Twenty-two children (age range of 3.5–10.92 years old) with autism spectrum disorder (ASD) were assessed using the Test of Gross Motor Development (Second Edition; TGMD-2) using three different protocols. The total duration of assessment time and the percentage of time engaged in on-task behavior during the assessments were measured and analyzed using within-subjects repeated measure ANOVA designs to compare performance across the three protocols. Significant differences emerged across the duration of assessment time by assessment protocol, while no significant differences emerged for time on-task during the assessments by protocol used. In addition, correlations were calculated between the TGMD-2 scores and the duration of assessment time and the percentage of time on-task. An inverse relationship was found between TGMD-2 scores and total duration of assessment time by protocol used, ($r = .726, .575, .686$), while a positive relationship was found between the TGMD-2 scores and time on-task ($r = -.570, -.535, -.798$). These results suggest a direct relationship between skill proficiency and contextually appropriate behaviors.

**Keywords:** visual supports, Test of Gross Motor Development (Second Edition), fundamental motor skills

According to the American Psychiatric Association (APA; 2013), autism spectrum disorder (ASD) includes a group of developmental disabilities that can cause significant difficulty in communication and social interaction. In addition, children with ASD often display repetitive and stereotypic behaviors that are deemed contextually inappropriate and may have an impact on performance in the classroom (McAllister & Maguire, 2012). Because individuals with ASD often have difficulty understanding the spoken word and processing auditory information, it is important...
for educators to understand and accommodate the special needs of this diverse population. One such strategy that can increase structure and thereby maximize time on-task is the use of visual supports (Bryan & Gast, 2000; MacDuff, Krantz, & McClannahan, 1993; Schneider & Goldstein, 2010).

Visual supports are instructional aides including concrete boundaries, picture task cards, and picture activity schedules. These are believed to direct attention to the relevant stimuli within the task and to display the abstract constructs of the task in concrete ways. These strategies help organize the surrounding environment, thereby increasing understanding of the social and verbal cues while reducing confusion (Breslin & Rudisill, 2011; Bryan & Gast, 2000; Schneider & Goldstein, 2010). Visual supports are recommended for use with children with ASD to increase student understanding in a variety of educational settings including physical education (Fittipaldi-Wert & Mowling, 2009; National Research Council, 2001; Odom et al., 2003; O’Reilly, Sigafoos, Lancioni, Edrisinha, & Andrews, 2005; Rao & Gagie, 2006). By helping students with ASD focus on task relevant stimuli, visual supports may help them reduce contextually inappropriate behavior and improve functioning in the environment. This is a particularly positive outcome for educators who are interested in maximizing students’ time engaged in on-task behaviors.

It is possible that the incorporation of visual supports into the administration of this motor skill assessment may lead to increased scores on assessments by children with ASD. Indeed, Breslin and Rudisill (2011) have found that when a picture task card was presented during the administration of the TGMD-2 and verbal instruction was minimized, higher gross motor quotient scores were found when compared with (a) a traditional approach utilizing conversational sentences and physical demonstrations or (b) an approach utilizing a picture activity schedule and reduced verbal instruction. A picture task card is defined as the visual depiction of an action provided to an individual when verbal communication is difficult, while a picture activity schedule is a visual representation of the order of the actions and behaviors to be undertaken to complete a task (Bryan & Gast, 2000; Breslin & Rudisill, 2011; Liu & Breslin, 2013). Liu and Breslin (2013) found the same outcome when a picture activity schedule was used during the administration of the Movement Assessment Battery for Children (Second Edition).

The performance of children with ASD in learning environments can be measured by exploring the incidence of stereotypic behaviors or by measuring positive outcomes such as the time engaged in on-task behavior. Time on-task can be defined as the time during which they (a) visually attend to appropriate equipment or demonstration for the task at hand, (b) look at and manipulate the appropriate materials for completing the task, (c) appropriately perform the activities within the task, or (d) are in transition from one skill to another (Bryan & Gast, 2000). Beyond engagement in on- or off-task behavior in a learning environment, children also may engage in administrative activities (such as tying shoes or obtaining supplies) that are not associated with the task at hand, but are necessary activities to proceed with the task at hand.

Although time on-task is most commonly measured during instruction, it was the purpose of this study to examine the construct during episodes of student assessment and motor skill assessment, particularly. The Test of Gross Motor Development (Second Edition; TGMD-2) is a tool used to assess the gross motor skill performance of children between the ages of 3–10 years. Within the TGMD-2,
children’s performance in 12 different skills is assessed. These skills include running, galloping, sliding, hopping, leaping, horizontal jumping, catching, kicking, dribbling, overarm throwing, underarm rolling, and striking a stationary object (Ulrich, 2000). Nonetheless, studies that have incorporated the TGMD-2 have not examined time engaged in on-task behavior during its administration. As children with ASD who exhibit problematic, off-task behaviors can disrupt the learning process for themselves and their peers in an educational setting (Donahue, Casey, Bicard & Bicard, 2012), educators are very interested in maximizing students’ time engaged in on-task behaviors. Little attention has been paid to the impact of on-task and off-task behaviors in performance during assessments. Thus, our study sought to answer three questions regarding the role of on-task performance during motor skill assessments.

There were three specific purposes of this study. First, we sought to explore if there were differences in the total amount of time it took children with ASD to complete the TGMD-2 assessments using the different administrative protocols. Specifically, it was hypothesized that a picture activity schedule protocol would result in the shortest duration of time to complete the TGMD-2 assessment, while a picture task card protocol would take more time. In addition, the TGMD-2 would take the longest to complete using a traditional protocol as described in the TGMD-2 Examiner’s Manual. The picture activity schedule protocol was hypothesized to yield the shortest duration of time for the TGMD-2 because it provides the most detailed information regarding what happens next, thereby reducing anxiety for the children with ASD (Downing & Peckham-Hardin, 2000).

The second purpose of this study was to determine if there were differences in the percentage of time engaged in on-task behaviors across the three different administrative protocols. It was hypothesized that the use of visual supports would increase the percentage of time engaged in on-task behaviors on the TGMD-2 compared with the recommended traditional protocol for this assessment.

The third purpose of this study was to examine the relationship between the amount of time it took to conduct the TGMD-2 assessments, the time on-task during the assessments, and the gross motor quotient scores on the TGMD-2 by children with ASD. It was hypothesized that an inverse relationship would exist between the performance scores and the length of time of assessments. That is, TGMD-2 assessments would take less time to complete with children who scored higher on the TGMD-2 assessments than children with lower TGMD-2 scores. It also was hypothesized that a positive relationship would exist between gross motor quotient scores and the time engaged in on-task behaviors during the TGMD-2 assessment.

Method

Participants

The participants in this study were also the participants in the Breslin and Rudisill (2011) study. Both studies were conducted during a summer supplemental educational program in a small city in the southeastern United States enrolling 35 students with ASD. Upon approval from the university institutional review board and the return of parental informed consent, N = 30 (male = 23, female = 7) participants were recruited. According to parent report data, the participants met the inclusion
criteria for ASD according to the *Diagnostic and Statistical Manual* (APA, 2013) as measured by a developmental pediatrician or a trained, licensed psychologist. Four of the participants were older than 10 years, and were not included in this study since the TGMD-2 would not be a valid assessment for children that age (Ulrich, 2000). Four additional participants were excluded from data analysis due to absences resulting in incomplete data, yielding a final sample of \(N = 22\) (male = 16, female = 6; African American = 5, Caucasian = 17), ranging in age from 3.5 years–10.92 years. According to parent report data, 72% of the participants included in the final sample were diagnosed with autistic disorder, 18% had pervasive developmental disorder not otherwise specified (PDD-NOS), and 4.5% had Asperger’s disorder or autistic disorder and PDD-NOS.

**Setting**

Data collection took place within a large multipurpose room located in the elementary school hosting a supplemental summer educational program for children with ASD. The doors to the multipurpose room were shut and signs were placed on the doors asking people to use alternate routes while testing was occurring. The dimensions of the area where data collection occurred was 10.97 m × 25.07 m. This was the same setting used in the Breslin and Rudisill (2011) study.

The TGMD-2 assessments were administered by a female motor development specialist with over one year experience working with children with ASD as the instructor of gross motor play and physical activity at a university-affiliated preschool program for children with ASD. Two other female research assistants were present in the testing room, one serving as the videographer, and one setting up and collecting equipment before and following the performance of the motor skill trials for each skill. The same research assistants fulfilled the same roles each day of data collection. Because children with ASD are known to have difficulty attending to relevant environmental stimuli, measures were taken to decrease the amount of irrelevant stimuli in the assessment environment (Larson, 2006). The researcher attempted to control for distractions in the assessment environment by creating a routine (Larson, 2006) and utilizing concrete boundaries (Simpson, Myles, & LaCava, 2008).

**Instruments**

**TGMD-2.** The TGMD-2 is a widely used assessment in physical education (Ulrich, 2000). The children’s performance is compared with criteria based on age and sex to determine a score of motor performance. TGMD-2 data collection and analysis was collected using the procedures described in the TGMD-2 Examiner’s Manual (Ulrich, 2000) or as described previously in the literature to include a picture activity schedule and picture task cards (Breslin & Rudisill, 2011). The principal investigator of this study and a research assistant blind to the purpose of the study coded to criterion all of the children’s TGMD-2 performance. Then, interrater reliability for TGMD-2 criterion coding was calculated at 95.74%.

**Behavior Evaluation Strategy and Taxonomy (BEST) Software.** BEST is a software system (developed by Educational Consulting, Inc.; Thousand Oaks, CA) that enables quantitative data to be collected and analyzed from a physical activity
environment (Sharpe & Koperwas, 1999). BEST uses computer codes programmed by the researcher to measure the frequency in which the participants engage or do not engage in specified motor and physical behaviors in a real time, observational format. Following software configuration, percentage of time engaged in on-task behavior recorded on the videotapes was analyzed using the BEST software. The computer keyboard was configured according to the researcher’s predetermined taxonomy which allowed for the recording of frequency and duration of operationally defined behaviors.

Specifically, BEST was programmed to produce an auditory tone every six seconds informing the researchers it was time to code the participant’s behavior shown on the videotape for the previous six seconds. In order for a participant’s behavior to be coded as on-task, the participant must have engaged in the behavior defined as on-task for the entire six second observation period. If the child engaged in behavior indicated he or she was off-task for part of the observation period, the child would be coded as off-task. Whole interval data were used to reduce the likelihood that more than one behavior would occur in each six second observation period (van der Mars, 1989).

Using the BEST software, the total duration of the assessment was calculated. To determine the total duration of the assessment, the BEST software kept a tally in minutes and seconds of how long the software was engaged in a data analysis session. The clock began when the key coded to mark the start of data collection was depressed, and stopped when the key to terminate data collection was pressed. To calculate the percentage of time engaged in on-task behavior, BEST software divided the time duration in seconds coded as “ontask” into the time duration in seconds of the entire assessment (Sharpe & Koperwas, 1999). This number, multiplied by 100, yielded a percentage of time on-task. Any time spent engaged in administrative tasks (e.g., tying shoes, retrieving equipment) necessary for the TGMD-2 to continue (albeit not a part of the TGMD-2) was excluded from data analysis.

**Design and Procedures**

TGMD-2 assessments were conducted with one participant at a time, one assessment per participant per day, in a counterbalanced random order to ensure learning would not impact results. The TGMD-2 was terminated upon completion of the assessment items. In the event that a child was sick and/or totally noncompliant during an assessment period, the child was given another opportunity to participate on the next scheduled day of data collection.

The videotapes were also analyzed by two research assistants to determine the participants’ percentages of time engaged in on-task behavior during the TGMD-2. To code this data, both researchers were provided a list explaining which trials (i.e., the last trial of three trials) were to be coded as extra trials, but the data were coded individually. After the quantitative data were coded for each observation period, total duration and/or frequency scores were be calculated for each key (representing a frequency and duration of an operationally defined behavior). Using these data, interrater agreement was calculated using Cohen’s kappa coefficient (Cohen, 1960). A kappa coefficient between .80 and 1.0 was considered appropriate, indicating high interrater agreement.
Statistical Analyses

To examine differences in assessment duration and the percentage of time engaged in on-task behavior, the independent variable in this study was the different protocols for the TGMD-2. The traditional protocol incorporated the format of the TGMD-2 protocol that uses conversational sentences to describe assessment instructions and physical demonstrations as described by the TGMD-2 Examiner’s Manual (Ulrich, 2000). The other two protocols were modified formats of the TGMD-2 protocol. These protocols incorporated the use of task cards with modeling and short verbal commands, and the use of a picture activity schedule with modeling and verbal commands. The dependent variable in this study was the percentage of time-on-task behavior. In addition, data regarding gross motor quotient scores on the TGMD-2, a measure of motor performance, from the Breslin and Rudisill (2011) study were used to calculate correlational relationships between assessment duration of time, percentage of time on-task, and motor performance as measured by the TGMD-2. Pearson-product moment correlations were calculated using a one-tailed test at $\alpha = .05$ (Thomas, Nelson, & Silverman, 2005).

To ensure interrater reliability, data coders were trained to code participant behavior using the BEST software for the percentage of time-on-task behavior. Training continued until the data coders reached 90% agreement in their codes, and then all of the assessments were checked for agreement in their codes of percentages of time on-task behavior to reduce the influence of expectancy bias because the primary investigator was one of the two data coders. Interrater reliability between the researcher and the research assistant trained to code the percentage of time on-task was 95.84%.

Results

All data analyses were conducted using the Statistical Package for the Social Sciences, Version 16 (IBM Corporation). Univariate ANOVAs testing the assumption that there was no statistically significant differences for the percentage of time on-task as a function of the order of protocols (i.e., traditional protocol, picture task card condition, picture activity schedule condition) were not significant, $F(5,16) = .371$, $\eta^2 = .023$, $p = .693$. This analysis indicated that the participants’ percentages of time engaged in on-task behavior were not influenced by the order in which they received the three protocols. Thus, the results of the study were not influenced by a learning effect or a change in routine during the supplementary summer education program.

Table 1 provides the gross motor quotient scores, time to completion, and percentage of time engaged in on-task behavior for each participant. The first research question examined if the protocol received by the children with ASD influenced the total duration of time on the TGMD-2 assessment. A repeated measures univariate ANOVA revealed a significant main effect for protocol, $F(2,20) = 6.039$, $\eta^2 = .860$, $p = .005$). The means and standard deviations for the total duration of time in minutes by protocol (i.e., traditional protocol, picture task card, and picture activity schedule) were 13.11 ± 3.88, 13.75 ± 4.77, and 15.43 ± 4.41, respectively. Bonferroni follow-up tests indicated that the picture activity schedule condition was significantly longer in duration than the traditional protocol ($p < .0001$), but
there were no differences in duration between the picture task card condition and the traditional protocol \((p = 1.000)\) or between the picture activity schedule condition and the picture task card condition \((p = .162)\). These findings show that the protocols produced significant differences in the total duration of time of TGMD-2 assessment, with the picture activity schedule condition significantly longer in duration than the other two protocols.

The second research question examined if the protocols received by children with ASD influenced performance as measured by the percentage of time engaged in on-task behavior. A repeated measures univariate ANOVA revealed the main effect for protocol was not significant, \(F(2,21) = .425, \eta^2 = .020, p = .657\). The means and standard deviations for the percentage of time engaged in on-task behavior by protocol (i.e., traditional protocol, picture task card, and picture activity schedule) were \(73.78\% \pm 15.29\%, 76.16\% \pm 18.86\%,\) and \(75.49\% \pm 18.71\%\), respectively. These findings show that the protocol did not produce differences in the percentage of time on-task.

The third research question explored the relationships between time engaged in on-task behavior and performance on the TGMD-2 across protocols (i.e., traditional, protocol, picture task card, and picture activity schedule). The Pearson product-moment correlation \((df = 20, p < .05,\) two tails) between the total duration of the assessment and gross motor quotients across protocols was \(-.570, -.535,\) and \(-.798\), respectively. The \(r^2\) values of \(.325, .286,\) and \(.637\), respectively, suggest a moderate inverse relationship between total duration and gross motor quotient using a traditional or picture task card protocol and a strong inverse relationship using a picture activity schedule protocol (see Figure 1). The Pearson product-moment correlation \((df = 20, p < .05,\) two tails) between the percentage of time engaged in on-task behavior and gross motor quotient across respective protocols was \(.726, .575,\) and \(.686\), while the \(r^2\) values of \(.527, .331,\) and \(.471\), respectively suggest a positive relationship between the percentage of time engaged in on-task behavior and gross motor quotient in the traditional and picture task card conditions (see Figure 2).

**Discussion**

The first hypothesis proposed that the picture activity schedule protocol would result in the shortest duration of time to complete the TGMD-2 assessment, while the picture task card protocol would take more time than the picture activity schedule protocol, and the TGMD-2 would take the longest to complete using a traditional protocol. This hypothesis was not supported. Significant differences were found with respect to duration in minutes of the TGMD-2 protocols. During both the traditional protocol and the picture task card conditions, the TGMD-2 took approximately 13 min to complete; however, the picture activity schedule condition took an additional two and a half minutes to complete. Therefore, the picture task card condition was almost the same duration as the traditional protocol, and the picture activity schedule protocol was significantly longer in duration than the other two protocols. It is possible that the children with ASD were unable to pay attention or lost interest during the picture activity schedule condition as a result of the extended time required for testing. Explaining this finding and the previous finding that revealed administering the TGMD-2 using a picture task yields higher gross motor quotient scores than a traditional protocol (Breslin & Rudisill, 2011),
Table 1  Participants’ Gross Motor Quotient Scores, Time to Completion in Minutes, and Percentage of Time Engaged in On-Task Behaviors as a Function of TGMD-2 Administrative Protocol

<table>
<thead>
<tr>
<th>Participant</th>
<th>GMQ</th>
<th>TtC</th>
<th>%ToT</th>
<th>GMQ</th>
<th>TtC</th>
<th>%ToT</th>
<th>GMQ</th>
<th>TtC</th>
<th>%ToT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46</td>
<td>17.50</td>
<td>53.69</td>
<td>46</td>
<td>11.31</td>
<td>69.22</td>
<td>46</td>
<td>21.26</td>
<td>46.50</td>
</tr>
<tr>
<td>2</td>
<td>88</td>
<td>9.41</td>
<td>83.80</td>
<td>103</td>
<td>10.2</td>
<td>84.01</td>
<td>91</td>
<td>12.20</td>
<td>82.55</td>
</tr>
<tr>
<td>3</td>
<td>46</td>
<td>16.70</td>
<td>40.95</td>
<td>46</td>
<td>18.8</td>
<td>29.05</td>
<td>46</td>
<td>19.06</td>
<td>31.49</td>
</tr>
<tr>
<td>4</td>
<td>82</td>
<td>10.72</td>
<td>88.46</td>
<td>91</td>
<td>8.00</td>
<td>94.51</td>
<td>97</td>
<td>7.90</td>
<td>93.36</td>
</tr>
<tr>
<td>5</td>
<td>73</td>
<td>8.32</td>
<td>91.19</td>
<td>91</td>
<td>8.52</td>
<td>89.10</td>
<td>88</td>
<td>9.51</td>
<td>94.02</td>
</tr>
<tr>
<td>6</td>
<td>64</td>
<td>15.49</td>
<td>87.29</td>
<td>70</td>
<td>14.71</td>
<td>80.84</td>
<td>64</td>
<td>21.99</td>
<td>51.50</td>
</tr>
<tr>
<td>7</td>
<td>58</td>
<td>14.86</td>
<td>62.15</td>
<td>61</td>
<td>23.49</td>
<td>43.16</td>
<td>61</td>
<td>18.29</td>
<td>61.38</td>
</tr>
<tr>
<td>8</td>
<td>52</td>
<td>14.00</td>
<td>70.31</td>
<td>52</td>
<td>14.69</td>
<td>85.62</td>
<td>55</td>
<td>19.30</td>
<td>76.24</td>
</tr>
<tr>
<td>9</td>
<td>73</td>
<td>11.10</td>
<td>84.64</td>
<td>76</td>
<td>12.92</td>
<td>71.73</td>
<td>82</td>
<td>13.52</td>
<td>76.45</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>10.80</td>
<td>82.80</td>
<td>97</td>
<td>12.4</td>
<td>93.72</td>
<td>91</td>
<td>11.50</td>
<td>91.06</td>
</tr>
<tr>
<td>11</td>
<td>46</td>
<td>14.36</td>
<td>76.62</td>
<td>73</td>
<td>15.86</td>
<td>68.07</td>
<td>58</td>
<td>13.62</td>
<td>62.24</td>
</tr>
<tr>
<td>12</td>
<td>61</td>
<td>11.09</td>
<td>63.90</td>
<td>79</td>
<td>8.90</td>
<td>97.53</td>
<td>79</td>
<td>12.71</td>
<td>91.41</td>
</tr>
<tr>
<td>13</td>
<td>64</td>
<td>13.52</td>
<td>78.02</td>
<td>61</td>
<td>17.77</td>
<td>70.70</td>
<td>61</td>
<td>15.39</td>
<td>85.46</td>
</tr>
<tr>
<td>14</td>
<td>85</td>
<td>8.30</td>
<td>86.20</td>
<td>88</td>
<td>9.61</td>
<td>95.49</td>
<td>85</td>
<td>14.61</td>
<td>92.63</td>
</tr>
<tr>
<td>15</td>
<td>46</td>
<td>16.77</td>
<td>59.59</td>
<td>46</td>
<td>16.70</td>
<td>78.79</td>
<td>46</td>
<td>20.41</td>
<td>73.26</td>
</tr>
<tr>
<td>16</td>
<td>70</td>
<td>9.72</td>
<td>84.51</td>
<td>79</td>
<td>10.20</td>
<td>80.78</td>
<td>79</td>
<td>12.50</td>
<td>84.25</td>
</tr>
<tr>
<td>17</td>
<td>49</td>
<td>17.10</td>
<td>60.12</td>
<td>55</td>
<td>25.10</td>
<td>45.16</td>
<td>49</td>
<td>18.59</td>
<td>64.04</td>
</tr>
<tr>
<td>18</td>
<td>46</td>
<td>21.90</td>
<td>44.64</td>
<td>46</td>
<td>18.00</td>
<td>53.93</td>
<td>46</td>
<td>23.71</td>
<td>44.63</td>
</tr>
<tr>
<td>19</td>
<td>73</td>
<td>18.71</td>
<td>70.86</td>
<td>73</td>
<td>15.09</td>
<td>77.30</td>
<td>58</td>
<td>17.90</td>
<td>84.82</td>
</tr>
<tr>
<td>20</td>
<td>46</td>
<td>10.30</td>
<td>72.37</td>
<td>46</td>
<td>11.19</td>
<td>77.32</td>
<td>46</td>
<td>13.60</td>
<td>91.34</td>
</tr>
<tr>
<td>21</td>
<td>58</td>
<td>8.10</td>
<td>84.18</td>
<td>58</td>
<td>8.63</td>
<td>96.31</td>
<td>70</td>
<td>11.52</td>
<td>87.05</td>
</tr>
<tr>
<td>22</td>
<td>67</td>
<td>9.60</td>
<td>96.82</td>
<td>79</td>
<td>10.52</td>
<td>93.18</td>
<td>79</td>
<td>10.30</td>
<td>95.10</td>
</tr>
</tbody>
</table>

M 63.3 13.11 73.78 68.9 13.76 76.16 67.1 15.43 75.49

SD 15.9 3.89 15.29 18.3 4.77 18.86 17.5 4.41 18.71
**Figure 1** — This plot shows the associations between TGMD-2 gross motor quotient scores and the assessment duration of time in minutes as a function of the assessment protocol used. There was a significant relationship between these variables across all three protocols ($r = -.570$, $-.535$, and $-.798$, $r^2 = .325$, $286$, and $637$), respectively.

**Figure 2** — This plot shows the associations between TGMD-2 gross motor quotient scores and the percentage of time engaged in on-task behavior for children with ASD as a function of the assessment protocol used. There was a significant relationship between these variables across all three protocols ($r = .726$, $.575$, $.686$, $r^2 = .527$, $.331$, and $.471$), respectively.
the picture task card is an adaptation that may result in improved performance with no change in assessment duration. This finding could be beneficial to clinicians who must assess many children in a very short amount of time or are interested with the time to assessment completion.

The location of the picture activity schedule could explain the differences in assessment duration in minutes by protocol. The picture activity schedule was placed in the center of the assessment area. After completing each skill on the TGMD-2, the participant was prompted to “check schedule.” In order for the participant to check the picture activity schedule, the participant had to leave the area where he or she was performing the motor skills to walk approximately 15 m to the schedule and remove the top card on the schedule. The participant then returned back to the motor skills area to continue the assessment. During the traditional and the picture task card conditions, children watched physical demonstrations of a skill close to where they would begin performing the skill. The picture activity schedule placement was ideal for collecting data on videotape, but was perhaps another placement would have enabled the assessment duration to be more consistent across conditions.

This study also sought to examine the effectiveness of visual supports on the percentage of time on-task during the TMGD-2 assessment. The percentage of time on-task was examined in this study because children with ASD frequently have problems attending to relevant stimuli in instructional settings. The results of this study did not support the hypothesis that the percentage of time on-task during the TGMD-2 would be significantly higher for children with ASD when the assessment protocol included the picture activity schedule condition than the picture task card condition, which would also yield higher scores than the traditional protocol. There were no statistically significant differences in the percentage of time on-task between the three protocols.

A possible explanation as to why the percentage of time on-task was not significantly different between protocols was because this study examined the influence of visual supports on the completion of an assessment, rather than an instructional intervention. Previous studies utilizing visual supports resulting in significant differences in the percentage of time on-task were derived from intervention based studies in which baseline, intervention, and follow-up data were collected regarding the percentage of time engaged in on-task behavior (Bryan & Gast, 2000; Fittipaldi-Wert, 2007, MacDuff, Krantz, & McClannahan, 1993). Specifically, in these studies, data regarding on- and off-task behavior of elementary school aged children was collected using time sampling procedures during physical education, language arts education, and leisure time. In all three studies, baseline data were collected regarding behavior, then visual supports were introduced into the environment and intervention data were collected. In the current study, no baseline data examining the percentage of time engaged in on-task behavior by the participants was collected. Furthermore, previous studies used single subject research design, rather than the within-subjects group design used in this study. These differences may help to explain why visual supports do not impact time on-task during the TGMD-2.

Perhaps the nonsignificant findings related to the percentage of time on-task can be attributed to the measures used to attempt to control for environmental distractions. Beginning with the acclimation period, a routine was established where the participant followed the researcher’s instructions and performed a series of motor skills after the researcher demonstrated the skill. The order in which the motor skills were completed did not vary between the acclimation period and the assessment
protocols. The physical environment was not changed between assessments. Therefore, it is possible that the researchers’ attempts to control for environmental distractions could have influenced the findings related to the percentage of time on-task. Specifically, the window coverings and the lack of decorations on the walls in the assessment area could have eliminated the distractions present in previous studies examining time on-task in classroom settings. Finally, it should be noted that previous studies that found differences in time on-task when visual supports were used to communicate instructions to a child with ASD in settings with other children present (Bryan & Gast, 2000; Fittipaldi-Wert, 2007; MacDuff et al., 1993; Reid, Collier, & Cauchon, 1991). This study was an assessment conducted in a setting in which there was one child and three adult researchers present, a factor that may have influenced these results.

The final purpose of this study was to examine the relationships between the duration of time, time on-task, and the gross motor quotient scores on the TGMD-2 assessments. It was hypothesized that an inverse relationship would exist between the performance scores and the duration of assessments. TGMD-2 assessments would take less time to complete with children who scored higher on the TGMD-2 assessments than children with lower TGMD-2 scores. The data supported a moderate to strong inverse relationship between the length of time to complete the assessment and the gross motor quotient scores for the TGMD-2, depending on the assessment protocol used. Between 28%–32% of the variance in performance by children with ASD can be explained by the relationship between the duration of time and the gross motor quotient scores during the traditional protocol and the picture task card conditions, respectively. A strong relationship emerged during the picture activity schedule condition, as 63% of the variance in TGMD-2 gross motor quotient scores could be explained by the assessment duration. This finding supports the suggestion in the literature that a picture activity schedule serves as a powerful communicative tool for children with ASD (Bryan & Gast, 2000; Downing & Peckham-Hardin, 2000; MacDuff et al., 1993; O’Reilly et al., 2005).

It was also hypothesized that a positive relationship existed between gross motor quotient scores and the time engaged in on-task behaviors during the TGMD-2 assessment. Thus, students who obtained high TGMD-2 gross motor quotient scores spent more time appropriately engaged in the assessment than children with lower gross motor quotient scores. This relationship was found to exist regardless of protocol used, and this correlation explained between 33%–52% of the variance of performance by children with ASD. Given that time engaged in on-task behaviors in educational settings is one of the best predictors of long term success in performance settings (Hume & Reynolds, 2010), it is important to consider the implications of the relationships between assessment duration, time on-task, and performance scores on assessment settings. It is possible that interventions to reduce the assessment duration, or increase time on-task, could result in higher assessment scores. A causal study, rather than a correlational study, would be necessary to determine this type of relationship. Parents and clinicians interested in both the process and end result of assessment may be particularly interested in an intervention of this type on additional behavioral measures beyond time engaged in on-task behaviors.

It would be worthwhile to replicate the study in an environment where fewer measures were taken to eliminate distractions to determine if the percentage of time
on-task changes between the three protocols. It could be that conducting this study in a more “real world” environment could elicit different results related to percentage of time on-task. Furthermore, future studies should seek to explore causation between time on-task and gross motor quotient scores. Another follow-up study could determine if moving the picture activity schedule to a location closer to the assessment area than the location used in this study, elicits significantly higher gross motor quotient scores compared with the traditional protocol.

This is the first study documenting the relationship between time engaged in on-task behavior and performance on a motor skills assessment. The study sought to examine the effects of protocol modifications (i.e., picture activity schedule and picture task cards) on the performance of the TGMD-2 by children with ASD as measured by total duration of assessment time and the percentage of time engaged in on-task behavior. Furthermore, this study explored the relationship between time on-task and performance on the TGMD-2 as measured by gross motor quotient. Although the incorporation of visual supports had no effect on the percentage of time on-task, a relationship was found between time on-task and performance on the TGMD-2, regardless of protocol used. Children with ASD who had high percentage of on-task behavior, also seemed to have high gross motor quotient scores on the TGMD-2, especially when a picture activity schedule protocol was used. Clinicians and teachers interested in the scores yielded during assessments, as well as the process by which scores are earned during assessment, should note these findings, and acknowledge that using visual supports to improve motor skills on assessments will not extend assessment time.

Acknowledgments

This study was completed as part of the first author’s doctoral dissertation project. Thanks to Alice M. Buchanan, Mark G. Fischman, David Shannon, Robert G. Simpson, and Leah E. Robinson for serving on her doctoral committee, and Elizabeth Gilchrist, Katie Hayley, John C. Jordan, Alison Link, and Lauren Stark for assisting with data collection and analysis.

References


