

Catching Fish and Avoiding Sharks: Investigating Factors That Influence Developmentally Appropriate Measurement of Preschoolers' Inhibitory Control

Journal of Psychoeducational Assessment

2015, Vol. 33(6) 585–596

© 2014 SAGE Publications 

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0734282914562933

jpa.sagepub.com



Steven J. Howard¹ and Anthony D. Okely¹

Abstract

Although researchers agree that the first 5 years of life are critical for children's developing executive functions (EFs), further advances are hindered by a lack of consensus on the design and selection of developmentally appropriate EF tasks for young children. Given this debate, well-established adult measures of EF routinely have been adapted for young children. Given young children's comparatively limited cognitive capacities, however, such adaptations do not guarantee that the task's critical EF demands are retained. To investigate this possibility, the current study examined the characteristics that optimize measurement of young children's EFs—specifically, their inhibitory control—using the go/no-go (GNG) task as an exemplar. Sixty preschoolers completed six GNG tasks differing in stimulus animation, presentation time, and response location. Comparison EF tasks were administered to examine concurrent validity of GNG variants. Results indicated effects of stimulus presentation time and response location, with animation further enhancing task validity and reliability. This suggests that current GNG tasks deflate estimates of young children's ability to inhibit, with implications for future design and selection of developmentally appropriate EF tasks.

Keywords

inhibition, executive functions, preschool, go/no-go, early years toolbox, iPad

Introduction

Executive functions (EFs) provide the foundation for our ability to adapt in novel, misleading, or complex situations, in which learned responses are often suboptimal. In early childhood research, the term *executive functions* (EFs) is often used to refer to the capacity (working memory capacity) and control (inhibition, shifting) of attention. That is, EFs serve to activate, coordinate, and control information and processes that are within the focus of mental attention—the causal component underlying developmental growth of working memory (Pascual-Leone & Johnson, 2011). EFs also make important contributions to higher order cognitive processing, such as problem

¹University of Wollongong, New South Wales, Australia

Corresponding Author:

Steven J. Howard, School of Education, University of Wollongong, Wollongong, New South Wales 2522, Australia.

Email: stevenh@uow.edu.au

solving and self-regulation (Hofmann, Schmeichel, & Baddeley, 2012; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000), which may explain why the developmental influence of EFs appears to extend beyond the cognitive domain. Specifically, research suggests that proficient executive functioning is related to school readiness and academic achievement (Müller, Lieberman, Frye, & Zelazo, 2008), social and emotional understanding and competence (Riggs, Jahromi, Razza, Dillworth-Bart, & Müeller, 2006), and the deficient cognitive function often found in a range of developmental disorders (e.g., attention-deficit/hyperactivity disorder [ADHD]; Happe, Booth, Charlton, & Hughes, 2006). EFs in childhood thus appear to form a cognitive foundation that contributes to a broad array of subsequent developments.

Despite a general consensus that EFs develop gradually until late adolescence or early adulthood (Best, Miller, & Jones, 2009), how EFs are conceptualized (quantity, composition, and interpretation) and measured continues to vary widely. For instance, despite widespread use of “Tower” tasks (e.g., Tower of Hanoi, Tower of London) to measure children’s problem solving abilities, research suggests that these tasks may not tap the same underlying abilities, even in children of a similar age (Bull, Espy, & Senn, 2004). These issues in measurement are compounded by the different tasks and task specifications that researchers use to measure the same abilities across studies, which together have contributed to the inconsistent construal of “Tower” tasks as measuring inhibitory control (Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Miller, Giesbrecht, Müller, McInerney, & Kerns, 2012), planning (Hughes, Ensor, & Wilson, 2010), and shifting (Bull et al., 2004). In such cases, it is difficult to ascertain exactly how (and when) EFs develop and how they relate to children’s emerging competencies.

As a result of these measurement complexities, there remains little consensus regarding design, selection, and administration of developmentally appropriate EF tasks for preschool-aged children. Even EF measures that have been adapted from well-established adult versions of these tasks present complications for valid and reliable assessment in this age group. This is at least partially due to young children’s comparatively limited capacity and duration of attentional focus (Kannass, Oakes, & Shaddy, 2006), control of attention (e.g., increased susceptibility to distraction; Best et al., 2009; Howard, Johnson, & Pascual-Leone, 2014), ability to understand instructions and communicate their response (Hughes, 1998), and knowledge base (Chi, 1978). As a consequence, adapting an established task for young children does not ensure that its critical EF demands are retained (Anderson & Reidy, 2012; Garon, Bryson, & Smith, 2008). For this reason, Blair, Zelazo, and Greenberg (2005) conclude “work on the early development of EF has been limited by the lack of suitable measures for assessing specific aspects of EF in young children” (p. 561).

Issues in Developmentally Appropriate EF Measurement With Young Children: Case of the Go/No-Go (GNG) Task

The GNG task is a prime example of a task initially used to measure aspects of adults’ executive functioning—specifically, their inhibitory control—that has been adapted for use with young children (Dowsett & Livesey, 2000; Miller et al., 2012; Simpson & Riggs, 2006; Wiebe, Sheffield, & Espy, 2012). In this task, participants respond to a more frequent “go” stimulus and withhold responses to a less frequent “no-go” stimulus. The high frequency of go trials in the context of speeded performance pre-potentiates responding (Simpson & Riggs, 2006). As a result, this pre-potent tendency to respond must be inhibited for successful performance on no-go trials (the accuracy of which is often used to index inhibitory control). While maintaining these central task features, researchers have made a range of modifications to adult GNG tasks to enhance their appropriateness for young children. Specifically, these modifications often use more engaging stimuli, fewer trials, and slower stimulus presentation times to account for young children’s comparatively limited cognitive capacities.

Although it is not a given that EF task adaptations will retain the critical features of the original, multiple lines of evidence support the strength of these GNG variants for use with young children. For instance, studies indicate that the neural networks underlying children's GNG performance (Booth et al., 2003; Durston et al., 2002) broadly mirror that found with adults (albeit with children typically displaying greater degrees of activation; Garavan, Ross, Murphy, Roche, & Stein, 2002; Rubia et al., 2001). These modified GNG tasks also display good inter-task correlations with other established measures of inhibition and good reliability estimates (Simpson & Riggs, 2006; Wiebe et al., 2012). This has led some to conclude that the GNG task represents "a relatively pure measure of inhibitory control" (Simpson & Riggs, 2006, p. 19).

Despite these strengths, however, GNG tasks that have been modified for young children continue to vary widely in task specifications. That is, these GNG tasks differ in stimuli (e.g., fish/sharks, red/blue lights), stimulus timing (ranging from 1.5 s to 7 s), number of trials (from 20 to 100), ratio of go to no-go trials (ranging from 52% to 75% go trials), and method of delivery (e.g., physical apparatus, computer; Dowsett & Livesey, 2000; Miller et al., 2012; Simpson & Riggs, 2006; Wiebe et al., 2012). This is problematic insofar as each of these factors has been suggested to influence young children's GNG performance (Anderson & Reidy, 2012; Garon et al., 2008; Simpson & Riggs, 2006). For instance, even a 1-s change in GNG stimulus presentation time was found to render the task either too difficult (providing insufficient time to respond) or too easy (inadequately time pressured to generate a pre-potent response). Different levels of performance across even highly similar tasks thus may be a product of irrelevant features of the task, such as features or processes extraneous to the processes of interest, which may serve to obscure our understanding and sequencing of children's developing cognitive capacities.

Issues With Current EF Measurement Practices

Although previous studies provide insights into some of the task characteristics that more accurately and reliably index young children's inhibitory control, it is likely that these are not the only characteristics that affect young children's GNG performance. For instance, research indicates that reorienting attention adds to participants' response time, even if this attentional redirection does not involve saccadic movement (Hunt & Kingstone, 2003; Posner & Cohen, 1984). This is problematic in that common computerized GNG tasks present young children with stimuli in one spatial location (i.e., on screen) and require them to respond with a button press that, although well identified, is in another spatial location (e.g., keyboard). This results in the need to continually shift focus between stimulus and response locations, even when this does not require overt movement of the eyes. In such cases, it is unclear to what extent young children's correct no-go performance results from successfully withholding the pre-potent response or failing to inhibit the pre-potent response, yet this response not being considered because it occurred outside the allowable response interval.

In the context of learning and instructional design, there is also evidence that dynamic, animated stimuli can enhance attentional focus and task engagement (in contrast to the static GNG stimuli commonly used; Hongpaisanwiwat & Lewis, 2003). That this animation is most commonly achieved through the use of technology aligns well with suggestions that adapting EF tasks for tablet computers could increase precision with which EFs are measured in young children (Anderson & Reidy, 2012; Best & Miller, 2010). These advantages have been shown with adults, such as Brunetti, Del Gatto, and Delogu's (2014) adaptation of the Corsi block-tapping task for tablet computers, which provided advantages in administration, presentation, and scoring, while retaining comparable performance with the original physical version of the task. However, the extent to which these factors influence young children's EF performance remains unclear. Given the recent growth of interest in young children's EFs, including extensive and

expensive interventions to promote preschoolers' EF development, careful and thoughtful analyses of optimal and developmentally appropriate EF task design and selection is necessary.

The Current Study

The current study examined the task characteristics that optimize measurement of young children's inhibitory control, using the GNG task as an exemplar. Specifically, we examined the effects of integrating the stimulus and response locations (touchscreen vs. button press), stimulus dynamicity (static vs. animated), and stimulus presentation time (1,000, 1,500, 2,000 ms) on young children's GNG performance. GNG task variants were evaluated on the basis of (a) whether they generated a sufficiently pre-potentiated "go" response (indicated by significantly poorer performance on no-go trials compared with go trials), (b) concurrent validity (correlation between no-go and other EF task performance), and (c) reliability (split-half reliability) of the data generated. In line with previous research on optimal timings for GNG tasks with young children, as well as research suggesting the benefits of animation and integrating stimulus and response locations, it was expected that the animated, 1,500 ms iPad version of the GNG task would optimize measurement of young children's inhibitory control. Although this study focused specifically on effects of GNG task manipulations, the insights generated were expected to inform principles of EF task design and selection more broadly.

Material and Method

Participants

Participants were 60 children aged 3.05 to 5.69 years ($M = 4.59$, $SD = 0.59$), recruited from three Australian preschool centers managed by a non-profit organization. Forty-three percent of the participants were girls ($n = 26$), with a relatively even split of boys and girls across ages. All children were native speakers of English. Census data reveal that participants were drawn from a multicultural, middle-class urban area that is below the state average in family income and employment rate.

Measures

GNG. The GNG task is a measure of effortful response inhibition that has been used extensively with young children (Dowsett & Livesey, 2000; Miller et al., 2012; Müller, Kerns, & Konkin, 2012; Simpson & Riggs, 2006; Wiebe et al., 2012). The GNG task requires participants to respond to "go" trials ("catch fish") and withhold responding on "no-go" trials ("avoid sharks"). That the majority of stimuli were go trials (80% fish) generated a pre-potent tendency to respond, thereby requiring participants to inhibit this response for no-go trials (20% sharks). Prior to the task commencing, participants were given instruction and practice in the following sequence: go instructions, followed by 5 practice go trials; no-go instructions, followed by 5 practice no-go trials; combined GNG instructions, followed by a mixed block of 10 practice trials (80% go trials); and a recap of instructions prior to the task commencing. Feedback in the form of auditory tones was provided on all practice trials. The task proceeded with 75 test stimuli divided evenly into three test blocks (each separated by a short break and a reiteration of instructions). Stimuli were presented in pseudo-random order, such that a block never began with a no-go stimulus and no more than two successive trials were no-go stimuli, separated by a 1,000 ms interval between stimuli. Scores represent proportional accuracy on go and no-go trials.

The GNG task variants developed for this study were based on the protocols of Wiebe et al. (2012), except varying in integration of stimulus and response location (i.e., laptop button press,

iPad touchscreen), stimulus presentation time (i.e., 1,000, 1,500, 2,000 ms), and stimulus dynamicity (i.e., static, animated). To examine the effects of these factors on the measurement of young children's inhibitory control, six GNG variants were developed and evaluated. For comparison with a task typical of those commonly used with young children, (1) a "Standard Laptop" GNG task presented static stimuli for 1,500 ms each on a laptop computer (for studies using a computer-based GNG task with young children, see Miller et al., 2012; Müller et al., 2012; Simpson & Riggs, 2006). To evaluate the effect of integrating stimulus and response locations, performance on this task was compared against (2) a "Standard iPad" version of the task, differing only in that responses were made via iPad screen tap rather than laptop button press. To further evaluate the effect of stimulus timing, this newly developed iPad GNG task was modified to yield two additional variants differing only in stimulus presentation time—(3) "Fast iPad" (1,000 ms) and (4) "Slow iPad" (2,000 ms). To evaluate the effect of stimulus dynamicity, previously static stimuli were also animated (fish and sharks swam from right to left across the screen), yielding additional (5) "Standard iPad Animated" (1,500 ms) and (6) "Slow iPad Animated" (2,000 ms) variants. An animated version of the 1,000 ms iPad task was not created because, in accordance with previous results (Simpson & Riggs, 2006), it was expected that these stimuli would be too rapid to generate a pre-potent response in preschool-aged children. Similarly, we did not appraise all stimulus presentation times for the laptop task given previous research establishing these effects for computerized GNG tasks (Simpson & Riggs, 2006). The iPad-based GNG task used for these manipulations is freely available for download on the iTunes App Store under the name EYT Go/NoGo.

Peg-tapping task. The peg-tapping task (following the protocols of Diamond & Taylor, 1996) is a measure of effortful inhibition that has been psychometrically established as valid and reliable for use with young children (Smith-Donald, Raver, Hayes, & Richardson, 2007). In this task, participants are required to tap a small wooden dowel ("peg") one time whenever the tester taps twice, or tap the peg twice when the tester taps one time. Given that the pre-potent tendency for young children is to replicate the action they have witnessed, successful performance on this task requires that this pre-potent tendency be inhibited. Following the protocols of Diamond and Taylor (1996), the task provided instruction, demonstration, brief practice, and 14 test trials presented in the same random order to all participants. Scores were the number of correct trials.

Corsi blocks (backward). Corsi blocks is a measure of working memory capacity that has also been validated for use with young children (Pagulayan, Busch, Medina, Bartok, & Krikorian, 2006). In this task, participants were required to replicate, in reverse, a sequence of blocks tapped by the tester. Stimuli were nine blocks arranged in irregular order on a board. The task began with instruction and two practice trials with feedback (repeated up to 3 times in the case of an incorrect response). Two trials at each level of difficulty were then administered (ranging from two to six block sequences), until the earlier of completion or incorrect responses on both trials at the same level of difficulty. Although this is a measure of working memory capacity, this task was administered given evidence of the correlation among EFs (Miyake et al., 2000). Scores were the number of correct trials.

Mr. Ant. The Mr. Ant task, adapted from Case's (1985) Mr. Cucumber task, is a newly developed iPad-based measure of working memory capacity. Following established protocols of Morra (1994), participants were required to remember the spatial locations of "stickers" and identify these after a brief retention interval. Instruction and three practice trials served to familiarize participants with task requirements. Test trials increased in difficulty (i.e., working memory demand) as the task progressed, with three trials at each level of difficulty (progressing from one to eight stickers). All test trials progressed in the following order: (a) Mr. Ant presented with n differently colored stickers (where n equals the current level of difficulty) for 5 s;

(b) presentation of a blank screen for 4 s; then (c) an image of Mr. Ant without stickers, along with an auditory prompt to respond, presented until the participant's response was complete. Participants responded by tapping the parts of Mr. Ant that previously had stickers. The test continued until the earlier of completion or failure on all trials of the same level of difficulty. Scores were the number of correct trials. As above, this task was administered given evidence of the relationship among EFs (Miyake et al., 2000). This task is also freely available for download from the iTunes App Store under the name EYT Mr Ant.

Procedure

All children received the EF task battery across 3 days in a single week. To optimize levels of engagement, testing was separated across morning, midday, and afternoon sessions. Each testing session lasted approximately 10 min and was separated by a minimum 2-hr break. To minimize the influence of practice effects, tasks were administered in pseudo-counterbalanced order. Specifically, morning and afternoon sessions were used to administer different configurations of the GNG task. On Day 1, consecutive administrations of the 1,500 ms laptop and iPad GNG task were administered in counterbalanced order. On Days 2 and 3, all other GNG variants were administered in counterbalanced order. Midday testing sessions consisted of one of the comparison EF tasks, administered in counterbalanced order. All tasks were administered individually, in a quiet area of the child's preschool.

Results

Data Screening

To ensure only valid responses were included in analyses, GNG data were removed in the case of extremely rapid responses (trials with response times < 300 ms were removed because the response was unlikely to have been in response to the stimulus), indiscriminant responding (individual blocks were removed from analyses if go trial accuracy exceeded 80% and no-go trial accuracy fell below 20%), or non-responsiveness (individual blocks were removed if go trial accuracy fell below 20% and no-go trial accuracy exceeded 80%). This screening did not result in the complete removal of any participant's data. However, differing data loss across GNG tasks resulted: Standard Laptop (5.7%; 256 trials), Standard iPad (4.0%; 180 trials), Fast iPad (3.4%; 153 trials), Slow iPad (2.0%; 90 trials), Standard iPad Animated (1.9%; 86 trials), and Slow iPad Animated (1.4%; 62 trials). This provides initial support for the hypothesized measurement issues associated with some of the GNG variants, insofar as these tasks yielded a greater proportion of invalid data. Although Shapiro–Wilk statistics indicated that the resultant data did not meet assumptions of normality, this skewness was not extreme ($z_{\text{skewness}} < 4$).

Validity and Reliability Evidence

To further evaluate the convergent validity of each GNG variant, correlations between no-go performance and our comparison inhibition and working memory tasks were examined (Table 1). Results indicated that the Standard iPad Animated GNG task was most strongly and consistently related to the other EF tasks (r s ranging from .31 to .37). To also examine the reliability of these task variants, Spearman–Brown split-half reliability coefficients (Kaplan & Saccuzzo, 2013) were calculated for proportional no-go accuracy on each task. Reliability coefficients were as follows, arranged highest to lowest: Slow iPad ($r = .85$), Standard iPad Animated ($r = .84$), Slow iPad Animated ($r = .74$), Fast iPad ($r = .74$), Standard iPad ($r = .73$), and Standard Laptop ($r = .54$). Only the Standard iPad Animated and Slow iPad variants were supported by good reliability estimates.

Table 1. Correlations Between GNG and Comparison EF Tasks.

	1	2	3	4	5	6	7	8	9
1. GNG Standard Laptop	—	.35*	.36*	.38*	.30*	.39*	.13	.20	.18
2. GNG Standard iPad		—	.49*	.66*	.61*	.33*	.19	.36*	.35*
3. GNG Fast iPad			—	.54*	.23	.44*	.06	.18	.32*
4. GNG Slow iPad				—	.43*	.32*	.31*	.36*	.27
5. GNG Standard iPad Animated					—	.57*	.31*	.37*	.36*
6. GNG Slow iPad Animated						—	.27	.35*	.33*
7. Mr. Ant							—	.38*	.45*
8. Peg tapping								—	.31*
9. Corsi blocks									—

Note. Scores for GNG tasks were proportional no-go accuracy. Scores for Mr. Ant were number of correct trials. Scores for peg tapping were number of correct trials. Scores for Corsi blocks were number of correct trials. GNG = go/no-go; EF = executive function.

* $p < .05$.

Effects of Attentional Reorientation

To examine the effect of integrating stimulus and response locations, performance on the Standard Laptop and Standard iPad tasks were compared. It was expected that the integration of stimulus and response locations in the iPad version would minimize non-essential response time, thus facilitating responding on trials requiring a response and more accurately capturing incorrect responses on no-go trials. A 2 (Response Location) \times 2 (Trial Type) ANOVA on proportional accuracy for go and no-go trials indicated a main effect of Trial Type, $F(1, 44) = 13.84, p = .001, \eta^2 = .14$, such that performance on go trials was superior to performance on no-go trials. There was no main effect of Response Location, $F(1, 44) = 1.78, p = .189, \eta^2 = .01$. However, results were conditioned by a Response Location \times Trial Type interaction, $F(1, 44) = 7.07, p = .011, \eta^2 = .03$. As expected, post hoc analyses indicated that go performance was superior in the Standard iPad task ($M = 0.94, SD = 0.06$) relative to the Standard Laptop task ($M = 0.84, SD = 0.16$). No-go performance was non-significantly lower on the Standard iPad task ($M = 0.79, SD = 0.20$) relative to the Standard Laptop task ($M = 0.82, SD = 0.17$).

Latency analyses indicated that these differences may be a product of slower responding on the laptop task ($M = 0.93, SD = 0.16$) relative to the iPad task ($M = 0.88, SD = 0.11$), $t(46) = 2.80, p = .007, \eta^2 = .15$. Given that these two tasks were near identical—designed to differ only in their need for attentional reorientation to respond—this pattern of results suggests the Standard iPad's improved capture of go and no-go responses (i.e., fewer valid responses falling outside the allowable response interval). The need for attentional reorientation in the Standard Laptop variant, in contrast, may have resulted in some correct go trials and incorrect no-go trials not being captured. If so, this would suggest that go performance was deflated, and no-go performance inflated, in the Standard Laptop GNG variant.

Effects of Stimulus Presentation Time

To examine the effect of stimulus presentation time (1,000 ms; 1,500 ms; 2,000 ms), a 2 (Trial Type) \times 3 (Stimulus Duration) repeated-measures ANOVA was run on the Standard, Fast, and Slow iPad variants with static stimuli. Consistent with Wiebe et al. (2012), it was expected that a 1,500 ms presentation time would generate a sufficiently pre-potent response without yielding performance at ceiling or floor. ANOVA results indicated main effects of Stimulus Duration, $F(2, 76) = 4.21, p = .018, \eta^2 = .02$, and Trial Type, $F(1, 38) = 19.55, p < .001, \eta^2 = .21$. Post hoc

analyses indicated greater no-go accuracy at 2,000 ms ($M = 0.89$, $SD = 0.10$) compared with the 1,000 ms ($M = 0.84$, $SD = 0.10$) or 1,500 ms presentation times ($M = 0.86$, $SD = 0.10$), which did not significantly differ. This was conditioned by a significant interaction, $F(2, 76) = 6.04$, $p = .004$, $\eta^2 = .03$. This can be understood as follows: (a) a significant difference in go trial accuracy across all Stimulus Duration conditions (1,000 ms: $M = 0.86$, $SD = 0.11$; 1,500 ms: $M = 0.94$, $SD = 0.06$; 2,000 ms: $M = 0.97$, $SD = 0.05$), (b) no significant differences in no-go accuracy among any of the Stimulus Duration conditions (1,000 ms: $M = 0.82$, $SD = 0.18$; 1,500 ms: $M = 0.79$, $SD = 0.20$; 2,000 ms: $M = 0.80$, $SD = 0.21$), and (c) no significant difference between go and no-go performance in the 1,000 ms task, but significant differences in the 1,500 ms and 2,000 ms conditions. Results thus suggest that the Fast iPad GNG task did not discriminate between go and no-go performance.

Effects of Stimulus Dynamicity

Last, 2 (Trial Type) $\times 2$ (Animation) repeated-measures ANOVAs were run separately for the Standard and Slow iPad tasks with and without animated stimuli, to evaluate the effects of stimulus dynamicity on performance. It was expected that animation would enhance young children's ability to sustain their on-task performance, thus creating a stronger pre-potent tendency toward responding. As such, reduced accuracy on the animated tasks was expected. In contrast to expectations, results for the 1,500 ms tasks indicated a non-significant effect of Animation in either GNG variant: 1,500 ms, $F(1, 40) = 0.75$, $p = .391$, $\eta^2 < .01$; 2,000 ms, $F(1, 44) = 0.02$, $p = .901$, $\eta^2 < .01$. There was a significant interaction for the 2,000 ms variant: 1,500 ms, $F(1, 40) = 0.39$, $p = .534$, $\eta^2 < .01$; 2,000 ms, $F(1, 44) = 8.39$, $p = .006$, $\eta^2 = .04$. Post hoc analyses indicated this was a product of superior go performance in the non-animated condition ($M = 0.97$, $SD = 0.05$) compared with the animated condition ($M = 0.92$, $SD = 0.10$), $t(44) = -3.75$, $p = .001$, $\eta^2 = .24$. There was no significant difference in no-go performance, $t(44) = 1.63$, $p = .111$, $\eta^2 = .06$.

Discussion

The current study provides novel and converging evidence that a diverse array of factors, beyond those previously considered in EF research, influence developmentally appropriate measurement of young children's inhibitory control. Specifically, using a repeated-measures experimental design, we administered multiple GNG task variants to investigate the influence of response method, stimulus dynamicity, and timing on young children's GNG performance. Results indicated significant effects of response method and stimulus presentation time, with animation further enhancing validity and reliability. These results highlight the problems associated with many current practices in preschool EF research, including adaptation of established adult EF measures that do not adequately cater for the unique complexities of assessing the cognitive development of young children.

Whereas contemporary studies of young children's EFs have frequently used computer-based "button press" tasks to assess inhibitory control (Miller et al., 2012; Müller et al., 2012; Simpson & Riggs, 2006; Wiebe et al., 2012), our results suggest that this mode of responding obscured accurate measurement of preschoolers' inhibitory control. Specifically, comparing a typical computer-based GNG task (following Wiebe et al., 2012) with an otherwise identical iPad version of this task demonstrated that go trial performance was significantly better and no-go performance non-significantly worse on the iPad task. In fact, only in the iPad version did go and no-go performance significantly differ. This finding is consistent with suggestions that the reorientation of attention (i.e., from on-screen GNG stimuli to the off-screen response button) requires both time and effort, even in the absence of an overt redirection of the eyes (Hunt & Kingstone, 2003;

Posner & Cohen, 1984). In this context, the integration of stimulus and response locations in the iPad task can be interpreted as fostering more accurate capture of correct go trials and incorrect no-go trials (a suggestion that is supported by our latency data). The relative superiority of the Standard iPad task was also evidenced by its improved reliability estimate (.73) compared with the computer-based version of the task (.54).

Stimulus timing was also found to have an effect on young children's GNG performance. Our stimulus timing results paralleled previous findings (Simpson & Riggs, 2006), such that overly rapid stimulus presentation (1,000 ms) was associated with lower rates of responding and an inability to discriminate between go and no-go performance. In contrast, Standard and Slow stimulus presentation times displayed good levels of discrimination. Animation did not have a similar effect on performance, although it did provide better reliability and stronger correlations with other EF measures (especially in the case of the Standard iPad Animated variant). Although the Slow iPad variant showed slightly better reliability (.85), it was not as highly correlated with other EF measures (ranging from .27 to .36). Thus, the Standard iPad Animated uniquely provided a sufficiently pre-potentiated response, good differentiation of go and no-go performance, consistently significant correlations with other EF tasks, and good reliability estimates. Although this finding supports Wiebe et al.'s (2012) assertion that the 1,500 ms stimulus presentation time would extend the utility of the task to a broader array of ages, the current results further suggest that the developmentally appropriate measurement of preschoolers' inhibitory control can be additionally enhanced through use of dynamic stimuli.

It is notable that inter-task EF correlations in the current study were only modest, ranging from .31 to .37 with the Standard iPad Animated GNG task. A possible interpretation of these correlations is that these tasks may fail to tap a common source of EF variance (Dempster, 1992; Nigg, 2000). Alternatively, the EF task correlations may reflect modest but meaningful relationships (Howard et al., 2014)—a proposal that is supported by the consistency of these correlations with studies that successfully extract EF latent variables (Friedman & Miyake, 2004; Miyake et al., 2000). Nevertheless, these correlations highlight the need for studies such as this, which aim to optimize the measurement of young children's EFs.

In support of this aim, although the current study was limited in its focus on the GNG paradigm, it is expected that the insights generated from this study can inform EF research more broadly. To illustrate, GNG tasks are not the only measures of inhibitory control that require a button press to respond (e.g., Continuous Performance Test, Simon Task; Martin-Rhee & Bialystok, 2008; Wiebe, Espy, & Charak, 2008). Many non-inhibitory EF measures also require response to static stimuli via button press (Stins et al., 2005; Tsujimoto, Kuwajima, & Sawaguchi, 2007), with unclear effects on the accuracy of the EF measurements they yield. Which, and to what extent, current EF tasks accurately and sensitively index young children's executive functioning thus remains unclear.

In contrast to some of the current practices in preschool EF research, our results advocate the adoption of less-common characteristics in the design and selection of EF tasks for young children. These include (a) minimizing irrelevant features in an EF task, which can introduce extraneous processes or processing (e.g., reorienting attention from stimulus to response); (b) optimizing domain-specific and domain-general demands of the task (e.g., the intuitiveness of catching fish and avoiding sharks reducing the working memory demands of remembering response rules compared with, for example, red and blue lights); (c) balancing empirical needs (to enhance reliability of generated data) with behavioral considerations (e.g., duration of young children's attentional focus, without which validity is threatened); and (d) generating a sufficiently pre-potent response, with responding sufficiently speeded so as to ensure that performance is not at ceiling or floor. Although further research is required to substantiate these suggestions across a range of EF tasks and contexts, our results nevertheless highlight the need for careful consideration in the design and selection of EF tasks for young children.

Acknowledgments

We are grateful to the participating preschools, the children who participated in the project, and to Jessica Tougher, Jacqueline Peperkamp, Kellie Mahar, Justine Harman, and Tamara Raso for help with data collection.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by Internal Funding from the University of Wollongong.

References

- Anderson, P. J., & Reidy, N. (2012). Assessing executive function in preschoolers. *Neuropsychology Review, 22*, 345-360.
- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*, 1641-1660.
- Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review, 29*, 180-200.
- Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology, 28*, 561-571.
- Booth, J. R., Burman, D. D., Meyer, J. R., Lei, Z., Trommer, B. L., Davenport, N. D., . . . Mesulam, M. M. (2003). Neural development of selected attention and response inhibition. *NeuroImage, 20*, 737-751.
- Brunetti, R., Del Gatto, C., & Delogu, F. (2014). eCorsi: Implementation and testing of the Corsi block-tapping task for digital tablets. *Frontiers in Psychology, 5*, 1-9.
- Bull, R., Espy, K. A., & Senn, T. E. (2004). A comparison of performance on the Towers of London and Hanoi in young children. *Journal of Child Psychology and Psychiatry, 45*, 743-754.
- Case, R. (1985). *Intellectual development: Birth to adulthood*. New York, NY: Academic Press.
- Chi, M. T. H. (1978). Knowledge students and memory development. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73-96). Hillsdale, NJ: Lawrence Erlbaum.
- Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12*, 45-75.
- Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "do as I say, not as I do." *Developmental Psychobiology, 29*, 315-334.
- Dowsett, S. M., & Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of "executive skills" training. *Developmental Psychobiology, 36*, 161-174.
- Durstun, S., Thomas, K. M., Yang, Y., Ulug, A. M., Zimmerman, R. D., & Casey, B. J. (2002). A neural basis for the development of inhibitory control. *Developmental Science, 5*, F9-F16.
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133*, 101-135.
- Garavan, H., Ross, T. J., Murphy, K., Roche, R. A. P., & Stein, E. A. (2002). Dissociable executive functions in the dynamic control of behavior: Inhibition, error detection, and correction. *NeuroImage, 17*, 1820-1829.
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*, 31-60.
- Happe, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition, 61*, 25-39.
- Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences, 16*, 174-180.

- Hongpaisanwiwat, C., & Lewis, M. (2003, October). The effects of animated character in multimedia presentation: Attention and comprehension. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, 2, 1350-1352.
- Howard, S. J., Johnson, J., & Pascual-Leone, J. (2014). Clarifying inhibitory control: Diversity and development of attentional inhibition. *Cognitive Development*, 31, 1-21.
- Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, 16, 233-253.
- Hughes, C., Ensor, R., & Wilson, A. (2010). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology*, 35, 20-36.
- Hunt, A. R., & Kingstone, A. (2003). Covert and overt voluntary attention: Linked or independent? *Cognitive Brain Research*, 18, 102-105.
- Kannass, K., Oakes, L., & Shaddy, J. (2006). A longitudinal investigation of the development of attention and distractibility. *Journal of Cognition and Development*, 7, 381-409.
- Kaplan, R. M., & Saccuzzo, D. P. (2013). *Psychological testing: Principles, applications, and issues* (8th ed.). Blemont, CA: Brooks/Cole.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21, 59-80.
- Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 11, 81-93.
- Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, 13, 395-423.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and the contributions to complete "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49-100.
- Morra, S. (1994). Issues in working memory measurement: Testing for M capacity. *International Journal of Behavioral Development*, 17, 143-159.
- Müller, U., Kerns, K. A., & Konkin, K. (2012). Test-retest reliability and practice effects of executive function tasks in preschool children. *The Clinical Neuropsychologist*, 26, 271-287. doi:10.1080/13854046.2011.645558
- Müller, U., Lieberman, D., Frye, D., & Zelazo, P. D. (2008). Executive function, school readiness, and school achievement. In S. K. Thurman & C. A. Fiorello (Eds.), *Applied cognitive research in K-3 classrooms* (pp. 41-84). New York, NY: Routledge.
- Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working memory taxonomy. *Psychological Bulletin*, 126, 220-246.
- Pagulayan, K. F., Busch, R. M., Medina, K. L., Bartok, J. A., & Krikorian, R. (2006). Developmental normative data for the Corsi block-tapping task. *Journal of Clinical and Experimental Neuropsychology*, 28, 1043-1052.
- Pascual-Leone, J., & Johnson, J. (2011). A developmental theory of mental attention: Its application to measurement and task analysis. In P. Barrouillet & V. Gaillard (Eds.), *Cognitive development and working memory: A dialogue between neo-Piagetian theories and cognitive approaches* (pp. 13-46). New York, NY: Psychology Press.
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In D. Bouma & D. Bonwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531-556). Hillsdale, NJ: Lawrence Erlbaum.
- Riggs, N. R., Jahromi, L. B., Razza, R. P., Dillworth-Bart, J. E., & Müeller, U. (2006). Executive function and the promotion of social-emotional competence. *Journal of Applied Developmental Psychology*, 27, 300-309.
- Rubia, K., Russell, T., Overmeyer, S., Brammer, M. J., Bullmore, E. T., Sharma, T., . . . Taylor, E. (2001). Mapping motor inhibition: Conjunctive brain activations across different versions of go/no-go and stop tasks. *NeuroImage*, 13, 250-261.
- Simpson, A., & Riggs, K. J. (2006). Conditions under which children experience inhibitory difficulty with a "button press" go/no-go task. *Journal of Experimental Child Psychology*, 94, 18-26.

- Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PRSA) for field-based research. *Early Childhood Research Quarterly, 22*, 173-187.
- Stins, J. F., de Sonneville, L. M. J., Groot, A. S., Polderman, T. C., van Baal, C. G. C. M., & Boomsma, D. I. (2005). Heritability of selective attention and working memory in preschoolers. *Behavior Genetics, 35*, 407-416.
- Tsujimoto, S., Kuwajima, M., & Sawaguchi, T. (2007). Developmental fractionation of working memory and response inhibition during childhood. *Experimental Psychology, 54*, 30-37.
- Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: Sources of variation in emergent mathematic achievement. *Developmental Psychology, 44*, 575-587.
- Wiebe, S. A., Sheffield, T. D., & Espy, K. A. (2012). Separating the fish from the sharks: A longitudinal study of preschool response inhibition. *Child Development, 83*, 1245-1261.