Supplemental Methods

ERP Sample Demographic and Perinatal Characteristics.

Sample characteristics for the children who provided useable ERP data are provided in Tables S1-S2. Excluded subjects generally did not differ from those included in the ERP dataset on demographic or perinatal characteristics described in these tables, although for full-term children, c-section delivery was associated with inclusion in the ERP dataset ($p < .02$). No other demographic or perinatal characteristics were associated with inclusion in the ERP dataset (p's > .05).

Across all children, despite the narrow age range utilized in the study (4.5-5 years), age at test was the only other variable associated with data inclusion. Children who were older at the time of the testing session were more likely to contribute useable ERP data ($p < .01$).

Supplemental Results

Effects of Moderate-to-Late Prematurity on ERP Measures of EF

We examined the impact of moderate-to-late preterm birth on ERP component amplitude and latency during the directional Stroop task. Models described in the FT cohort in the main paper were re-run as a mixed model with all original predictors, as well as an additional group factor (PT versus FT) and the interaction between group and trial type. Means and standard deviations are provided in Supplemental Tables S3-S4. Figures S1-S3 illustrate the average waveforms by condition in each group.

Task effects (i.e. significant effects of trial type) generally remained consistent in this larger sample of children with those described in the previous section, unless explicitly described below. However, we generally did not detect significant differences between PT and FT children in ERP amplitudes or latencies for either early attentional or later cognitive components.

Early attentional components.

N1. There was no effect of group (PT versus FT) on N1 amplitude, $F(1, 90) = .95$, $p < .33$, or latency, $F(1, 90) = .05$, $p < .82$, at Fz. Similarly, there was no interaction between group and trial type for N1 amplitude, $F(1, 90) = 1.14$, $p < .29$, or latency, $F(1, 90) = .12$, $p < .74$. Within the PT group, there were no statistically significant relationships between individual variation in N1 amplitudes and/or latencies and gestational age at birth.

P2. There was a trend-level effect of group on P2 amplitude, $F(1, 90) = 3.06$, $p < .08$, but no interaction between group and trial type, $F(1, 90) = 1.84$, $p < .18$, at Fz. PT children showed a trend-level decrease in P2 amplitudes in comparison to their FT peers at Fz. However, within the PT group there was no significant relationship between P2 amplitude and gestational age at birth.

The main effect of trial type for P2 reached significance in this larger sample, $F(1, 90) = 6.13$, $p < .02$. There was no effect of group on P2 latency, $F(1, 90) = .01$, $p < .91$. However, a trend-level interaction between group and trial type, $F(1, 90) = 3.14$, $p < .08$, suggested that PT children showed greater latency differences by trial type than their FT peers. Within the PT group there were no significant relationships between P2 latency and gestational age at birth.

Later cognitive components.

N2. There was no effect of group on N2 amplitude, $F(1, 90) = .00$, $p < .99$, and there were no significant interactions involving group. Similarly, there was no effect of group on N2 latency, $F(1, 90) = .57$, $p < .45$, and there were no significant interactions involving group. Within the PT group, there were no significant relationships between individual variation in N2 amplitudes and/or latencies and gestational age at birth.

P3 complex. There was no effect of group on P3 amplitude at Cz, $F(1, 90) = .19$, $p < .66$, and no interaction involving group, $F(1, 90) = .02$, $p < .89$. Although the effect of trial type on P3 latency reached trend-level significance in this larger data set, $F(1, 90) = 3.15$, $p < .06$, there was no effect of group on latency, $F(1, 90) = .20$, $p < .66$, and no interaction involving group, $F(1, 90) = .23$, $p < .63$. Within the PT group, there were also no significant relationships between individual variation in P3 amplitudes and/or latencies at Cz and gestational age at birth.

There was no effect of group on P3a amplitude at Pz, $F(1, 90) = .33$, $p < .57$, and no interaction between group and trial type, $F(1, 90) = .44$, $p < .51$. Similarly, there was no effect of group on P3b amplitude at Pz, $F(1, 90) = .03$, $p < .87$, and no interaction between group and trial type, $F(1, 90) = .73$, $p < .40$. There was no effect of group on P3a latency, $F(1, 90) = .78$, $p < .38$, but an interaction between group and trial type was present, $F(1, 90) = 3.86$, $p < .05$, whereby PT children showed greater latency differences by trial type than their FT peers. The main effect of

trial type for P3b latency reached significance in this larger data set, $F(1, 90) = 5.54$, $p < .02$. However, there was no effect of group, $F(1, 90) = 1.06$, $p < .31$, or interaction between group and trial type, $F(1, 90) = .01$, $p < .94$, for P3b latency. Within the PT group, higher gestational age at birth was related to a faster latency to peak at P3b for incongruent trials at Pz, $r(43) = -.33$, $p < .03$. All other relationships between gestational age at birth and individual variation in P3 complex amplitudes and latencies at Pz were non-significant.