SUPPORTING INFORMATION

Table S1. Weighted cueing levels used for each step. Each cue level is delivered twice when the participant is not following or engaged in the task. There is a 10 second delay between every cue, with the exception of cases where a child is distressed or in danger.

Cue (Level)	Description
No cues (Level 0)	The participant is able to complete step without support.
General verbal cue (Level 1) x 2	Participant requires prompting with open-ended question that will help him/her proceed with the step. e.g. "What is the next step?" or "What else do you need?"
Gestural cue (Level 2) x 2	Administrator may move his/her hand to demonstrate without words (e.g. demonstrating how to open ink pad) or use pointing (e.g. point to where the participant may find the item, point to the recipe book picture, or point to the appropriate place on the paper). However, the administrator does not handle any of the items or physically participate.
Direct verbal cue (Level 3) x 2	The participant requires a direct one-step instruction. E.g., "The recipe shows that the red circle is over here" "You need the timer for this part" or "You need the scissors to cut the grass"
Physical	The administrator physically assists the participant with a single part of the step. E.g. Retrieve a necessary item from the

assistance (Level 4) x 2	box or put glue on back of circle and wait for participant to stick in correct position.
Do for participant (Level 5) x1	The administrator completes the step that the participant is demonstrating difficulty with using self-talk, and then waits for the participant to proceed to next step.

Table S2. Task score descriptions

Score	Description
Total Score (TS)	Total score is a weighted score based on the number and level of cues required throughout task.
Total Cues (TC)	Total number of cues required throughout whole task.
Total task time (Time)	Time taken to complete the task.
Highest Cue Level	Highest level of cue required during task (5 levels as per Table S1).
Initiation	Amount of cues required to independently start task.
Learning	An improvement in the amount of cues required for part A and part C in a section that repeats the same step three times can be informative of an individual child's learning process.
Sequencing	Number of steps without any cues required.
Meta-cognition	Number of cues required to figure out that the timer is required to time for one minute whilst blowing on the ink
Judgment/Safety	Number of safety cues required throughout whole task

Completion	Number of cues required to finish task/realize task is completed
Working Memory	The child can be rated from 1 (poor working memory) to 3 (superior working memory) with 2 denoting a child who is observed to have typical working memory for their age-based on manual guidance.
Organisation	The child can be rated from 1 (low organisation skills) to 3 (highly organised) with 2 denoting a child who is observed to have organisational skills as typical for their age-based on manual guidance.
Emotional Lability	The child can be rated from 1 (very emotionally labile) to 3 (low lability) with 2 denoting a child who is observed to have emotional reactions as typical for their age-based on manual guidance.
Distractibility	The child can be rated from 1 (very distractible) to 3 (not distractible) with 2 denoting a child who is observed to typical levels of distractibility for their age-based on manual guidance.
Pre-task self-judgment	A child's judgment or prediction of their own ability to complete task

	independently
Post-task self-judgment	A child's judgment of how much help they received to complete task
Post-task self-review	A child's judgment of how good a job they did during task
Self-talk	Coded as Yes or No

Table S3. Normative data for quantitative PETA domains

Variable	Total Mean, SD (range)	3 year olds	4 year olds	5 year olds
Total Score	46.6 38.3 (4-202)	84.7 42.2 (20-202)	41.8 26.6 (4-139)	22.85 18.2 (4-84)
Total Cues	26.3 15.8 (4-73)	41.8 14.7 (12-73)	25.3 11.8 (4-59)	15.5 9.3 (4-44)
Time to Complete	13.96 3.9 (6-26)	16.7 4.0 (7.5-26)	13.9 3.4 (7-25)	11.9 2.8 (6-17.5)
Initiation	2.6 (0-9)	4.1 (0-9)	2.7 (0-9)	1.4 (0-9)
Sequencing	1.87 (0-7)	0.80 (0-6)	1.8 (0-6)	2.7 (0-7)
Meta-cognition	4.27 (0-9)	5.84 (2-9)	4.42 (1-9)	2.9 (0-7)

Judgment/Safety	43 (0-5)	.53 (0-5)	.46 (0-5)	.32 (0-5)
Completion	1.1 (0-6)	1.4 (0-6)	1.1 (0-5)	0.7 (0-3)

Table S4. Data for qualitative PETA domains

Domain	*%*
Working Memory	
Poor	*11*
Typical	*51*
Very Good	*38*
Organisation	
Poor	*17*
Typical	*38*
Very Good	*45*

Emotional Lability

Poor

6

Typical

62

Very Good

32

Distractibility

Poor

13

Typical

46

Very Good

41

Highest Level of Support	
Verbal Guidance	*4.9*
Gestural Guidance	*19.5*
Direct Verbal	*41.5*
Physical Assistance	*23.2*
Examiner Completes	*11*
Self-talk	
No	*51.1*
Yes	*48.8*

Table S5. PETA Self-ratings

How much help do you think you will need?	
None	*30%*
A little	*44%*
A lot	*26%*
How much help did you need?	
None	*42%*
A little	*44.8%*
A lot	*13.2%*
Do you think you did a good job?	
Yes	*78.7%*
No	*21.3%*

Table S6. Inter-correlations between quantitative PETA domains

	PETA TS	PETA TC	Initiation	Sequencing	Meta-cog	Judgment	Completion	Time
PETA TS	-	.972**	.639**	-.618**	.717**	.170*	.265**	.691**
PETA TC	.972**	-	.662**	-.703**	.731**	.205**	.319**	.738**
Initiation	.639**	.662**	-	-.465**	.481**	.018	.177*	.540**
Sequencing	-.618**	-.703**	-.465**	-	-.534**	-.146	-.305**	-.574**
Meta-cog	.717**	.731**	.481**	-.534**	-	.158*	.243**	.574**
Judgment	.170*	.205**	.018	-.146	.158*	-	.068	.118
Completion	.265**	.319**	.177**	-.305**	.243**	.068	-	-.369**
Time	.691**	.738**	.540**	-.574**	..574**	.118	.369**	-

*p<.05 ** p <.01 *** p <.005

Table S7. Total group demographics and separated for each age range

Variable	Total	3 year olds	4 year olds	5 year olds

Participants tested (N)	166	45	60	61
Male (N, %)	87 (53%)	23	32	32
White British (N, %)	99 (60%)	23	43	33
Black British (N, %)	24 (15%)	7	5	12
Other Ethnic Minority (N, %)	41 (25%)	15	10	16
Low SES	49 (29%)	16	12	21

Table S8. Task descriptives

Variable	Total Mean Mean (SD)	3 year olds Mean (SD)	4 year olds Mean (SD)	5 year olds Mean (SD)
PIQ	101.6(15.9)	103.9(15.5)	102.0(16.3)	99.1(14.7)

VIQ	107.7(16.5)	106.2(18.0)	107.3(16.6)	108.5(16.1)
BRIEF GEC	48.3(10.4)	49.2(9.6)	48.7(11.0)	47.5(10.4)
BRIEF EMI	49.3 (10.8)	49.9 (11.3)	50.0 (10.6)	48.6 (10.)
BRIEF ISCI	47.4 (8.7)	48.9 (7.6)	47.4 (9.1)	46.3 (9.2)
BRIEF FI	45.8 (8.7)	45.6 (8.6)	45.1 (5.9)	46.3 (10.4)
CBQ attention [b]	4.9 (0.94)	4.68 (.85)	4.89 (.94)	5.04 (.94)
CBQ inhibitory Control [b]	4.9 (0.8)	4.69 (.73)	4.98 (.80)	4.92 (.8)
NIH toolbox attention/inhibition [c]	102.9 (19.9)	101.3 (5.7)	110.2 (25)	98.5 (22.2)
		(N=8)	(N=8)	(N=10)

a BRIEF-P missing for n=44 children b CBQ missing for n=7 c Age-adjusted scores (3 year olds: N=8; 4 year olds: N=8, 5 year olds: N=10)

Influence of self-talk, gender, and socioeconomic status on performance

The influence of self-talk (yes/no), gender (make/female), and socio-economic status (SES; lower/higher)[1] were investigated separately for the Total Summary Score, Total Number of Cues, and Completion Time. Overall, the use of self-talk had no influence on Completion Time ($F_{1,134}$= .282, p =.596) but those who engaged in self-talk obtained a better Total Summary Score (M=35.26, SD=23.6vM=47.51, SD=40.5; $F_{1,134}$= 4.52, p =.035) and showed a trend for a difference in Total Number of Cues (M=22.14, SD=11.3vM=26.7, SD=17.5; $F_{1,134}$= 3.14, p =.079). There were no differences in the rates of self-talk between age-ranges. When investigated further, self-talk had no effect on performance for the four- and five-year-olds but the three-year olds who did not engage in self-talk were found to require more support to complete the task (Total Number of Cues, t(133)=1.77, p =.004; Total Summary Score, t(133)=2.13, p=.003) but no group difference for Completion Time. A non-significant trend for gender differences

---

[1] SES was based on home postcode to estimate total house income on a scale from the UK Office for National Statistic for children attending the laboratory with school postcode used for children assessed on school sites (Nation, Cocksey, Taylor, & Bishop, 2010). Children were divided into low and high SES groups based on the mean scale score. Children in the lower SES group were from households with an estimated average weekly net income (before housing costs) of less than £480 per week. The average weekly figure in the most recent London survey was £620 (Bond & Campos, 2010).

showed that girls tended to receive less cues overall (M=23.8, SD=14.8 vM=28.4, SD=16.3; t(162)=-1.71, *p* =.06) and a lower Total

Summary Score (M=41.2, SD=36.2 vM=51.4, SD=39.7; t(162)=-1.90, *p* =.09), but there was no difference for Completion Time.

Poorer performance in the "lower SES" group (N=47) was observed for Total Number of Cues (M=30.5, SD=18.9 vM=24.5,

SD=13.9; t(162)=-2.4; *p* =.04), Total Summary Score (M=58, SD=46.5 vM=42.1, SD=33.7; t(162)=-2.1; *p* =.05), and longer

Completion Time (M=14.9, SD=4.2 vM=13.5, SD=3.2; t(162)-2.0; *p* =.04).