Cohort-Sequential Study of Conflict Inhibition during Middle Childhood

Leslie Rollins

Christopher Newport University

Tracy Riggins

University of Maryland

Author Note

This research was supported by in part by a Doctoral Dissertation fellowship from the University of Minnesota, a grant from the National Institutes of Health (HD-R01-28425, PI: Patricia J. Bauer), and the Department of Psychology at the University of Maryland, College Park. The authors would like to thank Patricia J. Bauer for her generous support of this work, the members of the Cognition in Transition Laboratory at the University of Minnesota, in particular Marina Larkina, the members of the Neurocognitive Development Lab at the University of Maryland, especially Kelsey Woods, Eliana Sudikoff, Sarah Blankenship, Margaret-Ellen Johnson, Martin Folkoff, Jamie Rosen, Andrew Cherenzia, Lindsay Gerber, Tammy Tran, Rachel Peissner, Dayna Mazza, as well as the families who participated in this study. Portions of these data were presented in DeBoer, T. (2005). A neurobehavioral investigation of autobiographical memory development: Contributions of source memory and memory for temporal order. (Doctoral dissertation, University of Minnesota).

Abstract

This longitudinal study examined developmental changes in conflict inhibition and error correction in three cohorts of children (5, 7, and 9 years of age). At each point of assessment children completed three levels of Luria's tapping task (1980), which requires the inhibition of a dominant response and maintenance of task rules in working memory. Findings suggest that both conflict inhibition and error detection and correction improve significantly during middle childhood. When cognitive demands were high, conflict inhibition, as shown by initial response accuracy, improved steadily across middle childhood. In contrast, the ability to detect and correct for errors improved between 5 and 6 years of age. Further, variability in conflict inhibition decreased with age and individual differences in conflict inhibition were stable across the one-year period in 7- and 9-year-old, but not 5-year-old children. These findings are discussed in relation to previous research on the development of inhibition.

Keywords: Conflict inhibition, conflict tapping, error correction, executive functions, longitudinal study

Cohort-Sequential Study of Conflict Inhibition during Middle Childhood

The intentional inhibition of a dominant response is a core executive function (Miyake et al., 2000). Individual differences in inhibition are associated with reading and mathematics proficiency (e.g., Blair & Razza, 2007), social competence (e.g., Ciairano, Visu-Petra, & Settanni, 2007), pedestrian safety (Barton & Schwebel, 2007), and obesity (e.g., Anzman & Birch, 2009). Inhibition emerges by late infancy and develops throughout childhood (Best & Miller, 2010; Zelazo, Carlson, & Kesek, 2008). Most research on inhibition has been conducted cross-sectionally; few studies have examined inhibition longitudinally, especially during middle childhood (cf., Lee, Bull, & Ho, 2013; Van der Ven, Kroesbergen, Boom, & Leseman, 2012). Longitudinal studies are needed because only they can provide understanding of intra-individual change and stability of individual differences across time (Best & Miller, 2010; Robinson, Schmidt & Teti, 2005). To address this gap, the present study used a one-year longitudinal design to examine the development of conflict inhibition in three cohorts of children (5, 7, and 9 years of age).

The term inhibition has been widely used to refer to many functions. As stated above, the present use of the term is in reference to response inhibition, or the intentional inhibition of a dominant response. This ability has been differentiated from similar constructs such as resistance to proactive interference, reactive inhibition, and inhibition in spreading activation models and connectionist networks (Friedman & Miyake, 2004; Miyake et al., 2000). Principal components analysis on data from children has shown that response inhibition can be further divided into conflict inhibition and delay inhibition (Carlson & Moses, 2001). Conflict inhibition is the ability to act in opposition to a prepotent response (e.g., thanking someone for an undesirable gift) whereas delay inhibition is the ability to postpone providing a dominant response (e.g., waiting until the end of the party to open gifts; Carlson & Moses, 2001). Conflict and delay inhibition are functionally dissociable, contributing differentially to abilities, such as theory of mind, in childhood (Carlson & Moses, 2001).

The focus of the current study is on the development of conflict inhibition. Cross-sectional investigations suggest that during childhood there are improvements in conflict inhibition as well as error detection and correction, which are needed when conflict inhibition fails. Precisely when age-related differences are found depends upon the task utilized. During early childhood, performance robustly improves on the day/night, grass/snow, and bear/dragon tasks as well as Luria's hand games; performance reaches ceiling levels on these tasks by approximately 6 years (Carlson & Moses, 2001; Gerstadt, Hong, & Diamond, 1994; Hughes, 1998). However, when more complex tasks are used that demand a speeded response, cross-modal integration, and/or additional cognitive processes (e.g., working memory), conflict inhibition improves throughout childhood and into adolescence (Best & Miller, 2010). For example, performance on Luria's tapping task (Klenberg, Korkman, & Lahti-Nuuttila, 2001), the Flanker task (e.g., Simonds, Kieras, Rueda, & Rothbart, 2007), the Stop-Signal task (Bedard et al., 2002), and the Go/No-go task (e.g., Huizinga, Dolan, & van der Molen, 2006) shows improvement during middle childhood. Improvement on the Stroop and Erikson Flanker tasks continues during middle childhood and, depending on the task demands, into adolescence (e.g., Huizinga et al., 2006; Ikeda, Okuzumi, & Kokubun, 2014; Leon-Carrion, García-Orza, Perez-Santamaria, 2004; Klenberg, Närhi, Korkman, & Hokkanen, 2015; Macdonald, Beauchamp, Crigan, & Anderson, 2014).

The few longitudinal investigations of conflict inhibition extend cross-sectional studies by examining intra-individual change in conflict inhibition during middle childhood; together, they suggest small to moderate improvement during this period. For example, Lee and colleagues (2013) reported reductions in reaction time (RT) differences between congruent and incongruent trials in kindergarten, 2nd, 4th, and 6th grade cohorts across 4 year-long waves for Flanker, Simon, and Mickey tasks. Improvements in accuracy were less pronounced, only increasing across waves in the kindergarten cohort on the Flanker task (children were at ceiling-level performance (~95%) on the Simon and Mickey tasks). A 1-year longitudinal study using factor analysis did not show evidence for developmental improvements in the factor that reflected inhibition/shifting between 1st and 2nd grade (Van der Ven et al., 2012). One potential issue with these previous longitudinal studies is that the tasks utilized (Flanker, Mickey, and Simon tasks) may not have been sufficiently challenging to allow for the examination of developmental change in inhibition accuracy during middle childhood. Thus, more research in middle childhood is needed using complex tasks that require a speeded response, cross-modal integration, and/or additional cognitive processes (e.g., working memory) to assess conflict inhibition since children are less likely to achieve ceiling-level performance on such tasks.

Luria's (1980) tapping task is one measure that has been used successfully in cross-sectional studies across a wide age range. This task requires the inhibition of the prepotent response to imitate the experimenter's prompt (e.g., tapping twice when the experimenter taps twice). An advantage of this task is that level of difficulty can be manipulated by altering the modality of the experimenter prompt and the participant's response while keeping most task demands constant (e.g., task instructions). Three versions of the tapping task have been used in developmental research (Auditory/Visual-Same, Visual-Different, and Auditory-Different). The first term refers to the modality of the experimenter prompt and the second term refers to whether the child's response is in the same or different modality as the experimenter's prompt.

Specifically, in the auditory conditions the child hears the experimenter tap a rod on the table once or twice; in the visual conditions the child can see the experimenter tap on the table or raise her finger/arm(s). On the easiest version of the task (i.e., Auditory/Visual-Same), the child has to respond by tapping twice on the table when the experimenter taps once and vice versa. Previous studies using this version have either presented children with combined visual and auditory prompts (Diamond & Taylor, 1996; Klenberg et al., 2001) or just auditory prompts (Becker et al., 1987; Passler, Isaac, & Hynd, 1985). On this version of the task, cross-sectional studies have found that children rapidly improve between 3.5 and 4 years of age and reach ceiling-level performance at approximately 6 years (Becker et al., 1987; Diamond & Taylor, 1996; Klenberg et al., 2001; Passler et al., 1985). In contrast, performance on more complex versions follows an extended developmental trajectory. Adults and 7-to 9-year-old children perform similarly on the Visual-Different version (i.e., experimenter prompts by raising a finger, child responds by tapping table twice), but children commit more errors than adults on the Auditory-Different version (i.e., experimenter prompts with one tap, child responds by lifting two fingers; Cycowicz et al., 2001).

The present study was designed to examine conflict inhibition as well as error detection and correction in 5-, 7-, and 9-year-old children longitudinally over a one year period using all three versions of Luria's (1980) tapping task (Auditory-Same, Visual-Different, Auditory-Different). We hypothesized that developmental improvements in conflict inhibition as well as error detection and correction would be present during middle childhood and that these improvements would be more prominent on more complex levels of the tapping task (especially the Auditory-Different level). We explored whether different age-related improvements were present for initial and final response accuracy. Developmental change in initial response accuracy would suggest improvement in conflict inhibition. Change in final response accuracy, relative to initial response accuracy, would suggest improvement in error detection and correction because, if the child initially responded incorrectly, the child had to identify that an error was made and then provide the correct response. We also hypothesized that variability in performance accuracy would decrease with age and that individual differences in conflict inhibition would remain stable over a one year period. Stability in individual differences would suggest that, overall, children develop at the same rate whereas instability would suggest that children develop at various rates. The goals of the present study were to examine the development of conflict inhibition in middle childhood by 1) utilizing the tapping task, which should reduce the probability of participants reaching ceiling-level performance as they did in a previous longitudinal investigation (Lee et al., 2013), at least on the Auditory-Different Level, 2) assessing initial and final response accuracy to allow for the examination of change in error detection and correction 3) examining changes in variability in conflict inhibition performance and 4) investigating how stable individual differences in conflict inhibition are over a period of one year.

Method

Participants

Participants were recruited from a database maintained by a large midwestern University in the United States. Participants included 90 children taking part in a 3-year longitudinal study who contributed data for all three levels of the tapping task (see Procedure) at the second and third year assessments (i.e., Time 1 and 2; data from the initial year were excluded because too few children from the youngest Cohort contributed data for all 3 levels of the task). Consistent with the community from which the sample was drawn, 91% of the participants were of

Caucasian, non-Hispanic descent. During Time 1, 26 participants (13 males) were 5-years-old (*M* = 5.23 years, *SD* = .1), 34 participants (16 males) were 7-years-old (*M* = 7.2 years, *SD* = .11), and 30 participants (13 males) were 9-years-old (*M* = 9.23 years, *SD* = .1). Time 2 data were collected after a 1-year delay (*M* = 330.37 days, *SD* = 26.12). Delay did not differ between cohorts, $F(2, 87) = .23$, $p = .795$. Of the children that participated in the second and third year assessments, 21 children were excluded from present analyses due to attrition between Time 1 and Time 2 ($n = 6$), incomplete performance ($n = 9$), experimenter error ($n = 4$), or failure of the practice trials ($n = 2$). There were no differences in gender or race between children included or excluded from the present analyses. The University Institutional Review Board approved all procedures before the study began, and parents provided consent for their children. Children received a small gift and parents received a gift certificate for participating.

Procedure

Each year children completed a cognitive assessment battery. Luria's (1980) tapping task was performed at the end of the 1-2 hour testing session. The administration procedure for the tapping task resembled previous research (Becker et al., 1987; Cycowicz et al., 2001; Diamond & Taylor, 1996; Klenberg et al., 2001; Passler et al., 1985). Children were required to provide a response (e.g., tapping twice when the experimenter taps once) that conflicted with the prepotent response to mimic the experimenter (e.g., tapping twice when the experimenter taps twice). Children completed task levels in order of ascending difficulty: Auditory-Same, Visual-Different, Auditory-Different. The first term refers to the experimenter prompt modality; the second term refers to whether the participant response modality was the same as or different from the experimenter's prompt modality. Thus, during the Auditory-Same level, participants tapped a wooden rod once when the experimenter tapped twice and twice when the experimenter tapped once. In the Visual-Different level, participants tapped a wooden rod once when the experimenter raised two arms and twice when the experimenter raised one arm. In the Auditory-Different level, participants raised one arm when the experimenter tapped twice and two arms when the experimenter tapped once.

Before each level, the experimenter administered training and practice trials. If necessary, the experimenter re-explained the instructions and provided a maximum of 3 additional practice trials. Test trials were only administered if children passed the practice trials. The Auditory-Same level included 10 test trials and the Visual-Different and Auditory-Different levels included 20 test trials. Test trial order was standard across participants, and whether participants began with a one or two tap/arm raise trial was counterbalanced across participants. Coders (*N* = 10) identified from video whether the child accurately responded, self-corrected, or incorrectly responded. Coder reliability was assessed for a random sample (approximately distributed across Cohorts) of 20% of the videos at each time point. Average agreement between coders was 97% (*SD* ± 2). Dependent measures included the proportion of accurate initial responses (i.e., children's first response), accurate final responses (i.e., last response, allowing for self-corrections), and self-corrections. These measures were assessed for all trials and specifically switch trials (i.e., trials for which the prompt differed from the previous trial).

Analytic Approach

The effects of Time, Cohort, and Level on initial and final response accuracy were assessed using generalized estimating equation (GEE) analyses, an extension of the generalized linear model, because they account for the potential correlation between repeated measurements, which is common in longitudinal research (Ballinger, 2004; Hardin & Hilbe, 2003) and do not assume that measures are normally distributed. Separate 2 Time x 3 Cohort (5-, 7-, 9-year-old) x

3 Level (Auditory-Same, Visual-Different, Auditory-Different) GEE analyses were conducted for initial and final response accuracy. We utilized an unstructured correlation matrix and took a Wald Chi-Square Type III sum of squares approach. Main effects and interactions were assessed with the least significant difference method for pairwise comparisons of estimated marginal means. One-tailed, directional tests were used because we hypothesized improvements with age. Because Cohort-related improvement was linear for initial response accuracy but non-linear for final response accuracy (final response accuracy only improved between the 5- and 7-year-old Cohorts), we examined the proportion of self-corrections on switch trials (i.e., trials for which the prompt differed from the previous trial) during the most complex level of the tapping task (i.e., Auditory-Different) using a 2 Time x 3 Cohort GEE analysis. Changes in response variability were also assessed for the Auditory-Different Level by calculating the coefficient of variation (CV) and conducting a Levene's test to assess homogeneity of adjusted coefficient of variation scores. The Levene's test assessed main effects of Time and Cohort as well as the Time x Cohort interaction. Relative performance stability was measured using Spearman's rho correlation coefficients on Time 1 and Time 2 performance for each cohort.

Results

Changes in Response Accuracy

GEE results for initial and final response accuracy are shown in Table 1.[1] Results for level are always presented in order of ascending difficulty (i.e., Auditory-Same, Visual-Different, Auditory-Different).

Initial response accuracy. The GEE analysis on initial response accuracy revealed significant main effects of Time, Cohort, and Level (Table 1, Figure 1). Children overall performed better at Time 2 (M = .9) than Time 1 (M = .85; p < .001). Similarly, pairwise

[1] Analyses for switch trials revealed the same pattern of results and are not presented for the sake of brevity.

comparisons associated with the main effect of Cohort revealed that age was linearly related to initial response accuracy; the 9-year-old Cohort (*M* = .92) performed better than the 7-year-old Cohort (*M* = .89; *p* = .038) and the 7-year-old Cohort performed better than the 5-year-old Cohort (*M* = .83; *p* = .005). Lastly, the manipulation of task difficulty was successful because performance was highest on the Auditory-Same Level (*M* = .95), intermediate on the Visual-Different Level (*M* = .92), and lowest on the Auditory-Different Level (*M* = .76; *ps* < .001).

Significant Time x Cohort and Time x Level interactions also emerged. The Time x Cohort interaction suggested that the 5-year-old (*p* < .001) and the 7-year-old (*p* = .003) Cohorts but not the 9-year-old Cohort (*p* = .278) significantly improved between time points. This finding suggests that the development of conflict inhibition occurs at a faster rate earlier in middle childhood.

The Time x Level interaction suggested that improvement between time points was greater as the complexity of the tapping task increased. Improvement in original accuracy between Time 1 and Time 2 was not significant for the Auditory-Same Level (*p* = .057) but was for the Visual-Different (*p* = .013) and the Auditory-Different Levels (*p* < .001).

Final response accuracy. The GEE analysis revealed significant main effects of Time, Cohort, and Level (Table 1, Figure 1). Similar to the finding for initial response accuracy, children performed better at Time 2 (*M* = .96) than Time 1 (*M* = .92; *p* < .001). The pairwise comparisons associated with the main effect of Cohort revealed that the 7- year-old Cohort (*M* = .96) and 9-year-old Cohort (*M* = .97) significantly outperformed the 5-year-old Cohort (*M* = .9; *ps* < .001), but that they did not differ from one another (*p* = .21)., Lastly, performance was highest on the Auditory-Same Level (*M* = .97), intermediate on the Visual-Different Level (*M* = .94), and lowest on the Auditory-Different Level (*M* = .92; *ps* < .05).

There was also a significant Time x Cohort interaction. Children in the 5-year-old ($p < .001$) and the 7-year-old ($p = .04$) but not the 9-year-old Cohort ($p = .158$) demonstrated significant improvement in performance between time points. Again, this finding suggests that the development of conflict inhibition occurs at a faster rate earlier in middle childhood.

Improvement between time points was greater as the complexity of the tapping task increased, as indicated by a Time x Level interaction. Change within individuals was not significant for the Auditory-Same Level ($p = .152$) or the Visual-Different Level ($p = .089$) but was for the Auditory-Different Level ($p < .001$). Thus, improvement within individuals was greatest on the most challenging level of the tapping task.

Lastly, there was a Cohort x Level interaction because age-related differences were most pronounced on more challenging levels of the task. This interaction was followed-up by conducting GEEs separately for each level of the tapping task. No significant Cohort effect was present for the Auditory-Same Level, Wald χ^2 (2), = 2.951, $p = .229$. However, there was a main effect of Cohort for the Visual-Different Level, Wald χ^2 (2), 7.114, $p = .029$; pairwise comparisons showed that the 9-year-old Cohort performed significantly better than the 5-year-old ($p = .009$) and 7-year-old Cohorts ($p = .04$) and that the 7-year-old Cohort performed marginally better than the 5-year-old Cohort ($p = .057$). There was also a main effect of Cohort for the Auditory-Different Level, Wald χ^2 (2), = 17.073, $p < .001$; pairwise comparisons revealed that the 7- and 9-year-old Cohorts performed significantly better than the 5-year-old Cohort ($ps < .001$), but did not differ from one another ($p = .33$). This pattern suggests that age-related differences were larger on the more complex versions of the tapping task.

Improving final response accuracy by self-correcting. The GEE analysis on the proportion of self-corrections during the Auditory-Different Level revealed a significant Time x

Cohort interaction, Wald χ^2 (2) = 6.29, *p* = .043. Self-corrections were greatest in the 5-year-old Cohort at Time 2 (i.e., 6-year-old children) and 7-year-old Cohort at Time 1 (i.e., 7-year-old children; see Figure 2). There was an increase in self-corrections for the 5-year-old Cohort between Time 1 and Time 2 (*p* = .034). Self-corrections were similar between the 5-year-old Cohort at Time 2 and the 7-year-old Cohort at Time 1 (*p* = .117). The 7-year-old Cohort self-corrected more at Time 1 than Time 2 (*p* = .044), and more than the 9-year-old Cohort at both Time 1 (*p* = .029) and Time 2 (*p* = .024). This pattern suggests that 6- and 7-year-old children's performance benefitted most from their ability to detect and correct errors.

Changes in Response Variability

Figure 3 illustrates individual growth trajectories, which, taken together across Cohorts provides graphical depiction of changes in response variability. The Levene's test revealed a significant main effect of Time, *F*(1, 174) = 17.77, *p* < .001; children's responses were more variable at Time 1 (CV = 26.5%) than Time 2 (CV = 16.4%). There was also a significant main effect of Cohort, *F*(2, 174) = 5.25, *p* = .006. Response variability was significantly higher in the 5-year-old Cohort (CV = 32%) than the 9-year-old Cohort (CV = 18.2%), *p* = .004. Response variability in the 7-year-old Cohort (CV = 25.8%) did not differ from either the 5-year-old (*p* = .251) or 9-year-old Cohort (*p* = .298). The Time x Cohort interaction was not significant, (*p* = .42). Together, these results suggest that response variability decreases with age.

Stability in Individual Differences

In addition to understanding changes in response accuracy and variability, it is important to assess stability in individual differences (i.e., whether children that perform well relative to their peers at Time 1 continue to perform well relative to their peers at Time 2). Stability in individual differences was significant in the 7-year-old (r_s = .42, *p* = .014) and 9-year-old

Cohorts (r_s = .51, p = .004), but not the 5-year-old Cohort (r_s = .25, p = .252), which suggests that individual differences in performance are stable by middle childhood.

Discussion

The goal of the present study was to assess the development of conflict inhibition longitudinally during middle childhood. Using a cohort-sequential design, we examined developmental changes in conflict inhibition using three versions of Luria's (1980) tapping task between 5 and 10 years of age. Our results suggest that 1) conflict inhibition improves steadily during middle childhood, 2) error detection and correction increase substantially between 5 and 6 years of age, 3) response variability in conflict inhibition decreases during middle childhood, and 4) individual differences across a one year period were stable in the 7- and 9-year-old cohorts but not the 5-year-old cohort. These findings, with a longitudinal sample, are consistent with and extend previous studies of middle childhood that have examined conflict inhibition. Consistent with previous research, we demonstrated improvements in conflict inhibition during middle childhood (Bedard et al., 2002; Huizinga et al., 2006; Ikeda et al. 2014; Klenberg et al., 2001, 2015; Lee et al., 2013; Leon-Carrion et al., 2004; Macdonald et al. 2014; Simonds et al., 2007). We extended this research by demonstrating that the pattern of improvement observed in cross-sectional studies (Bedard et al., 2002; Huizinga et al., 2006; Klenberg et al., 2001; Leon-Carrion et al., 2004; Simonds et al., 2007) also accurately characterizes growth trajectories in individual participants. Further, the present findings showed that the rate of developmental change in initial and final response accuracy was greater earlier in middle childhood. One advantage to assessing both initial and final response accuracy was that we were able to also examine the development of error detection and correction. All children engaged in error detection and correction; however, there was a marked improvement in final response accuracy between 5 and 6 years of age. This difference may be due to improvements in performance monitoring, error detection, the deployment of error correction, or a combination of these abilities. Unfortunately, these cannot be distinguished in the present study. However, our findings are consistent with previous studies showing that error detection (as indexed by the error related negativity (ERN) ERP component) increases between 7- and 25 years of age (Wiersema, van der Meere, & Roeyers, 2007). The current study also showed that variability in response accuracy decreased with age. This finding is consistent with and extends cross-sectional research showing that response variability within-individuals on Stroop and Flanker tasks follows a U-shaped pattern across the lifespan with children and older adults demonstrating greater variability than young adults (Li, Hämmerer, Müller, Hommel, & Lindenberger, 2009; Williams, Strauss, Huitsch, & Hunter, 2007).

One major advantage of the longitudinal design of the present study was that it allowed for the examination of stability of responses with age. Individual differences in performance on the Auditory-Different version of the task were stable over the course of a year in the 7- and 9-year-old Cohorts but not the 5-year-old Cohort. This result suggests that children in the 5-year-old Cohort are developing at different rates, whereas improvements in the 7- and 9-year-old Cohorts were similar across participants. This pattern of results may be due to the significant improvement in conflict inhibition occurring early in middle childhood and children undergoing development at differing times.

Previous longitudinal investigations of conflict inhibition during middle childhood have suffered from children achieving ceiling-level performance on response accuracy measures (e.g., Lee et al., 2013). The present study was partially able to overcome this problem by utilizing three versions of Luria's (1980) tapping task. Task difficulty was manipulated by altering the modality of the experimenter prompt and the participants' response. Consistent with previous cross-sectional studies, participants reached ceiling-level performance in regards to accuracy on the Auditory-Same and Visual-Different levels of the task by approximately 6 years of age (Becker et al., 1987; Cycowicz et al., 2001; Diamond & Taylor, 1996; Klenberg et al., 2001; Passler et al., 1985). However, original response accuracy on the Auditory-Different level increased linearly during middle childhood and reached only 80% by 10 years of age (Cycowicz et al., 2001). Thus, the manipulation of task difficulty was successful and allowed us to assess developmental change in conflict inhibition on the Auditory-Different version of the tapping task.

We acknowledge that the present study has limitations that should be addressed in future research. First, the current research only assessed conflict inhibition. Future research should directly compare the developmental trajectory of conflict and delay inhibition. Second, the entire period of middle childhood was assessed through use of a cohort-sequential design. Thus, differences as a function of age reflect a combination of longitudinal and cross-sectional comparisons. Future studies could use an entirely longitudinal design by sampling children each year throughout middle childhood. There are also some limitations associated with the use of Luria’s (1980) tapping task. For example, it is currently unclear whether the versions of the tapping task differed merely in difficulty (i.e., inhibitory demand) and not in task complexity (i.e., required abilities beyond inhibition). Although task instructions remained constant across versions, it is possible that the degree of response automaticity or demand on working memory, set-shifting, or cross-modal integration may have differed across versions. Additionally, as stated above, although ceiling-level performance was less problematic in the current study than previous investigations, children’s final response accuracy was quite high on all levels of the task by the end of middle childhood. Lastly, the present study specifically examined changes in response accuracy with age. In the future, longitudinal studies should examine developmental changes in both accuracy and reaction time because cross-sectional research suggests that reaction time changes more dramatically than response accuracy during middle childhood (Ikeda et al., 2014; Klenberg et al., 2015).

In conclusion, the current findings suggest that conflict inhibition, error detection, and error correction undergo significant development during middle childhood. This finding is consistent with previous cross-sectional research on the development of executive functioning and extends this work by showing that 1) rates of improvement were greatest in the youngest cohort, 2) variability in performance is substantially reduced with age, and 3) individual differences in conflict inhibition show considerable stability toward the end of middle childhood.

References

Anzman, S. L., & Birch, L. L. (2009). Low inhibitory control and restrictive feeding practices predict weight outcomes. *Journal of Pediatrics*, *155*(5), 651-656. doi:10.1016/j.jpeds.2009.04.052

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods*, *7*(2), 127-150. doi: 10.1177/1094428104263672

Barton, B. K., & Schwebel, D. C. (2007). The roles of age, gender, inhibitory control, and parental supervision in children's pedestrian safety. *Journal of Pediatric Psychology*, *32*(5), 517-526. doi: doi:10.1093/jpepsy/jsm014

Becker, M. G., Isaac, W., & Hynd, G. W. (1987). Neuropsychological development of nonverbal behaviors attributed to "frontal lobe" functioning. *Developmental Neuropsychology*, *3*(3/4), 275-298. doi: 10.1080/87565648709540381

Bedard, A-C., Nichols, S., Barbosa, J. A., Schachar, R., Logan, G. D., Tannock, R. (2002). The development of selective inhibitory control across the lifespan. *Developmental Neuropsychology*, *21*(1), 93-111. doi: 10.1207/S15326942DN2101_5

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641-1660. doi:10.1111/j.1467-8624.2010.01499.x

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647-663. doi: 10.1111/j.1467-8624.2007.01019.x

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032-1053. doi: 10.1111/1467-8624.00333

Ciairano, S., Visu-Petra, L., & Settanni, M. (2007). Executive inhibitory control and cooperative behavior during early school years: A follow-up study. *Journal of Abnormal Child Psychology*, *35*, 335-345. doi: 10.1007/s10802-006-9094-z

Cycowicz, Y. M., Friedman, D., Snodgrass, J. G., & Duff, M. (2001). Recognition and source memory for pictures in children and adults. *Neuropsychologia*, *39*, 255-267. doi: 10.1016/S0028-3932(00)00108-1

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*(4), 315-334. doi: 10.1002/(SICI)1098-2302(199605)29:4%3C315::AID-DEV2%3E3.3.CO;2-C

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*(1), 101-135. doi: 10.1037/0096-3445.133.1.101

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 ½-7 years old on a Stroop-like day-night test. *Cognition*, *53*, 129-153. doi: 10.1016/0010-0277(94)90068-X

Hardin, J. W., & Hilbe, J. M. (2003). *Generalized estimating equations*. Boca Raton, FL: Chapman and Hall/CRC Press.

Hughes, C. (1998). Finding your marbles: Does preschoolers’ strategic behavior predict later understanding of mind? *Developmental Psychology*, *34*(6), 1326-1339. doi: 10.1037/0012-1649.34.6.1326

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*, 2017-2036.

Ikeda, Y., Okuzumi, H., & 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. *Frontiers in Psychology*, *5*, 1-6. doi: 10.3389/fpsyg.2014.00227

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- and 12-year-old Finnish children. *Developmental Neuropsychology*, *20*, 407-428. doi: 10.1207/S15326942DN2001_6

Klenberg, L., Närhi, V., Korkman, M., & Hokkanen, L. (2015). Examining methodological variation in response inhibition: The effects of outcome measures and task characteristics on age-related differences. *Child Neuropsychology*, *21*(5), 586-602. doi: 10.1080/09297049.2014.950215

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental change in executive functioning. *Child Development*, *84*(6), 1933-1953. doi: 10.1111/cdev.12096

Leon-Carrion, J., García-Orza, J., Perez-Santamaria, F. J. (2004). Development of the inhibitory component of the executive functions in children and adolescents. *International Journal of Neuroscience*, *114*(10), 1291-1311. doi:10.1080/00207450490476066

Li, S. -C., Hämmerer, D., Müller, V., Hommel, B., & Lindenberger, U. (2009). Lifespan development of stimulus-response conflict cost: Similarities and differences between maturation and senescence. *Psychological Research*, *73*(6), 777-785. doi: 10.1007/s00426-008-0190-2

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed., B. Haigh, Trans.). New York: Basic.

MacDonald, J. A., Beauchamp, M. H., Crigan, J. A., & Anderson, P. J. (2014). Age-related differences in inhibitory control in the early school years. *Child Neuropsychology*, *20*(5), 509-526. doi: 10.1080/09297049.2013.822060

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49-100. doi: doi:10.1006/cogp.1999.0734

Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology*, *1*(4), 349-370. doi: 10.1080/87565648509540320

Robinson, K., Schmidt, T., & Teti, D. M. (2005). Issues in the use of longitudinal and cross-sectional designs. In D. M. Teti (Ed.), *Handbook of Research Methods in Developmental Science* (pp. 3-20). Malden, MA: Blackwell Publishing.

Simonds, J., Kieras, J. E., Rueda, M. R., & Rothbart, M. K. (2007). Effortful control, executive attention, and emotional regulation in 7-10-year-old children. *Cognitive Development*, *22*, 474-488. doi: 10.1016/j.cogdev.2007.08.009

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology*, *82*, 100-119. doi: 10.1111/j.2044-8279.2011.02035.x

Wiersema, J. R., van der Meere, J. J., & Roeyers, H. (2007). Developmental changes in error monitoring: An event-related potential study. *Neuropsychologia*, *45*(8), 1649-1657. doi: 10.1016/j.neuropsychologia.2007.01.004

Williams, B. R., Strauss, E. H., Huitsch, D. F., & Hunter, M. A. (2007). Reaction time inconsistency in a spatial Stroop task: Age-related differences throughout childhood and adulthood. *Aging, Neuropsychology, and Cognition: A Journal on Normal and Dysfunctional Development*, *14*(4), 417-439. doi: 10.1080/13825580600584590

Zelazo, P. D., Carlson, S. M., & Kesek, A. (2008). The development of executive function in childhood. In C. Nelson & M. Luciana (Eds), *Handbook of Developmental Cognitive Neuroscience (2nd Ed.)*. Cambridge, MA: MIT Press.

Table 1

Summary of GEE Type III Wald Chi-Square values (bold indicates significant effects).

Dependent measure	Time	Cohort	Level	Time x Cohort	Cohort x Level	Time x Level	Time x Cohort x Level
Initial Accuracy	**22.78**	**20.33**	**203.87**	**13.68**	8.03	**13.21**	1.25
Final Accuracy	**17.62**	**18.55**	**26.32**	**10.76**	**9.58**	**19.75**	5.94

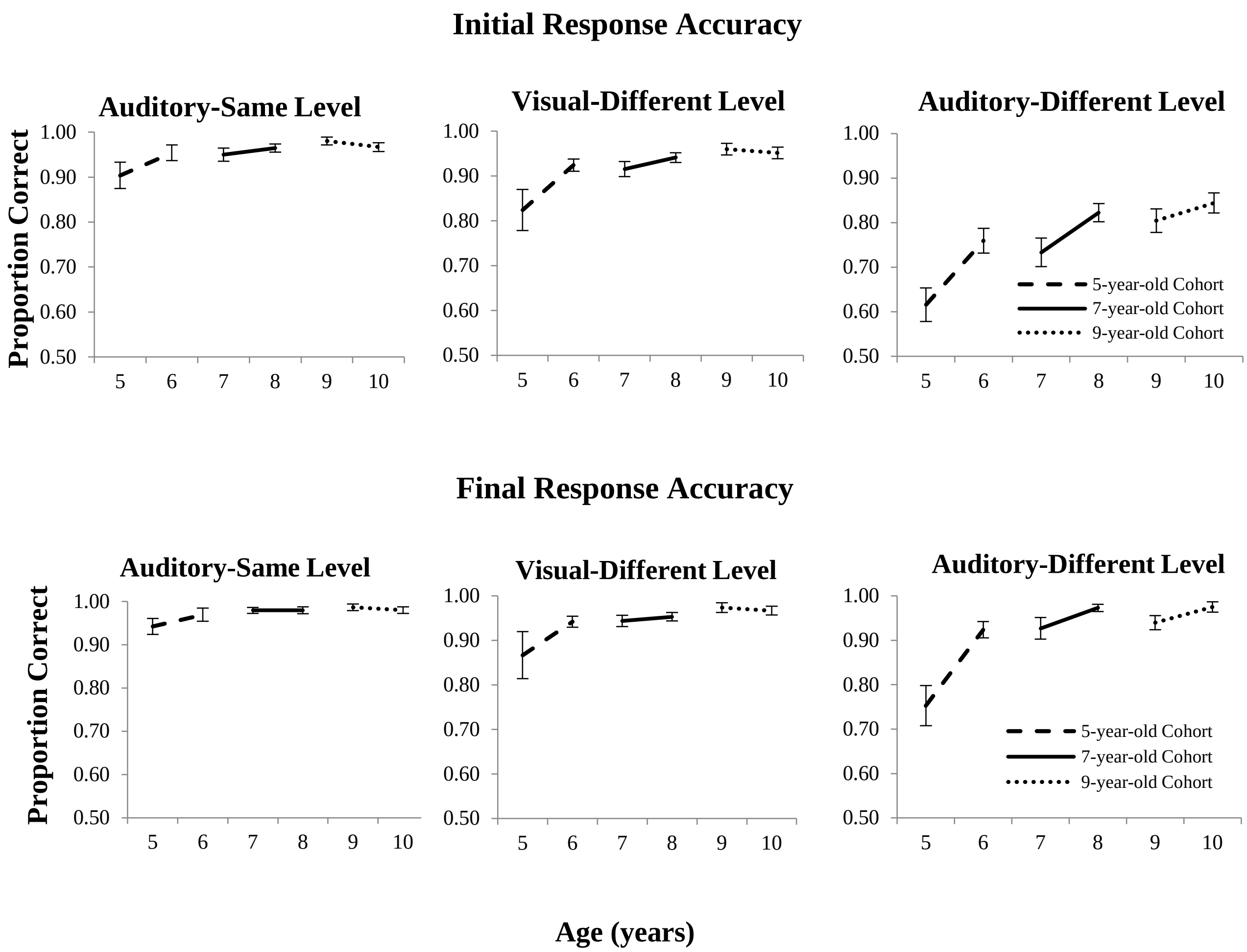

Figure 1. Initial and final response accuracy rates for each level of the tapping task. Error bars represent standard errors.

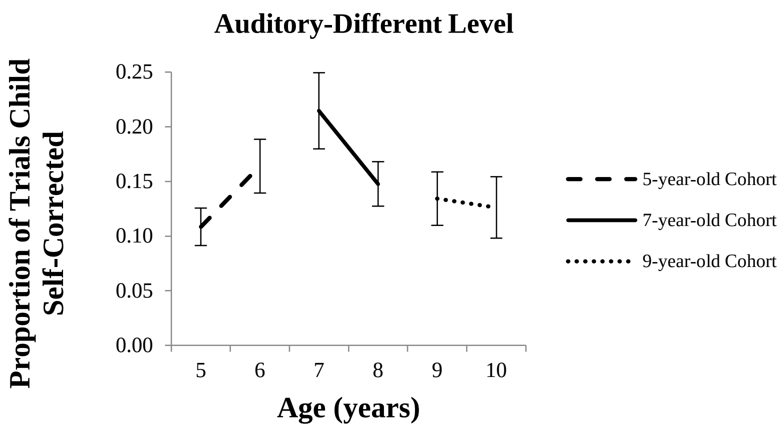

Figure 2. Proportion of switch trials on which children self-corrected on the Auditory-Different Level of the tapping task. Error bars represent standard errors.

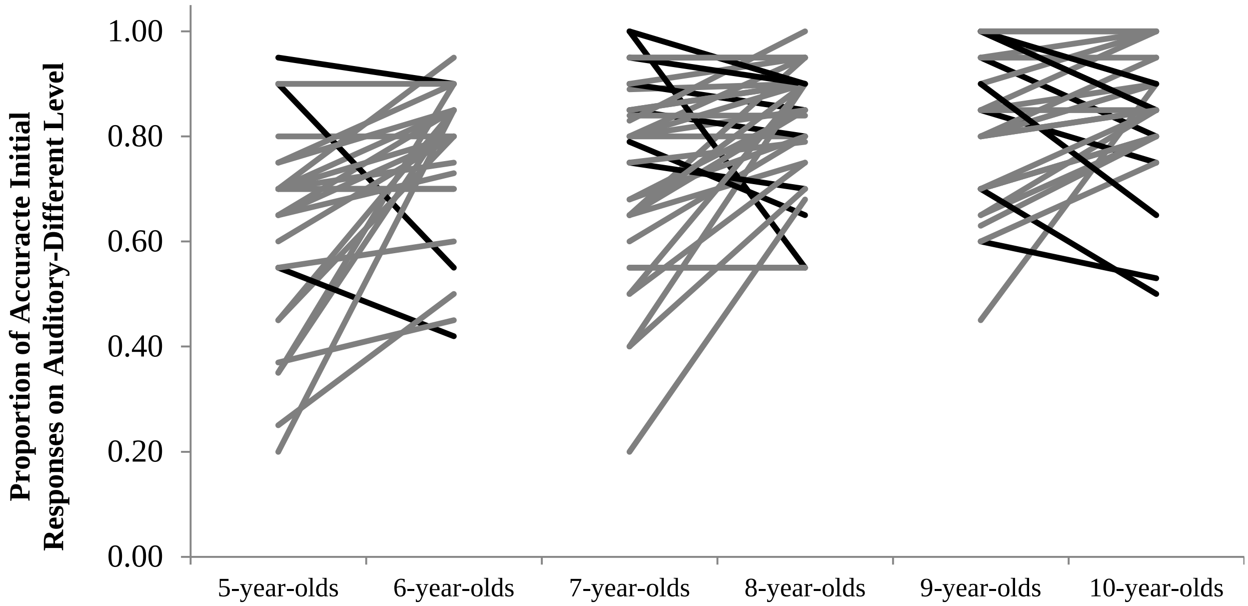

Figure 3. Growth trajectories for initial response accuracy on the Auditory-Different Level of the tapping task. Children whose performance remained the same or improved from Time 1 to Time 2 are depicted by gray lines and children whose performance decreased from Time 1 to Time 2 are depicted by black lines.