Effects of Wh-Question Graphic Organizers on Reading Comprehension Skills of Students with Autism Spectrum Disorders

Keri S. Bethune  
James Madison University

Charles L. Wood  
University of North Carolina at Charlotte

Abstract: Students with autism spectrum disorders often have difficulty with reading comprehension. This study used a delayed multiple baseline across participants design to evaluate the effects of graphic organizers on the accuracy of wh-questions answered following short passage reading. Participants were three elementary-age students with autism spectrum disorder. Results indicated improved accuracy of responses to wh-questions, generalization, and maintenance of gains following intervention. Implications for future research and practice are discussed.

Many students with autism spectrum disorders (ASD) acquire reading decoding skills, but continue to struggle with reading comprehension (Nation, Clarke, Wright, & Williams, 2006). Comprehension occurs when the reader actively obtains meaning from written text (Bursuck & Damer, 2011). Newman et al. (2007) found that students with ASD had statistically significant lower scores on reading comprehension compared to scores by their typically developing peers, even when controlling for sight word recognition. Saalasti and colleagues (2008) showed that students with Asperger syndrome had significantly lower scores on a comprehension of instructions subtest compared to scores of their typically developing peers. Walberg and Magliano (2004) identified possible reasons for the discrepancy between word reading and comprehension skills for students with ASD: (a) inability to use background knowledge to interpret information and ambiguities presented in text, (b) fundamental deficits in language abilities, (c) difficulties with linguistic processing at the sentence level, and (d) inability to resolve ambiguity in text.

Several researchers have examined ways to teach reading comprehension to students with ASD. Chiang and Lin (2007) conducted a literature review to investigate reading comprehension strategies for students with ASD. They included 11 studies that examined participants, setting, text (academic reading) comprehension, functional sight words, comprehension, and instructional methods. Reviewed studies included 49 students with ASD (many of whom also were diagnosed with an intellectual disability) who were 3 to 17 years old. Four studies occurred in traditional schools (two in inclusion classrooms), three in private clinics or specialized schools, one in a home setting, and three studies did not specify the setting. Four studies examined text comprehension and the remaining seven studies examined sight word comprehension. A wide variety of instructional methods were represented including progressive time-delay, discrete-trial training, peer tutoring, cooperative learning groups, incidental teaching, computer-assisted instruction, priming, use of a cloze task (filling in blanks), and cueing students to attend. The authors stated that future research needs to apply National Reading Panel (2000) identified strategies for instructing students with ASD, further exploration of literacy instruction of students with Asperger syndrome, and further research on instructional methods for teaching reading comprehension to students with high func-
tioning autism or Asperger syndrome in the general education setting.

Other studies have examined reading comprehension skills of students with ASD. First, Reynhout and Carter (2008) studied the effects of a social story intervention on appropriate group reading behavior and specific attention on reading comprehension difficulties. They used a single-subject, ABC design to evaluate the impact of the social story review prior to group read aloud on the behavior of looking at the book. The participant was an 8-year-old diagnosed with ASD, intellectual disability, and limited language skills. The results failed to demonstrate a functional relation between the social story and appropriate group reading behavior. The targeted behavior also failed to improve after adding a verbal prompt to read the story. The authors suggest that one possible explanation for the lack of a functional relation could be that the participant was unable to comprehend the social story text and/or the text being read to the class. The authors stated that future research is needed to examine comprehension skills that act as prerequisites to social stories and to implement rigorous designs including multiple baseline across behaviors, settings, or participants.

Flores and Ganz (2007) analyzed the effects of a Direct Instruction (DI) reading program on reading comprehension skills. DI involves carefully designed explicit instruction, scripted lessons, high rates of student responses, and immediate feedback (Watkins & Slocum, 2004). DI has been successful with students with disabilities and students at risk for failure (Bursuck & Damer, 2011; Carnine, Silbert, Kame’enui, & Tarver, 2010). They used a multiple probe across behaviors design to measure the impact of DI on reading comprehension skills and behaviors of students with developmental disabilities, including ASD. Results showed a functional relation between DI and reading comprehension skills. The authors stated that future research is needed to examine the effects of the complete DI program (rather than just the selected portions used in this study), the modifications needed for students with ASD, and feasibility of implementation of this program.

Songlee, Miller, Tincani, Sileo, and Perkins (2008) examined the effects of a test-taking strategy on reading comprehension test scores. They used a multiple probe across participants design with four phases to analyze the impact of teaching the test-taking strategy to students with high-functioning ASD. Results showed a functional relation between acquisition of the test-taking strategy and an increase in reading comprehension test scores. Results also demonstrated generalization and maintenance of the test-taking strategy.

Finally, Hundert and van Delft (2009) used a multiple probe across behaviors design to analyze the effects of a system of least prompts on students’ ability to answer three types of inferential “why” questions: (a) questions based on a three-card picture sequence (a visual representation of a short sequence of events paired with a short text), (b) questions based on a vocally presented story, and (c) general information questions. Participants were three children with high functioning autism. Results showed a functional relation between the embedded instruction and the students’ ability to answer the three identified types of inferential “why” questions; however, students’ responses did not generalize to untrained inferential “why” questions. One notable limitation of this study is the insufficient baseline data collected prior to implementation of intervention. At times there were insufficient data to establish a trend, and at other times an increasing trend was established in baseline. The authors stated that future research is needed to replicate these results in school settings during classroom routines and implemented by general education teachers, special education teachers, or paraprofessionals. The authors also stated that further research is needed to determine if an increase in ability to answer inferential “why” questions would translate into an increase in overall reading comprehension skills.

One method suggested for increasing reading comprehension skills for struggling readers (e.g., students with learning disabilities) is graphic organizers (Jiang & Grabe, 2007). Graphic organizers are visual representations of the information conveyed in a text (Jiang & Grabe, 2007). Graphic organizers have been used to increase reading comprehension skills in young readers and second language students (Jiang & Grabe, 2007).
ers are currently in wide use by reading experts and classroom teachers, and studies have shown a positive effect on reading comprehension (Carnine et al., 2010; Jiang & Grabe, 2007).

Graphic organizers and visual supports have also been recommended to teach reading comprehension skills to students with ASD (Gately, 2008). Visually cued instruction, including graphics, story/visual maps, and goal structure mapping, can help students with ASD focus on key information, increase independence and memory (Gately, 2008). Although graphic organizers and visual supports have been successful teaching reading comprehension skills to nondisabled young readers, and have been recommended as an effective teaching tool for students with ASD, further empirical research is needed to demonstrate its use to teach reading comprehension to students with ASD. Therefore, the purpose of this study was to examine the effects of wh-question (i.e., who, what, where, what doing) graphic organizers on reading comprehension skills in students with ASD.

Method

Participants

Participants were three elementary students diagnosed with ASD who demonstrated deficits in the area of reading comprehension. Convenience sampling was used to identify (a) ability to orally read text at a minimum of a 1st grade level, (b) ability to match written nouns to picture representations, (c) inability to accurately answer literal wh-questions (who, what, where, and what doing) about a previously read text, (d) inability to sort written text into a graphic organizer according to the corresponding type of literal wh-question (who, what, where, and what doing), and (e) signed parental consent and student assent. Although the participants of this study represent a range of races and ethnicities, participants were not selected based on race or ethnicity, rather this range was the result of the convenience sampling used to identify participants meeting the above criteria.

Aaron (pseudonyms used throughout) was an 8-year-old boy diagnosed with autism and was Hispanic. Aaron attended a public elementary school in a separate classroom designed for children with autism due to a combination of his challenging behaviors, such as aggression, tantrums, and lack of academic progress. Aaron’s IQ was tested with the Reynolds Intellectual Assessment Scale and was reported as 94. Aaron was able to vocally engage in conversational language when the topic or person was of interest to him; however, he most often used his language to request basic wants and needs. Aaron received reading instruction using a leveled Direct Instruction program. The teacher reported that while Aaron could read sentences and paragraphs fluently, he was unable to answer comprehension questions about a passage.

Mark was a 10-year-old boy diagnosed with autism and was Caucasian. Mark attended the same public elementary school as Aaron, but was placed in a different separate classroom designed for children with autism. Mark demonstrated severely restricted interests and stereotypic behaviors as well as delays in academic progress. Mark’s IQ was tested and reported as 67. Mark was able to vocally communicate his wants and needs and respond to some intraverbal questions (e.g., “What’s on your head?” “A hat.”). Mark was beginning to read sentences and paragraphs, but his fluency and intonation were not yet mastered. Mark’s teacher reported that Mark was unable to answer any comprehension questions about a passage he had previously read.

Joe was a 10-year-old boy who was also diagnosed with autism and was African American. Joe also attended the same public elementary school and was placed in a third classroom designed to meet the needs of students with autism. Joe demonstrated restricted interests and stereotypic behaviors as well as delays in academic progress. Additionally, Joe had some limitations in his articulation, making it difficult to understand his vocal communication. Joe demonstrated high levels of receptive language; however, he relied on short words/phrases, gestures, and writing words to express his wants and needs. Joe’s IQ was tested using the Leiter International Performance Scale and was reported as 90. Joe was able to read sentences and paragraphs; however, when he read aloud it was difficult to understand his speech. The teacher reported Joe was
unable to answer comprehension questions about passages he had read.

Setting
The study took place in a public elementary school located in an urban area of the southeastern United States. Instruction occurred at the students’ desk or at a designated reading instruction table in each student’s classroom. The classrooms were each designed specifically for students with autism, but were not otherwise modified for this study. Each classroom was staffed with one special education teacher and one paraprofessional. The first author served as the experimenter and conducted the intervention and probes throughout the study. A graduate student from a local university served as a data collector for interobserver reliability and procedural fidelity.

Materials
The graphic organizer was an 8.5" by 11" paper divided into four columns. Each column was labeled in large font according to the type of wh-question as follows: Who? (person), Where? (place), What? (thing), and What doing? (event). Reading materials were selected from the students’ current level in a Direct Instruction reading program and was between two to four pages in length. During each intervention and probe session, students read from the next story, ensuring that they never encountered the same story twice. Targeted text was changed for each session and probe to prevent students from memorizing responses.

Data Collection

Dependent variables. Data were collected on two dependent variables. First, data were collected on the accuracy of sorting words (presented as text) into the graphic organizer corresponding to the type of wh-question. In these probes, occurring after reading the passage but prior to answering wh-questions, students were shown a list of words by the researcher and were asked to sort (i.e., write) each word into the corresponding category. For example, when presented with the word “mom,” the response was scored correct if the student wrote the word “mom” into the “who” category. When presented with the word “running,” the response was scored correct if the student wrote “running” in the “what doing” category. This resulted in the student having a completed graphic organizer where key text from the story was sorted into the appropriate categories.

The primary dependent variable was accuracy answering eight, literal recall (i.e., stated in the text) wh-comprehension questions. This occurred after the student completed the graphic organizer. The eight questions comprised two questions from each category: who, what, where, and what doing. The researcher asked two questions of each type to ensure that students had to respond according to the type of information requested and ensure that the response was correct according to the passage. For example, if the passage read “Nancy played with dolls on the floor. Mom cooked dinner in the kitchen.” The student would have sorted “Nancy” and “mom” into the “who” category and “floor” and “kitchen” into the “where” category of the graphic organizer. Students had to use knowledge from the story to answer “Who cooked dinner?” If the student responded “mom,” his response was scored correct. If the student responded “Nancy,” it was scored incorrect and might indicate that the student did not comprehend the passage and answered based on the information found on the graphic organizer alone. If the student responded “kitchen,” it was counted incorrect and might indicate that he did not comprehend the type of question asked. Aaron and Mark answered the questions vocally. Joe answered vocally, but also pointed to the written word on the graphic organizer. Responses were recorded as correct or incorrect, and the number of independent correct responses was graphed on an equal interval graph (see Figure 1).

Interobserver reliability. Interobserver reliability was conducted by comparing the second observer’s scores with the experimenter’s scores. Reliability sessions were evenly distributed across baseline, intervention, and maintenance phases for all students. Interobserver reliability was scored item-by-item (Cooper, Heron, & Heward, 2007) and calculated by dividing the total disagreements by the total opportunities and multiplying by 100. For
Aaron, interobserver reliability was conducted in 31.25% of total sessions and averaged 97.5% (range 87.5%–100%). For Mark, interobserver reliability was also conducted in 31.25% of total sessions and averaged 97.5% (range 87.5%–100%). For Joe, interobserver reliability was conducted in 33.33% of total sessions and averaged 100%.

Social validity. Social validity questionnaires can obtain stakeholders’ (e.g., participants, teachers, parents) opinions about an intervention’s goals, outcomes, and methods (Cooper et al., 2007). After the maintenance phase of the study the students and their classroom teachers completed questions that addressed social validity. The teachers’ questionnaire included six items ranked on a Likert scale with possible scores ranging from one through five with one representing strongly disagree and five representing strongly agree. The experimenter read five questions to students and they responded by circling the word (paired with a symbol) yes or no.

Experimental Design
This study used a delayed multiple baseline across participants design (Heward, 1978 as cited in Cooper et al., 2007). Multiple baseline designs are a common research tactic in single-subject research and ideal for research with special needs populations when a large sample of participants is not available (Cooper et al., 2007; Kennedy, 2005; Tawney & Gast, 1984). There was a baseline phase, followed by an intervention phase and a final maintenance phase. The second student was introduced to the intervention once the first student’s data path showed a change in level and/or trend. The third student was introduced to the intervention in the same manner once the second student’s data path demonstrated a change in level and/or trend. The intervention ended once the student correctly answered at least seven out of eight comprehension questions about a text in three consecutive sessions and at least five data points per phase were established.

Procedure
Baseline. Data were collected during each baseline session by asking the student to complete two types of probes: (a) sorting eight words into corresponding wh-question categories on a graphic organizer (two of each type of wh-question), and (b) answering eight wh-questions about a text (two of each type of wh-question). The experimenter did not provide corrective feedback on students’ responses during probes. Once a minimum of five data points and a stable or decreasing trend was established, the first student began the intervention phase. Baseline probes were conducted daily following completion of the reading passage. Reading passages were changed for each probe to ensure students were unable to memorize responses. Each probe was conducted in the classroom during typical instructional time and lasted approximately 10 min.

Graphic organizer. The intervention using the graphic organizer was conducted in the
same manner as the baseline sessions with the exception of the introduction of prompting to teach the students to use the graphic organizer correctly. The experimenter used a least to most prompting hierarchy (independent, verbal, gesture, and physical) to teach each student to sort words into corresponding categories on a graphic organizer and answer corresponding comprehension questions from a short passage. For example, the experimenter presented the graphic organizer and a list of eight words (two from each category). The experimenter asked the student to write (i.e., sort) the words into the corresponding categories, providing least to most prompting as needed. After the words were successfully sorted, the experimenter presented a short reading passage (passages were never repeated), the student read it aloud, and then the experimenter presented eight corresponding wh-questions about the text, and prompted the student to use his graphic organizer. As described previously, the eight questions included two questions from each category: who, what, where, and what doing. This was to ensure that students had to respond according to the type of information requested and ensure that the response was correct according to the passage. Students were able to look back at the text or graphic organizer at any point.

Generalization and maintenance. Generalization data were taken once in baseline and once in maintenance by scoring permanent products from each student’s responses to questions from their reading in their special education classrooms. The students participated in a leveled Direct Instruction reading program that contained reading comprehension worksheets. The examiner scored the students’ responses to literal wh-questions on these worksheets that were completed with their regular teachers during reading instruction. The students did not use the graphic organizers during their teachers’ reading instruction. Maintenance data were collected once per week after the student completed the intervention phase. Aaron had five weeks of maintenance data, Mark had four weeks of maintenance data, and Joe had three weeks of maintenance data.

Procedural reliability. Procedural reliability served as the primary method of assessing accuracy of implementation of the intervention. The second observer watched sessions and completed a procedural reliability checklist. Procedural reliability was evenly distributed across baseline, intervention, and maintenance sessions. Reliability was conducted in 31.25% of Aaron and Mark’s sessions and averaged 100%. Procedural reliability was conducted in 33.33% of Joe’s sessions and averaged 100%.

Results

Figure 1 shows the number of independent correct responses to the wh-questions for Aaron, Mark, and Joe. The x-axis represents probes and the y-axis represents the number of independent correct responses during each probe. The solid, horizontal lines indicate the mean of words sorted correctly in the graphic organizer per phase. All three participants demonstrated an immediate change in level after instruction on the graphic organizer began, and quickly met the criteria to move to maintenance.

During baseline, Aaron’s performance averaged 3.8 correct answers on comprehension probes. The data were low and somewhat variable. During baseline Aaron averaged 4.8 independent correct words sorted in the graphic organizer. During the intervention phase Aaron averaged 7.2 correct answers to wh-questions. The data in this phase were high and stable. During the intervention phase Aaron 7.8 independent correct words sorted in the graphic organizer. During maintenance Aaron averaged 7.6 independent correct answers to wh-questions per probe. His data in this phase remained high and stable. During maintenance Aaron averaged 7.8 independent correct words sorted onto the graphic organizer.

During baseline, Mark’s performance averaged 2.7 correct answers to wh-questions per probe. The data were low and showed some variability. During baseline Mark averaged 4.7 independent correct words sorted in the graphic organizer. During the intervention phase Mark averaged 6.4 correct answers to wh-questions per probe. Mark’s data in this phase were high, stable, and showed an increasing trend. During the intervention phase Mark averaged 7.0 independent correct words
sorted in the graphic organizer. During main-
tenance Mark averaged 7.25 correct answers
to wh-questions per probe. His data in this
phase remained high and stable. During
maintenance Mark averaged 7.5 independent
correct words sorted onto the graphic organ-
izer.

Joe’s performance in baseline averaged 3.6
correct answers to wh-questions per probe. The
data were low with a fair amount of vari-
ability. During baseline Joe averaged 5.6 inde-
pendent correct words sorted in the graphic
organizer. During the intervention phase Joe
averaged 6.1 independent correct answers to
wh-questions per probe. His data in this phase
were higher than the baseline phase, but ini-
tially demonstrated some variability before
stabilizing. During the intervention phase
Joe averaged 7.4 independent correct words
sorted in the graphic organizer. During main-
tenance Joe averaged 7.0 correct answers to
wh-questions per probe. Joe’s data in this phase
were higher than the baseline phase, but ini-
tially demonstrated some variability before
stabilizing. During the intervention phase
Joe averaged 7.4 independent correct words
sorted in the graphic organizer. During main-
tenance Joe averaged 7.3 independent correct words
sorted in the graphic organizer.

Generalization

Data were collected in the form of percent of
wh-questions answered correctly from stu-
dents’ reading instruction from their special
education teachers. All three students’ scores
improved from baseline to generalization. Aaron’s baseline score was 0%, and improved
to 100% during the generalization probe that
was conducted during the maintenance phase.
Mark’s baseline score was 33%, and improved
to 75% during the generalization probe that
was conducted during the maintenance phase.
Joe’s baseline score was 40%, and improved
to 100% during the generalization probe that
was conducted during the maintenance phase.

Social Validity

Teachers reported that they strongly agreed
that reading comprehension is useful in the
student’s daily life. The teachers rated the
following questions as either a 4 or 5 on a
5-point scale: (a) they agreed that graphic
organizers helped the student answer compre-
hension questions, (b) graphic organizers are
an appropriate way to help teach wh-questions
about a text, (c) the student has answered
wh-questions in the classroom, and (d) the student increased his ability to answer wh-
questions from reading in general. Finally,
teachers had some disagreement on whether
the student had been able to accurately an-
swer reading comprehension questions dur-
ing regular English language arts instruction/
home. One teacher rated this as a 2,
another as a 4, and the last as a 5. Teachers
also had the opportunity to make open ended
comments. Two teachers responded, “The
graphic organizer has been very helpful to my
student in making progress on reading com-
prehension. The hands-on visual approach is
great,” and “Using the graphic organizer
helped clarify the important information from
extraneous information.”

Students all responded “yes” to the follow-
ing questions: (a) “Did the graphic organizer
help you learn about the types of wh-ques-
tions?,” (b) “Do you read a lot during your
day?,” and (c) “Did you like learning this way?”
Two of the three students responded “yes” to
the following two questions: (a) “Did you like
the lessons?,” and (b) “Did the graphic orga-
nizers help you answer questions about what
you read?”

Discussion

The purpose of this study was to examine the
effects of wh-question graphic organizers on
reading comprehension skills in students with
ASD. A functional relationship was shown be-
tween the use of graphic organizers and the
students’ correct responses to literal wh-ques-
tions from short passages. The students’ per-
formance also demonstrated a high level of
maintenance, and they were able to answer
wh-questions from three to five weeks after
stopping the intervention. Additionally, all
three students were able to demonstrate gen-
eralization to their reading program work-
sheets administered during regular reading
instruction.

Graphic organizers have been used success-
fully to teach reading comprehension skills to
a wide range of students such as young readers
and second language students (Carnine et al.,
2010; Jiang & Grabe, 2007). Additionally,
graphic organizers and visual supports have
been recommended to teach reading comprehension skills to students with ASD as a way to utilize visually cued instruction (Gately, 2008). This study demonstrates that use of graphic organizers can help students with ASD improve their ability to answer literal wh-questions about a text. However, this study focused on students with ASD who met a narrow set of criteria and the extent to which graphic organizers can help other students with ASD improve reading comprehension is unclear. This study focused on students with ASD who were vocal, able to communicate basic wants and needs, able to read aloud, but had difficulty with reading comprehension. Specifically, they lacked skills at answering literal comprehension questions. Therefore, the extent to which a graphic organizer could help students with ASD answer inferential comprehension questions is not clear.

Limitations and Future Research

This study was the first to use graphic organizers to teach reading comprehension skills to students with ASD. Additional replications of this study with a variety of participants, settings, and researchers are required in order for the use of graphic organizers to teach reading comprehension, specifically answering literal wh-questions, to become an established intervention.

A second limitation of this study is the experimenter’s role as primary instructor throughout implementation of the study. Classroom teachers were not asked to administer the intervention. This limits the authors’ ability to predict whether teacher implementation would maintain high treatment fidelity. Future studies could use classroom teachers to deliver the graphic organizer intervention.

A third limitation was that instruction was delivered during one-to-one sessions. Since group instruction is commonly used in classrooms, future research could evaluate instruction with graphic organizers during small or whole group instruction.

Implications for Practice

Results of this study suggest that graphic organizers might help students with ASD improve their literal reading comprehension. The four-column wh-question graphic organizer (available from the corresponding author) used in the study was low cost and easy to make on a word processor, making it easy for teachers to supplement their reading instruction. Social validity findings from teachers were generally positive, suggesting that graphic organizers might be a strategy they would use when teaching reading comprehension to students with ASD.

References

Bursuck, W. D., & Damer, M. (2011). Reading instruction for students who are at risk or have disabilities (2nd ed.). Boston, MA: Pearson Education, Inc.
National Reading Panel. (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for


Received: 9 February 2012
Initial Acceptance: 5 April 2012
Final Acceptance: 10 June 2012

---

244 / Education and Training in Autism and Developmental Disabilities-June 2013