(1), 6-22.

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C. & Tor geson, J., (2010). Improving reading comprehension in kindergarten through 3rd grade: IES practice guide. Retrieved from https://files.eric.ed.gov/fulltext/ED512029.pdf

Shapiro, E. S. (2004). *Academic skills problems: Direct assessment and intervention* (3rd ed.). Guildford.

Shelton, A., Wexler, J., Silverman, R. D., & Stapleton, L. M. (2019). A synthesis of reading comprehension interventions for persons with mild intellectual disability. *Review of Edu cational Research*, *89*(4), 612-651.

Sideridis, G. D., Mouzaki, A., Simos, P., & Protopapas, A. (2006). Classification of students with reading comprehension difficulties: The roles of motivation, affect, and psycho pathology. *Learning Disability Quarterly*, 29, 159-180.

Skinner, C. H. (1998). Preventing academic skills deficits. In T. S. Watson & F. Gresham (Eds.), *Handbook of child behavior therapy* (pp. 61-83). Plenum.

Society for Neuroscience. (2020). *Dyslexia: What brain research reveals about reading*. Retrieved from http://www.ldonline.org/article/10784/#:~:text=About%203.5%20per cent%20of%20American, %2C%20writing%2C%20and%20spelling%20difficulties.

Spargo, E. (1989). *Timed readings* (3rd ed.). Jamestown Publishers.

Stanovich, K. E. (1986). Matthew Effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360-407.

Stevens, J. P. (2007). *Intermediate statistics: A modern approach* (3rd ed.). Taylor & Francis Group.

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

Swanson, H. L., & Siegel, L. S. (2001). Learning disabilities as a working memory deficit. *Issues in Education: Contributions from Educational Psychology, 7,* 1–48.

Taylor, L. K., Alber, S. R., & Walker, D. W. (2002). The comparative effects of a modified self-questioning strategy and story mapping on the reading comprehension of elementary stu dents with disabilities. *Journal of Behavioral Education*, *11*(2), 69-87.

Vaughn, S., Gersten, R., & Chard, D. J. (2000). The underlying message in LD intervention

    research: Findings from research syntheses. *The Council for Exceptional Children*, *67*(1),

    99-114.

Waas, G. A., & Anderson, G. P. (1991). Outcome expectancy and treatment acceptability:

    Perceptions of school-based interventions. *Professional Psychology: Research and*

    Practice, *22*(2), 149.

Walker, B.J. (2003). The cultivation of student self-efficacy in reading and writing. *Reading &*

    Writing Quarterly, *19*, 173-187.

What Works Clearinghouse. (2017). *Self-regulated strategy development*. Retrieved from

    https://ies.ed.gov/ncee/wwc/InterventionReport/680

Williams, J. P. (1993). Comprehension of students with and without learning disabilities: Identif

    ication of narrative themes and idiosyncratic text representations. *Journal of Educational*

    Psychology, *85*, 631-641.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied

    behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, *11*(2), 203–

    214. https://doi.org/10.1901/jaba.1978.11-203

Wong, B. Y. L., & Wilson, M. (1984). Investigating awareness of and teaching passage organi

    zation in learning disabled children. *Journal of Learning Disabilities*, *17*, 477-482.

Zimmerman, B. J. (2008). Investigation self-regulation and motivation: Historical background,

    methodological developments, and future prospects. *American Educational Research*

    Journal, *45*, 166-183.

Zimmermann, L. M., & Reed, D. K. (2020). Improving reading comprehension of informational

    text: Text structure instruction for students with or at risk for learning disabilities.

    Teaching Exceptional Children, 52(4), 232-241.

Table 2.1 *Mean, SD, and Effect Size for Each Participant Across Conditions*

| Participant	Dependent Variable	Baseline Mean (*SD*)	ART Mean (*SD*)	ART+PD Mean (*SD*)	BL-ART NAP	BL-ART+PD NAP
Bryson	Comprehension Level (%C)	54.00 (15.17)	74.00 (8.94)	81.67 (11.69)	0.92	0.97
	Comprehension Rate (%C/M)	8.86 (3.37)	10.65 (4.15)	12.33 (2.85)	0.68	0.80
Charlotte	Comprehension Level (%C)	74.00 (13.42)	83.33 (17.51)	76.00 (5.48)	0.72	0.52
	Comprehension Rate (%C/M)	16.20 (3.91)	19.80 (5.36)	15.00 (4.75)	0.70	0.52
Cole	Comprehension Level (%C)	78.00 (13.04)	82.00 (13.04)	93.30 (5.16)	0.58	0.87
	Comprehension Rate (%C/M)	13.03 (2.17)	12.51 (2.44)	14.59 (2.53)	0.40	0.67

Note. SD = Standard deviation; BL = Baseline; NAP = Nonoverlap of all pairs. Effect size

interpretations are based on recommendations from Parker & Vannest (2009) with tentative NAP

ranges: Weak effects (0-0.65), medium effects (0.66-0.92), large effects (0.93-1.00).

**Table 2.2** *Mean and Standard Deviation for KIP Items Across Administrations*

Item	Administration Time		
	After Training Phase	After 2 ART & 2 ART+PD Sessions	After Intervention Phase
	M_1 (SD$_1$)	M_2 (SD$_2$)	M_3 (SD$_3$)
How much do you like learning to understand what you read each week?	4.33 (0.52)	4.67 (0.52)	4.67 (0.82)
How much do you like being told how to read so you can better understand what you read?	4.17 (0.75)	4.17 (1.17)	4.50 (0.84)
Were there times when you didn't want to learn these skills with us?	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
Were there times when you wished you could work more with us on understanding what you read better?	3.00 (1.41)	3.50 (1.38)	4.00 (1.41)
How much do you like being told how long it took you to read and how many questions you answered correctly?	4.00 (1.15)	4.83 (0.41)	4.83 (0.41)
How much do you think it helps you to know how long it took you to read and how many answers you got correct?	4.67 (0.52)	4.67 (0.52)	4.50 (0.55)
Do you think your understanding of what you read has improved?	4.67 (0.82)	3.83 (1.17)	4.50 (0.55)
Do you think your understanding of what you read has gotten worse?	1.17 (0.41)	1.00 (0.00)	1.17 (0.41)

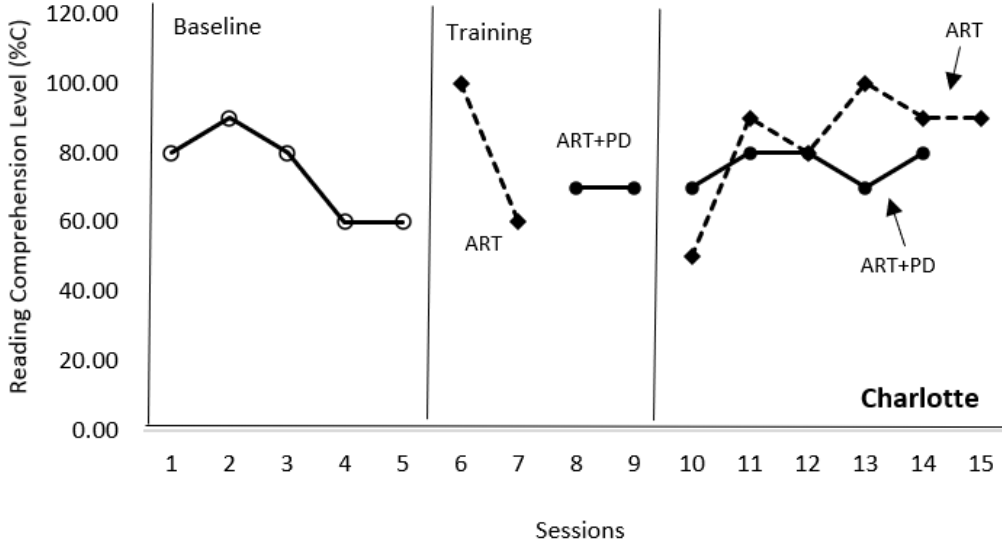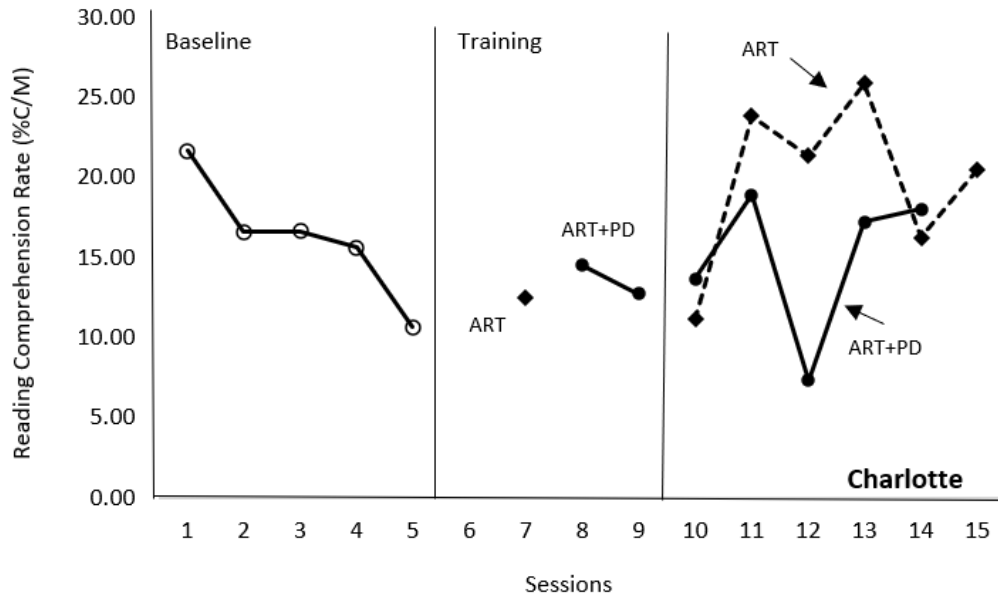*Note*. 1 = Not at all; 2 = A little bit; 3 = Some; 4 = A lot; 5 = Very, very much
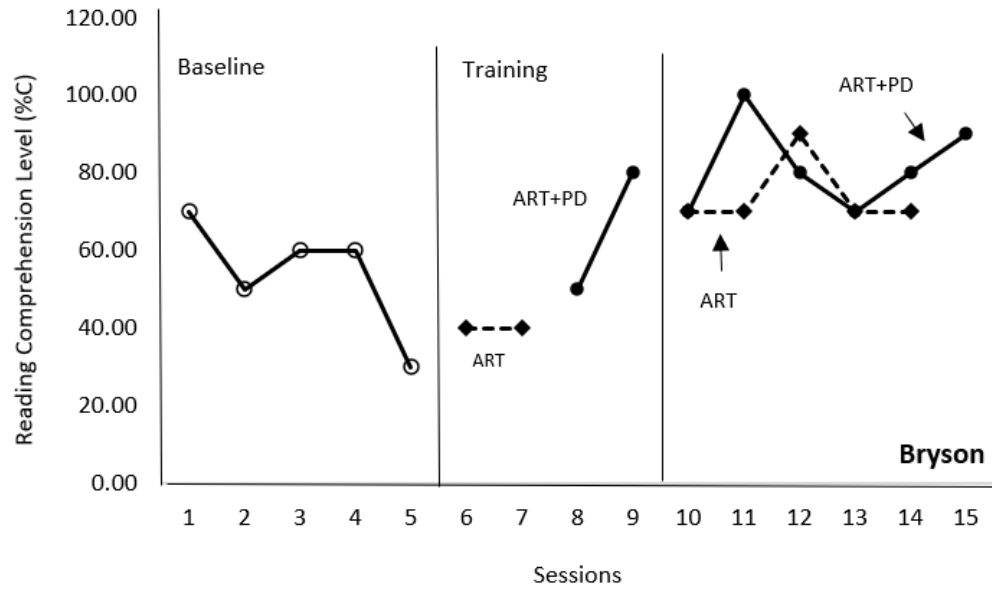
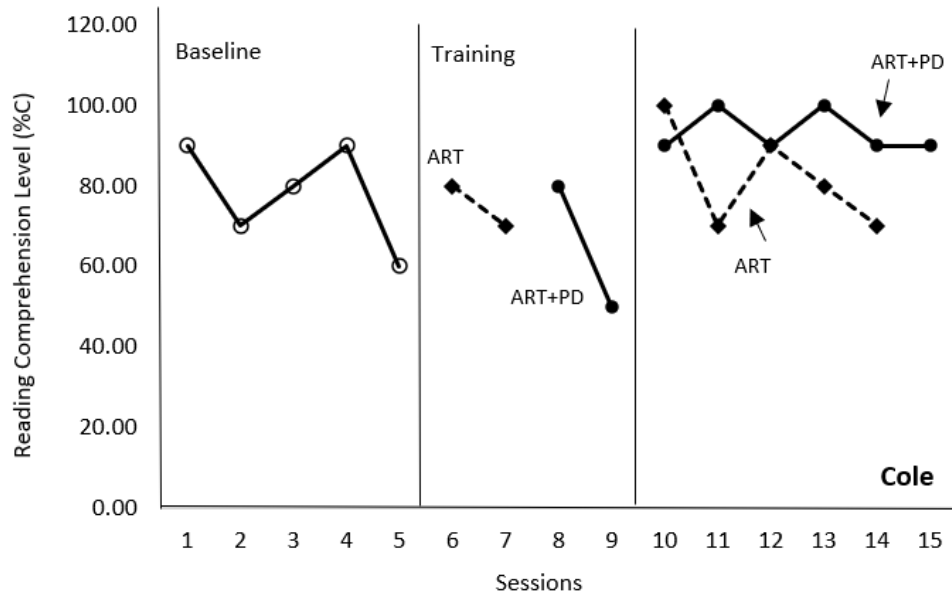**Figure 2.1** *Comprehension Level for Bryson*
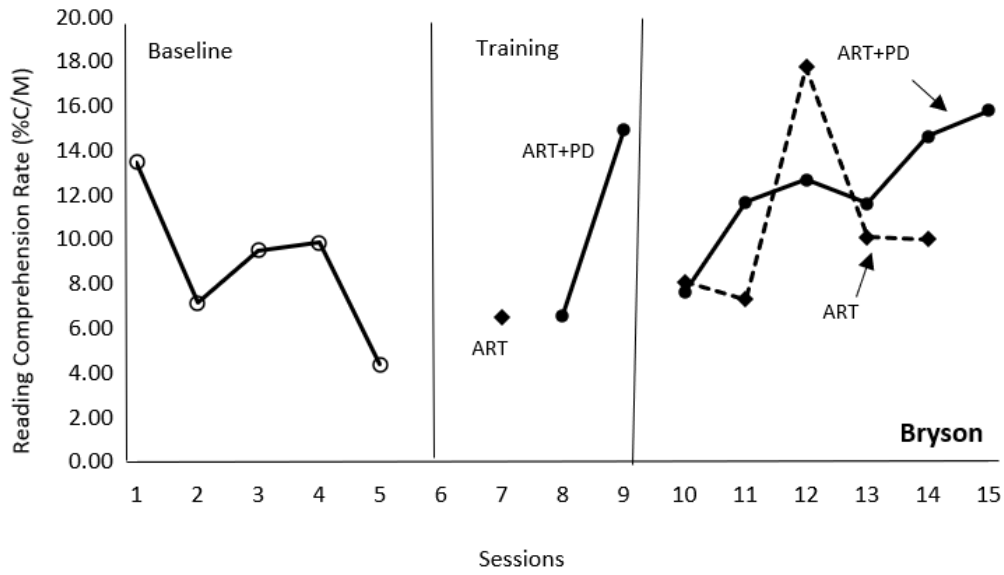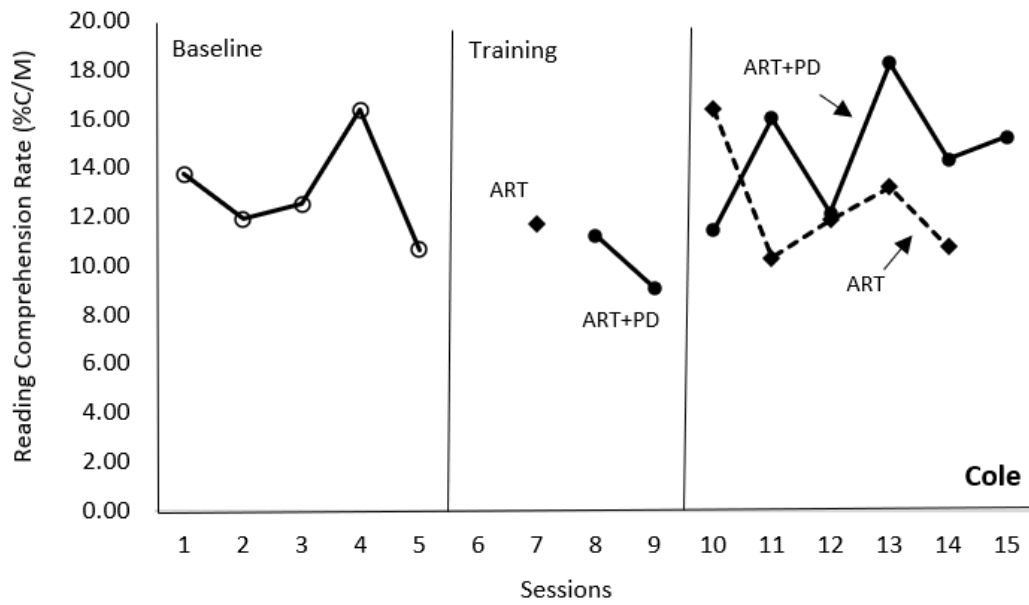
Figure 2.2 *Comprehension Rate for Bryson*

Figure 2.3 *Comprehension Level for Charlotte*

Figure 2.4 *Comprehension Rate for Charlotte*

Figure 2.5 *Comprehension Level for Cole*

Figure 2.6 *Comprehension Rate for Cole*

APPENDICES

Appendix A

The Kids Intervention Profile (KIP; Eckert et al., 2017) – Adapted Version

Item	Not at all	A little bit	Some	A lot	Very, very much
1. How much do you like learning to understand what you read each week?	1	2	3	4	5
2. How much do you like being told how to read so you can better understand what you read?	1	2	3	4	5
3. Were there times when you didn't want to learn these skills with us?	1	2	3	4	5
4. Were there any times when you wished you could work more with us on understanding what you read better?	1	2	3	4	5
5. How much do you like being told how long it took you to read and how many questions you answered correctly?	1	2	3	4	5
6. How much do you think it helps you to know how long it took you to read and how many answers you got correct?	1	2	3	4	5
7. Do you think your understanding of what you read has improved?	1	2	3	4	5
8. Do you think your understanding of what you read has gotten worse?	1	2	3	4	5

Boxes of increasing sizes are included to help students have a more concrete understanding of each rating (Eckert et al., 2017). Below is a sample response format with a visual illustration for each Likert rating.
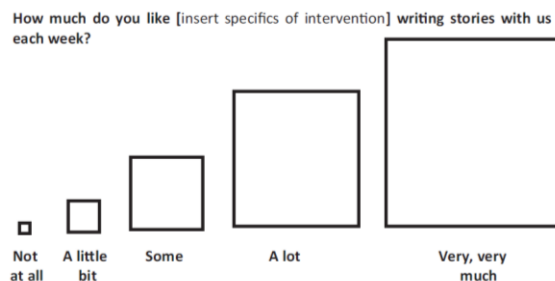
Appendix B

KIP Script

Step 1: Provide participant with the KIP

Step 2: Read instructions without deviating. Say, *"This is a short survey about the ART/ART+PD intervention. I want to know what you think of it. There are no right or wrong answers. I will read each statement to you. Circle the number that best represents your answer for each statement. "Not at all" is #1, "A little bit" is #2, "Some" is #3, "A lot" is #4, and "Very, very much" is #5.*

Step 3: Explain the visual representation for each rating. Say to participant, *"Look at the different squares at the bottom. This is another way to help you choose the rating for each statement. Each rating is shown with a different size square. "Not at all" is the smallest square on your left and "Very, very much" is the largest square to your right. If you forget what each rating means, you can use the squares to help you. Do you have any questions?"*

Step 4: As soon as participant is ready, read each statement and each rating. Repeat if necessary.

Step 5: Say to participant, *"Now circle your answer. Remember, there are no right or wrong answers."*

Step 6: Once all statements are read, check to make sure all items are answered.

Step 7: Collect KIP form.

Appendix C

Baseline (Control) Condition Protocol

Step 1: Provide participant with reading passage at his/her instructional level

Step 2: Read instructions without deviating. Say, "*Take your time and read the passage aloud when you are told to begin. When you finish reading, you should give the reading passage back to me. Next, you will be given ten multiple-choice questions about what you have read. I will not be able to help you. Do the best you can.*"

Step 3: Have timer ready and say to participant, "*Remember to read your passage aloud. You may begin.*"

Step 4: As soon as participant begins reading, start the timer.

Step 5: If participant pauses for more than 3 seconds, encourage participants to keep reading.

Step 6: As soon as participant finishes reading, stop the timer.

Step 7: Record time in seconds.

Step 8: Collect participant's reading passage.

Step 9: Hand out sheet with comprehension questions.

Step 10: Say to participant, "*Here are ten multiple-choice questions about what you have read. For each question, you have three choices. Circle the best choice for your answer. Be sure to do all ten questions. I will not be able to help you. Do the best you can.*"

Step 11: Check to make sure participant has answered all ten questions.

Step 12: Collect sheet with comprehension questions.

Appendix D

ART Training

DAY 1

Step 1: Develop background knowledge through initial conferences with participants and pre-
skill development,

Step 2: Discuss the ART strategy, teach mnemonic device, identify areas/set goals for
improvement, and provide feedback

Step 3: Model the ART strategy, model for students on how to fill out the Student Worksheet
when reading passages, model "fix-it" strategies such as re-read paragraph, slow reading,
underline unfamiliar words, and use context clues

Step 4: Help participants to learn and use the ART strategy, provide prompts if necessary,
provide ART Student Worksheet to help guide reading, and offer feedback

DAY 2

Step 5: Support the ART strategy, student-teacher collaborative practice, and provide feed-
back   *Step 6*:
Support the ART strategy, gradually release responsibility to students, and provide
feedback

Step 7: Support the ART strategy, gradually release more responsibility to students, encourage
independent student practice, and provide feedback

Step 8: Continue with independent student practice and offer positive reinforcement

Appendix E

ART+PD Training

<u>DAY 3</u>

Step 1: Discuss the ART+PD strategy, teach mnemonic device, review goals for improvement, and provide feedback

Step 2: Review with students on how to fill out the Student Worksheet when reading passages, review "fix-it" strategies such as re-read paragraph, slow reading, underline unfamiliar words, and use context clues

Step 3: Discuss the peer discussion (PD) component of ART+PD, model the PD for students by sharing what was learned from practice passage, questions/answers, and any interesting information from reading

Step 4: Help participants to learn and use the ART+PD strategy, provide prompts if necessary, provide ART+PD Student Worksheet to help guide reading and PD, and offer feedback

<u>DAY 4</u>

Step 5: Support the ART +PD strategy, student-teacher collaborative practice, and provide constructive feedback

Step 6: Support the ART+PD strategy, gradually release responsibility to students, implement student-peer collaborative practice, and provide feedback

Step 7: Support the ART+PD strategy, gradually release more responsibility to students, encourage independent student practice, and provide feedback

Step 8: Continue with independent student practice and offer positive reinforcement

Appendix F

ART Student Worksheet

ASK-READ-TELL (ART): Student Worksheet (McCallum et al., 2010)

Name: _____ Passage/Page Numbers: _____ Date:_____

Directions: Use the checklist below to guide your reading of this passage. Check off each step when completed.

Step 1: Goal Before Reading: I look at the title of the passage and ASK myself these questions:

☐ What is the **main topic** of the passage? What does it discuss?

☐ What **information** do I **already know** about this topic?

Based on the title, what are **two** questions about this passage's topic that I would like to have answered in my reading?:

1. _____ ?

2. _____ ?

Step 2: Goal While Reading: I READ the passage carefully for full understanding:

☐ While reading, I stop after each paragraph to ask, "Did I **understand** what I just read?"

☐ If I **do** understand the paragraph, I mark it with a plus sign (+) and continue reading.
   If I **do not** understand the paragraph, I mark it with a minus (-) sign and:
   - reread the paragraph;
   - slow my reading;
   - focus my *full* attention on what I am reading;
   - underline any words that I do not know and try to figure them out from the reading (context).

Step 3: Goal After Reading: I TELL what I learned from the passage:

☐ Based on my reading, here are answers to my **two** questions from Step 1:

1. _____

   _____

2. _____

   _____

Appendix G

ART Condition Protocol

Step 1: Provide participant with reading passage at his/her instructional level and ART Student Worksheet

Step 2: Read instructions without deviating. Say, "*Our goal is to help you use skills to help you better understand what you read. You will be using the 3-step strategy called ART. Look at the ART Student Worksheet in front of you. Step 1 is "A" which stands for "Ask". Look at the title of the passage and ask yourself some questions about the passage. What do you think is the main topic of the passage and what information do you already know? Based on the title of the passage, think of two questions about the passage topic that you would like to have answered in your reading. Now write the two questions in the blanks.*"

Step 3: Check to make sure participant has written down two questions on the Student Worksheet.

Step 4: Read instructions without deviating. Say, "*Step 2 is "R" which stands for "Read for full understanding." After you read each paragraph, ask yourself if you understand what you just read. If you understand the paragraph, mark it with a plus sign (+). If you do not understand the paragraph, mark it with a minus sign (-). Be sure to use your "fix-it" strategies like reread the paragraph, slow your reading, focus your full attention on what you read, underline unfamiliar words, and use clues from the reading to help you figure it out. Do you have any questions?*"

Step 5: Read instructions without deviating. Say, "*Take your time and read the passage aloud when you are told to begin. Remember to use the skills I just went over to help you better understand your passage.*"

Step 3: Have timer ready and say to participant, "*Remember to read your passage aloud. You may begin.*"

Step 4: As soon as participant begins reading, start the timer.

Step 5: If participant pauses for more than 3 seconds, encourage participants to keep reading. Check to make sure participant is marking each paragraph during reading. Monitor for use of "fix-it" strategies.

Step 6: As soon as participant finishes reading, stop the timer.

Step 7: Record time in seconds.

Step 8: Read instructions without deviating. Say, "*Step 3 is "T" which stands for "Tell yourself what you learned from the passage." Based on your reading, write the answers to your two questions from Step 1 in the blanks.*

Step 9: Check to make sure participant has written down two answers on Student Worksheet.

Step 10: Collect participant's reading passage.

Step 11: Hand out sheet with comprehension questions.

Step 12: Read instructions without deviating. Say, "*Here are ten multiple-choice questions about what you have read. For each question, you have three choices. Circle the best choice for your answer. Be sure to do all ten questions. I will not be able to help you. Do the best you can.*"

Step 13: Check to make sure participant has answered all ten questions.

Step 14: Collect sheet with comprehension questions and Student Worksheet.

Appendix H

ART+PD Student Worksheet

ASK-READ-TELL (ART): Student Worksheet (McCallum et al., 2010)

Name: _____ Passage/Page Numbers: _____ Date:_____

Directions: Use the checklist below to guide your reading of this passage. Check off each step when completed.

Step 1: Goal Before Reading: I look at the title of the passage and ASK myself these questions:

☐ What is the **main topic** of the passage? What does it discuss?

☐ What **information** do I **already know** about this topic?

Based on the title, what are **two** questions about this passage's topic that I would like to have answered in my reading?:

1. _____ ?

2. _____ ?

Step 2: Goal While Reading: I READ the passage carefully for full understanding:

☐ While reading, I stop after each paragraph to ask, "Did I **understand** what I just read?"

☐ If I **do** understand the paragraph, I mark it with a plus sign (+) and continue reading.
   If I **do not** understand the paragraph, I mark it with a minus (-) sign and:
   - reread the paragraph;
   - slow my reading;
   - focus my *full* attention on what I am reading;
   - underline any words that I do not know and try to figure them out from the reading (context).

Step 3: Goal After Reading: I TELL what I learned from the passage:

☐ Based on my reading, here are answers to my **two** questions from Step 1:

1. _____

   _____

2. _____

   _____

☐ When I meet with my peer partner, we **TELL** each other **what we learned** from the passage, sharing our questions and answers. Then we talk about any other interesting information from the reading.

Appendix I

ART+PD Condition Protocol

Step 1: Provide participant and a randomly selected discussion partner with the reading passage at participant's instructional level and ART+PD Student Worksheet.

Step 2: Read instructions without deviating. Say, "*Our goal is to help you use skills to help you better understand what you read. You will be using the 3-step strategy called ART. Look at the ART Student Worksheet in front of you. Step 1 is "A" which stands for "Ask". Look at the title of the passage and ask yourself some questions about the passage. What do you think is the main topic of the passage and what information do you already know? Based on the title of the passage, think of two questions about the passage topic that you would like to have answered in your reading. Now write the two questions in the blanks.*"

Step 3: Check to make sure participant has written down two questions on the Student Worksheet.

Step 4: Read instructions without deviating. Say, "*Step 2 is "R" which stands for "Read for full understanding." After you read each paragraph, ask yourself if you understand what you just read. If you understand the paragraph, mark it with a plus sign (+). If you do not understand the paragraph, mark it with a minus sign (-). Be sure to use your "fix-it" strategies like reread the paragraph, slow your reading, focus your full attention on what you read, underline unfamiliar words, and use clues from the reading to help you figure it out. Do you have any questions?*"

Step 5: Read instructions without deviating. Say, "*Take your time and read the passage aloud when you are told to begin. Remember to use the skills I just went over to help you better understand your passage.*"

Step 3: Have timer ready and say to participant, "*Remember to read your passage aloud. You may begin.*"

Step 4: As soon as participant begins reading, start the timer.

Step 5: If participant pauses for more than 3 seconds, encourage participants to keep reading. Check to make sure participant is marking each paragraph during reading. Monitor for use of "fix-it" strategies.

Step 6: As soon as participant finishes reading, stop the timer.

Step 7: Record time in seconds.

Step 8: Read instructions without deviating. Say, "*Step 3 is "T" which stands for "Tell yourself what you learned from the passage." Based on your reading, write the answers to your two questions from Step 1 in the blanks.*

Step 9: Check to make sure participant has written down two answers on Student Worksheet.

Step 10: Pair participant with discussion partner.

Step 11: Read instructions without deviating. Say, "*Now, you will have a discussion partner. You will tell your partner what you learned from the passage. Tell your partner your two questions and answers. Then tell your partner anything you found interesting from the reading. You will go first. Do you have any questions?*"

Step 12: Repeat instructions if necessary. Say, "*You may begin your discussion with your partner.*"

Step 13: Monitor to make sure participant is actively participating in peer discussion. Allow time for discussion partner to participate in peer discussion. Facilitate when needed.

Step 14: Collect reading passages from participant and discussion partner.

Step 15: Hand out sheet with comprehension questions.

Step 16: Read instructions without deviating. Say, "*Here are ten multiple-choice questions about what you have read. For each question, you have three choices. Circle the best choice for your answer. Be sure to do all ten questions. I will not be able to help you. Do the best you can.*"

Step 17: Check to make sure participant has answered all ten questions.

Step 18: Collect sheet with comprehension questions and Student Worksheet.

Appendix J

Procedural Fidelity

Observer: _____   Implementer: _____   Study Participant: _____

Date: _____   Start / End Time: _____/_____   Session: _____

Condition/Phase (Circle one that applies):   Baseline   ART   ART+PD

Directions: While observing implementer, please record whether implementer emitted behavior during instructional procedure for each session

Key: (+) = occurrence; (-) = nonoccurrence

Implementer Behaviors	Mark + or -
Baseline Phase	
1. Gather materials: reading passage, comprehension question sheet, video recorder & timer	
2. Begin video recording	
3. Hand out reading passage to participant	
4. Read Baseline (Control) Script without deviating	
5. Time participant's reading	
6. Stop the timer after participant finishes reading	
7. Record participant's time	
8. Collect reading passage	
9. Hand out comprehension question sheet	
10. Read instructions for comprehension question sheet without deviating	
11. Check to see that all questions were completed	
12. Collect comprehension question sheet	

13. Stop video recording	
ART Condition	
1. Gather materials: reading passage, ART Student Worksheet, comprehension question worksheet, video recorder, & timer	
2. Begin video recording	
3. Hand out reading passage & ART Student Worksheet to participant	
4. Read Step 1 instructions without deviating	
5. Check to make sure participant writes 2 questions	
6. Read Step 2 instructions without deviating	
7. Time participant's reading	
8. Stop the timer after participant finishes reading	
9. Record participant's time	
10. Read Step 3 instructions without deviating	
11. Check to make sure participant writes 2 answers for the 2 questions	
12. Collect reading passage	
13. Hand out comprehension question sheet	
14. Read instructions for comprehension question sheet without deviating	
15. Check to see that all questions were completed	
16. Collect comprehension question sheet.	
17. Stop video recording	
ART+PD Condition	
1. Gather materials: reading passage, ART Student Worksheet, comprehension question worksheet, video recorder, & timer	
2. Begin video recording	
3. Hand out reading passage & ART+PD Student Worksheet to participant & discussion partner	

4. Read Step 1 instructions without deviating	
5. Check to make sure participant writes 2 questions	
6. Read Step 2 instructions without deviating	
7. Time participant's reading	
8. Stop the timer after participant finishes reading	
9. Record participant's time	
10. Read Step 3 instructions without deviating	
11. Check to make sure participant writes 2 answers for the 2 questions	
12. Pair participant with discussion partner	
13. Read instruction for peer discussion without deviating	
14. Monitor participant's active participation in peer discussion & facilitate turn-taking/ex-changes between participant and discussion partner	
15. Collect reading passage	
16. Hand out comprehension question sheet	
17. Read instructions for comprehension question sheet without deviating	
18. Check to see that all questions were completed	
19. Collect comprehension question sheet.	
20. Stop video recording	

Notes:

Fidelity Percentage: _____

Appendix K

Interview Protocol
Questions adapted from Gresham & Lopez (1996) semi-structured interview for social validation

<u>Student Script</u>: *Thank you for your participation in this study. After completing the ART and ART+PD interventions, the next step in the study is to conduct a student interview. It is important for us to learn from your experiences so that the ART and ART+PD can be beneficial and successful for other students. This interview will ask questions related to your thoughts, perceptions, and experiences with the ART and ART+PD interventions. Your responses will not be linked to you or your school. This interview will take approximately 30 minutes.*

Opening question: Tell me the name of/tell me about your favorite book/magazine/etc..

Social Significance of Goals

1. Why is reading hard for you? (Participant) Why do you think reading is hard for some students? (Discussion Partner)
2. Describe how this reading problem affects your schoolwork?
3. Do you feel that it is important to learn skills to help you better understand your reading? Why or why not?
4. If you learn skills to help you better understand what you read, how will this affect your schoolwork?

A. Social Acceptability of Procedures
   <u>Script</u>: *Now I'm going to ask you questions about the ART/ART+PD interventions.*

5. How do you feel about Step 1 – in which you look at the title and ask yourself questions about the reading passage?
6. How do you feel about Step 2 – in which you ask yourself if you understand each paragraph that you read by marking it with a plus or minus sign?
7. How do you feel about using "fix-it" strategies while you read?
8. How do you feel about Step 3 – on which you answer the two questions from Step 1 based on your reading?
9. How do you feel about the discussion part in which you and a partner talk about what you read?
10. What part of the intervention do you like the most? Why?
    What part of the intervention do you like the least? Why?
11. If you could change any part of the ART+PD intervention to make it better/easier, what would you change? Why?

B. Social Importance of Outcomes

12. In what ways do you think the ART/ART+PD interventions have helped you when you read?

In what ways do you think the ART/ART+PD interventions have <u>not</u> helped you when you read?

13. After learning the skills from the ART/ART+PD interventions, describe how these affect your schoolwork?
14. Are you satisfied (happy) with the effects of the ART/ART+PD interventions? Why or why not?
15. On a scale of 1 to 10, with 1 being not satisfied (happy) at all and 10 being super satisfied (happy), how would you rate the ART/ART+PD?
16. Would you recommend the ART+PD intervention for other students? Why or why not?

Closing questions: What is one last thing you would like for me to know about students like you? What would you like for me to know about helping students like reading?