INTRODUCTION

Inhibitory control is the ability to suppress competing, dominant, automatic, or prepotent cognitive processing at perceptual, intermediate, and output stages. Inhibitory control is a key cognitive function of typical and atypical child development. This study examined age-related trends of Stroop-like interference in 3 to 12-year-old children and young adults by administration of a computerized Stroop-like big–small task with reduced working memory demand. This task used a set of pictures displaying a big and small circle in black and included the same condition and the opposite condition. In the same condition, each participant was instructed to say “big” when viewing the big circle and to say “small” when viewing the small circle. In the opposite condition, each participant was instructed to say “small” when viewing the big circle and to say “big” when viewing the small circle. The opposite condition required participants to inhibit the prepotent response of saying the same, a familiar response to a perceptual stimulus. The results of this study showed that Stroop-like interference decreased markedly in children in terms of error rates and correct response time. There was no deterioration of performance occurring between the early trials and the late trials in the sessions of the day–night task. Moreover, pretest failure rate was relatively low in this study. The Stroop-like big–small task is a useful tool to assess the development of inhibitory control in young children in that the task is easy to understand and has small working memory demand.

Keywords: inhibition, executive function, cognitive control, day–night task, Stroop
potency of the incorrect response is magnified on each trial because of its activation during testing (by virtue of its inclusion in the response set) coupled with its depiction throughout the testing. In other words, the structure of the day–night task elevates response competition because (a) the response alternative that must be suppressed is depicted on the test card (e.g., “night” for the night card) and (b) the incorrect response alternative was previously activated on previous trials (e.g., “night” for day cards) in the case of a correct response, i.e., a response set effect (Simpson and Riggs, 2005a; Montgomery et al., 2008; Simpson et al., 2012).

Working memory is also presumed to be an important factor related to the performance in the day–night paradigm. Working memory may be involved because resolving which of the conflicting responses must be suppressed entails holding the task rules in mind (“say ‘day’ for night card” and “say ‘night’ for day card”). In fact, some studies report deterioration in performance occurring between the first four trials and the last four trials in the sessions of the day–night task in young children (e.g., Gerstadt et al., 1994). These reports suggest that young children may have forgotten rules or that working memory demands add to the processing requirements of the task and consequently compromise inhibitory mechanisms (Montgomery and Koeltzow, 2010). Then, it is expected to reduce the working memory demands in the day–night paradigm so that the task primarily measures inhibitory control.

Working memory demands may also be related to learning the task rules. It might be true that the day–night task recruits working memory to a certain extent because learning the combination between words and pictures (“day” for a night picture of the moon and stars and “night” for a day picture of the sun) is not easy to understand for young children. Actually, previous studies with the day–night task had a great pretest failure rate, especially in children aged between 3 and 4 (e.g., Gerstadt et al., 1994). A problem is that with more children failing the pretest, the sample of children whose RT and accuracy data are analyzed is not representative of the population. Probably, the children with the poorest inhibitory control get excluded.

This study was conducted to examine age-related trends of inhibitory control in various age groups by administration of a Stroop-like day–night variant with reduced working memory demand. For this study, 3 to 12-year-old children and young adults were administered a Stroop-like big–small task. The task used a set of pictures displaying a big or small circle in black and required participants to produce sized-based responses following instructions given by the experimenter. The size labels “big” and “small” were used because they are well understood even by very young children and because they are distinctive and opposite, both of which may facilitate learning and holding the task rules. The task has two conditions: the opposite condition, in which a participant says the opposite of what is shown with card pairs, and the same task condition, in which a participant simply names what the stimulus represents. Because the original study (Gerstadt et al., 1994) used a different combination of words and pictures in the opposite condition (saying the opposite of what is shown for day–night cards) to that used in the control condition (saying “day” or “night” to abstract shapes), the degree to which a picture evokes a particular response was not controlled. Comparison between opposite and same conditions, as in this study, indicates the inhibitory processes, controlling for a difference in the degree to which a picture evokes a particular response. In this study, unlike the standard “card” version of the day–night task, the task was computerized to evaluate the correct RT more precisely.

**MATERIALS AND METHODS**

**PARTICIPANTS**

Participants were 113 typically developing people who were divided into six age groups: (a) 3–4 year, 20 children (10 boys, 10 girls; M age = 51.5 months, age range = 43–59); (b) 5–6 year, 14 children (7 boys, 7 girls; M age = 68.6 months, age range = 60–83); (c) 7–8 year, 20 children (11 boys, 9 girls; M age = 95.1 months, age range = 87–107); (d) 9–10 year, 19 children (9 boys, 10 girls; M age = 119.6 months, age range = 108–131); (e) 11–12 year, 17 children (9 boys, 8 girls; M age = 144.4 months, age range = 133–153); and (f) 23 young adults (9 men, 14 women; M age = 21.1 year, age range = 18–24). Children were recruited through local mainstream preschool and elementary school programs. Young adults were recruited from a university. All participants speak Japanese as a first language. Criteria for inclusion were the absence of bilingualism and absence of behavioral or educational problems, which would affect the study of inhibitory control.

**MATERIALS AND METHODS**

The Stroop-like big–small task used a set of pictures displaying either a big (12 cm diameter) or a small (1 cm diameter) circle in black. The same set of pictures was used in the same and opposite conditions. SuperLab (Cedrus Corp., San Pedro, CA, USA) controlled the task, presenting stimuli and recording participants’ vocal responses (error and time).

**PROCEDURES**

Participants were tested individually in quiet rooms at their respective schools. At arrival, a participant was asked to be seated next to the experimenter at the table and approximately 50 cm in front of a monitor with a headset microphone. Subsequently, the experimenter explained that they were going to play a “game” in which they would see two pictures. The experimenter showed the participant the big and small circles at the same time on the screen and asked him or her to point to the big circle and the small circle in turn. All participants were able to do this. Then, each was administered the Stroop-like big–small task. In this task, the same condition was arranged to precede the opposite condition in an attempt to elicit robust interference.

Prior to the test phase, participants were trained on how to play each “game.” For the same condition, the experimenter said, “Here is a picture of a big circle (show a big circle on the monitor). When this picture is shown, I want you to say ‘big.’ And, here is a picture of a small circle (show a small circle on the monitor). When this picture is shown, I want you to say ‘small.’” The participant did four practice trials (big–small–small–big). If the participant made any error, then the participant was corrected, reminded of the rules, and administered another four practice trials. The task did not commence until the participant was 100% correct for a set. For the opposite condition,
the experimenter said, “We are going to play an opposite game. Here is a picture of big circle (show a big circle on the monitor). When this picture is shown, I want you to say ‘small’. And, here is a picture of small circle (show a small circle on the monitor). When this picture is shown, I want you to say ‘big’.” The opposite condition required participants to inhibit the prepotent response of saying the same, a familiar response to a perceptual stimulus. The practice trials were identical to the same condition.

During the test phase, the participant was asked to respond as quickly and accurately as possible to a series of 20 stimuli (10 big circles and 10 small circles) for each task condition. All stimuli were presented one at a time and randomly at the center of the white screen on the monitor. At the instant a participant’s voice key was input, each stimulus was replaced by a fixation cross until the participant was judged by the experimenter to be ready to proceed to the next trial, looking at the fixation cross. The interstimulus interval was not controlled by SuperLab, as it was in a card version of the task, because some younger children have difficulty engaging in the task continuously. No feedback reminding participants of the task rules was given during testing.

**ANALYSIS**

Numbers of errors and RT for correct responses were recorded. Trials were counted as incorrect when participants’ initial responses were errors, even if they self-corrected. The RT was measured as the interval in milliseconds between the presentation of a stimulus and the onset of the participant’s vocal response by the microphone. Analysis of RT was conducted only for the correct response. Mean and standard deviations of error rates and correct RT on the whole trials were calculated for each task condition. To examine changes of performance over the course of a session for each task condition, mean and standard deviations of error rates and correct RT were calculated for the first five trials and the last five trials, respectively.

The data were analyzed using analysis of variance (ANOVA). Specifically, two-way ANOVAs with the within-participant factor of condition (same and opposite) and between-participant factor of age group (3 to 4-year olds, 5 to 6-year olds, 7 to 8-year olds, 9 to 10-year olds, 11 to 12-year olds, and young adults) were conducted for error rates and correct RT. Also, three-way ANOVAs with the within-participant factors of condition (same and opposite) and serial position (first five trials and last five trials) and between-participant factor of age group (3 to 4-year olds, 5 to 6-year olds, 7 to 8-year olds, 9 to 10-year olds, 11 to 12-year olds, and young adults) were conducted for error rates and correct RT. Software was used for statistical analyses (SPSS 19.0 for Windows; SPSS Japan Inc., Tokyo, Japan). Unless otherwise noted, a 0.05 level of significance was adopted for all statistical analyses.

**ETHICAL APPROVAL**

Informed consent was obtained from all adult participants and from a parent of each child participant before the assessment session. Our experimental protocol was administered in accordance with the guidelines of the Declaration of Helsinki and was approved by the institutional review board.

**RESULTS**

**TREATMENT OF UNUSED DATA**

An additional 14 participants were tested. Data from 9 participants were not included in this study because they showed results more than 3 SD from the mean of each age group (i.e., outliers). One 3-year-old child was not able to pass a pretest for the saying-opposite condition. Another 3-year-old child was not able to complete the task because his voice was too small to record. Three school-age children were excluded because of experimental error in recording the data.

**ACCURACY OF RESPONSE**

Figure 1 depicts the mean and standard deviations for the error rates in the Stroop-like big–small task by age group and condition. A 6 (age group) × 2 (condition) mixed-model ANOVA was conducted of the error rates. The analysis showed significant main effects for the age group ($F_{5,107} = 18.40, p < 0.001$, partial $\eta^2 = 0.46$), for condition ($F_{1,107} = 26.73, p < 0.001$, partial $\eta^2 = 0.20$), and for interaction of age group and condition ($F_{5,107} = 4.02, p < 0.01$, partial $\eta^2 = 0.16$). Post hoc Bonferroni tests yielded significant differences between-age group comparisons between 3 to 4-year olds and other age groups and between 5 to 6-year olds and other age groups for each condition ($p < 0.05$). For each condition, between-age group comparisons among 7 to 8-year olds, 9 to 10-year olds, and 11 to 12-year olds and young adults were not significant. Post hoc Bonferroni tests also yielded significant differences between conditions for 3 to 4-year olds and 5 to 6-year olds ($p < 0.05$), but not for other age groups (7 to 8-year olds, $p = 0.084$; 9 to 10-year olds, $p = 0.107$; 11 to 12-year olds, $p = 0.588$; young adults, $p = 1.00$). These results clarified that the interaction between age group and condition reflected age-related convergence of error rates in the same condition and in the opposite condition.

Figure 2 depicts the mean and standard deviations for the error rates early (the first five trials) and late (the last five trials) in the session in the Stroop-like big–small task by age group and condition. A 6 (age group) × 2 (condition) × 2 (serial position) mixed-model ANOVA was conducted of the error rates. The analysis showed significant main effects for the age group ($F_{5,107} = 18.33, p < 0.001$, $\eta^2 = 0.46$), for condition ($F_{1,107} = 26.73, p < 0.001$, partial $\eta^2 = 0.20$), and for interaction of age group and condition ($F_{5,107} = 4.02, p < 0.01$, partial $\eta^2 = 0.16$). Post hoc Bonferroni tests yielded significant differences between-age group comparisons between 3 to 4-year olds and other age groups and between 5 to 6-year olds and other age groups for each condition ($p < 0.05$). For each condition, between-age group comparisons among 7 to 8-year olds, 9 to 10-year olds, and 11 to 12-year olds and young adults were not significant. Post hoc Bonferroni tests also yielded significant differences between conditions for 3 to 4-year olds and 5 to 6-year olds ($p < 0.05$), but not for other age groups (7 to 8-year olds, $p = 0.084$; 9 to 10-year olds, $p = 0.107$; 11 to 12-year olds, $p = 0.588$; young adults, $p = 1.00$). These results clarified that the interaction between age group and condition reflected age-related convergence of error rates in the same condition and in the opposite condition.

![Figure 1](image-url) | Mean error rates in the Stroop-like big–small. Error bars represent standard deviations. No adults showed an error.
partial $\eta^2 = 0.46$), for condition ($F_{1,107} = 15.97, p < 0.001$, partial $\eta^2 = 0.13$), and for interaction of age group and condition ($F_{5,107} = 2.84, p < 0.05$, partial $\eta^2 = 0.12$). Main effect for the serial position, the other two-way interactions, and the three-way interaction were not significant. These results clarified that there were no deterioration of performance over the course of a session in terms of the error rates.

CORRECT RESPONSE TIME

Figure 3 depicts the mean and standard deviations for the correct RTs in the Stroop-like big–small task by age group and condition. A 6 (age group) × 2 (condition) mixed-model ANOVA was conducted of the correct RT. The analysis showed significant main effects for the age group ($F_{5,107} = 46.25, p < 0.001$, partial $\eta^2 = 0.68$), for condition ($F_{1,107} = 153.51, p < 0.001$, partial $\eta^2 = 0.59$), and for interaction of age group and condition ($F_{5,107} = 25.24, p < 0.001$, partial $\eta^2 = 0.54$). Main effect for the serial position, the other two-way interactions, and the three-way interaction were not significant. These results clarified that there were no deterioration of performance over the course of a session in terms of the correct RTs.

DISCUSSION

This study examined age-related trends of Stroop-like interference in 3 to 12-year-old children and young adults by administration of a computerized Stroop-like big–small task. In this study, the differences between the opposite and same conditions were compared among age groups for error rates and correct RT. It was hypothesized that working memory demand is reduced in the Stroop-like big–small task.

Results show that Stroop-like interference decreased markedly in children. The difference between conditions in error rates was significant for 3 to 4-year olds and 5 to 6-year olds but not for the older age groups although there were trends toward significance for some older age groups, which may be due to relatively small sample size. These results are consistent with the results obtained from previous studies using the day–night task and other variants of this task (Gerstadt et al., 1994; Simpson and Riggs, 2005b). However, this difference in correct RT was significant for 3 to 4-year olds, 5 to 6-year olds, 7 to 8-year olds, and 9 to 10-year olds. This result is consistent with the results reported by Simpson and

![Image of Figure 2: Mean error rates early and late in the session in the Stroop-like big–small.](image)

![Image of Figure 3: Mean RT in the Stroop-like big–small.](image)
In conclusion, this study demonstrated that the difference between naming what stimuli represent and naming of the “opposite” of the stimuli was decreased significantly during young childhood in the Stroop-like big-small task that has smaller working memory demands than the original version of the day–night task. In other words, this study showed that inhibitory control develops rapidly in young children. The Stroop-like big–small task is a useful tool to investigate the development of inhibitory control in young children in that the task is easy to understand and has small working memory demand.

ACKNOWLEDGMENTS
The authors thank all who participated in the study. This research was supported by the Japan Society for the Promotion of Science Research Fellowship for Young Scientists (to Yoshifumi Ikeda).

REFERENCES
Anderson, M. (2001). Annotation: conceptions of intelligence. J. Child Psychol. Psychiatry 42, 287–298. doi: 10.1111/1469-7610.00722
Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. Child Neuropsychol. 8, 71–82. doi: 10.1076/chin.8.2.71.8724
Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychol. Bull. 121, 65–94. doi: 10.1037/0033-2909.121.1.65
Diamond, A., Kirkham, N., and Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. Dev. Psychol. 38, 352–362. doi: 10.1037/0012-1649.38.3.352
Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. J. Exp. Psychol. Gen. 133, 101–135. doi: 10.1037/0096-7434.133.1.101
Gerstadt, C. L., Hong, Y. J., and Diamond, A. (1994). The relationship between cognition and action: performance of children 3 1/2–7 years old on a Stroop-like day–night test. Cognition 53, 129–153. doi: 10.1016/0010-0277(94)90068-X
Harnishfeger, K. K., and Bjorklund, D. F. (1994). A developmental perspective on individual differences in inhibition. Learn. Individ. Differ. 6, 331–355. doi: 10.1016/1041-6080(94)90201-3
Ikeda, Y., Okuzumi, H., and Kokubun, M. (2013a). Age-related trends of Stroop-like interference in animal size tests in 5–12-year-old children and young adults. Child Neuropsychol. 19, 276–291. doi: 10.1080/09297049.2012.658364
Ikeda, Y., Okuzumi, H., and Kokubun, M. (2013b). Stroop/reverse-Stroop interference in typical development and its relation to symptoms of ADHD. Res. Dev. Disabil. 34, 2391–2398. doi: 10.1016/j.ridd.2013.04.019
Ikeda, Y., Okuzumi, H., Kokubun, M., and Haishi, K. (2011). Age-related trends of interference control in school-age children and young adults in the Stroop color-word test. Psychol. Rep. 108, 577–584. doi: 10.2466/04.10.22.PR0.108.2.577-584
Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. Curr. Dir. Psychol. Sci. 21, 8–14. doi: 10.1177/0963721411429458
Montgomery, D. E., Anderson, M., and Uhrl, E. (2008). Interference control in preschoolers: factors influencing performance on the day–night task. Infant Child Dev. 17, 457–470. doi: 10.1002/icd.559
Montgomery, D. E., and Koeltzow, T. E. (2010). A review of the day–night task: the Stroop paradigm and interference control in young children. Dev. Rev. 30, 398–393. doi: 10.1016/j.dr.2010.07.001
Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. Psychol. Bull. 126, 220–246. doi: 10.1037/0033-2909.126.2.220

FIGURE 4 | Mean RT early and late in the session in the Stroop-like big–small. (A) mean RT on the first five trials, (B) mean RT on the last five trials. Error bars represent standard deviations.
Ozonoff, S., and Jensen, J. (1999). Brief report: specific executive function profiles in three neurodevelopmental disorders. *J. Autism Dev.* 29, 171–177. doi: 10.1023/A:1023052913110

Simpson, A., and Riggs, K. J. (2005a). Factors responsible for performance on the day–night task: response set or semantics? *Dev. Sci.* 8, 360–371. doi: 10.1111/j.1467-7687.2005.00424.x

Simpson, A., and Riggs, K. J. (2005b). Inhibitory and working memory demands of the day–night task in children. *Br. J. Dev. Psychol.* 23, 471–486. doi: 10.1348/02615005X28712

Simpson, A., Riggs, K. J., Beck, S. R., Gorniak, S. L., Wu, Y., Abbott, D. et al. (2012). Refining the understanding of inhibitory processes: how response prepotency is created and overcome. *Dev. Sci.* 15, 62–73. doi: 10.1111/j.1467-7687.2011.01105.x

Song, Y., and Hakoda, Y. (2011). An asymmetric Stroop/reverse-Stroop interference phenomenon in ADHD. *J. Atten. Disord.* 15, 499–505. doi: 10.1177/1087054710367607

Spronk, M., Jonkman, L. M., and Kemner, C. (2008). Response inhibition and attention processing in 5- to 7-year-old children with and without symptoms of ADHD: an ERP study. *Clin. Neurophysiol.* 119, 2738–2752. doi: 10.1016/j.clinph.2008.09.010

Yasumura, A., Kokubo, N., Yamamoto, H., Yasumura, Y., Nakagawa, E., Kaga, M., et al. (2014). Neurobehavioral and hemodynamic evaluation of Stroop and reverse Stroop interference in children with attention-deficit/hyperactivity disorder. *Brain Dev.* 36, 97–106. doi: 10.1016/j.braindev.2013.01.003

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 November 2013; accepted: 28 February 2014; published online: 18 March 2014.

Citation: Ikeda Y, Okuzumi H and Kokubun M (2014) Age-related trends of inhibitory control in Stroop-like big–small task in 3 to 12-year-old children and young adults. *Front. Psychol.* 5:227. doi: 10.3389/fpsyg.2014.00227

This article was submitted to Developmental Psychology, a section of the journal *Frontiers in Psychology.*

Copyright © 2014 Ikeda, Okuzumi and Kokubun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.