Bernardi, M., Leonard, H.C., Hill, E.L., Botting, N. & Henry, L.A. (2017). A two-year follow-up study of executive functions in children with developmental coordination disorder. *Developmental Medicine & Child Neurology.* (Wiley).

A two-year follow-up study of executive functions in children with Developmental Coordination Disorder

Marialivia Bernardi[a*], Hayley C. Leonard[b], Elisabeth L. Hill[c], Nicola Botting[a] & Lucy A. Henry[a]
[a]Division of Language and Communication Science, City, University of London, UK
[b]School of Psychology, University of Surrey, Surrey, UK
[c]Department of Psychology, Goldsmiths, University of London, UK

[*]Corresponding author:
Marialivia Bernardi; Division of Language and Communication Science, City, University of London, London, EC1V 0HB, United Kingdom. Marialivia.Bernardi@city.ac.uk

Abstract

Aim: Executive Function (EF) impairments have been identified in children with motor difficulties, with and without a diagnosis of developmental coordination disorder (DCD). However, most studies are cross-sectional. This study investigates the development of EF in children with poor motor skills over two years.

Method: Children aged 7-11 years (*N*=51) were assessed twice, two years apart, on verbal and nonverbal measures of EFs: executive-loaded working memory; fluency; response inhibition; planning; and cognitive flexibility. Typically developing children (TD: *n*=17) were compared to those with a clinical diagnosis of DCD (*n*=17) and those with identified motor difficulties (MD: *n*=17), but no formal diagnosis of DCD.

Results: Developmental gains in EF were similar between groups, although a gap between children with poor motor skills and TD children on nonverbal EFs persisted. Specifically, children with DCD performed significantly more poorly than TD children on all nonverbal EF tasks and verbal fluency tasks at both time points; and children with MD but no diagnosis showed persistent EF difficulties in nonverbal tasks of working memory and fluency.

Interpretation: Children with DCD and MD demonstrated EF difficulties over two years, which may impact on activities of daily living and academic achievement, in addition to their motor deficit.

What this paper adds

- EF difficulties in children with poor motor skills persist throughout middle childhood.
- Children with motor difficulties (MD), without a DCD diagnosis, demonstrate less pervasive EF difficulties than children with DCD.
- EF difficulties in MD and DCD groups affect mostly nonverbal domains.
- All groups showed similar developmental gains in EF.

Developmental Coordination Disorder (DCD) is a condition affecting 5% of the population[1] diagnosed on the basis of a significant motor coordination impairment impacting on activities of daily living, in the absence of any physical, neurological or intellectual disability. Individuals with DCD not only experience a motor coordination deficit but also report difficulties with personal organisation, planning, time management, memory, and decision making, which continue into adulthood[2]. These skills are underpinned by cognitive processes known as executive functions (EFs) that regulate, monitor and control behaviour towards a goal[3]. EFs are a strong predictor of academic achievement throughout childhood[4] and continue to predict general success in life during adulthood[5]. Therefore, understanding EFs in DCD is crucial for improving life outcomes for individuals with motor coordination impairments.

Previous research has identified EF deficits in children with DCD or poor motor skills (see Wilson et al.[6], and Leonard and Hill[7] for recent reviews). However, this research is largely cross-sectional. To date, two studies have assessed EF longitudinally in early childhood: in 5-6 year-old children with poor manual dexterity skills[8]; and in 4-6 year-olds screened for motor difficulties[9]. In both studies, children were followed-up one year later and those with persistent motor impairments demonstrated performance gains with age in EF tasks. However, poorer EFs were identified at both time points when compared to a sample of children with average or above average motor coordination scores, matched for age, gender and intellectual ability.

It is currently not understood whether EFs in children with DCD or poor motor skills follow a developmental trajectory similar to that of their typically-developing peers, who demonstrate continued improvement in EF skills throughout middle childhood and adolescence[10]. Importantly, different EF constructs mature at different ages[11], and some seem to reach adult levels between 8-12 years[12]. A longitudinal perspective reflecting developmental change in later childhood is essential to better understand the nature of EF difficulties in children with motor impairments.

The current study is a follow-up of previous research conducted by Leonard and colleagues[13]. They recruited children of between 7-11 years by screening for movement difficulties, as well as through clinical diagnoses of DCD. Two groups of children with poor motor skills, namely a DCD group and a motor difficulty (MD) group, were compared separately with a group of typically developing (TD) children. A comprehensive EF assessment battery was administered including parallel verbal and non-verbal measures in five EF domains. The battery included measures of executive-loaded working memory (ELWM; concurrently storing and processing information), response inhibition (suppressing unhelpful, yet automatic, prepotent responses), and cognitive flexibility (switching flexibly between strategies or tasks in response to feedback). Although these three domains are identified as ‘core’ EF skills[14], a three-factor model is not as strong when applied to children, for whom a broader set of five factors may be more appropriate[15]. Therefore, measures of planning (strategically organising a sequence of actions) and fluency (generating responses in response to instruction), which have previously been used in populations with neurodevelopmental disorders[16,17] were also included in the battery. Leonard and colleagues[13] reported that both the MD and DCD groups performed significantly more poorly than TD children on *nonverbal* tests of ELWM, inhibition and fluency. There were no reported differences in performance on switching tasks, but the MD group scored significantly below TD children on the task measuring nonverbal planning abilities. Critically, no differences in performance were found on any *verbal* EF tasks.

Two years later these children were followed up with the same EF assessment battery, and these data are presented here to provide a longitudinal perspective on EF in children with poor motor skills (DCD and MD). Three research questions were put forward: RQ1) Do children with poor motor skills show persistent EF difficulties at each time point compared to TD children? RQ2) Do children with poor motor skills demonstrate gains in EF? RQ3) If so, how do these gains compare to those of TD children?

Based on the original study findings[13], it was expected that children with DCD and MD would demonstrate difficulties in *nonverbal* EF tasks compared to TD children, and that these difficulties would be evident at both time points. It was predicted that at least some gains in EF performance would be apparent for both groups, but that these may vary between EF domains, as well as between verbal versus nonverbal task types.

Method

Participants

Ethical approval was obtained from the Language and Communication Science Proportionate Review Board at City, University of London. Parents of children who participated in the original study[13] were then approached. Informed consent was obtained from 56 parents and their children (61.5 % of the original sample) to take part in this follow-up study.

At Time 1, participants with DCD were recruited on the basis of an existing diagnosis from a qualified professional, which was corroborated by the research team using the Movement Assessment Battery for Children (2nd ed.; MABC-2)[18] and Checklist, along with parent reports and a standardised IQ assessment, the British Abilities Scales 3rd Edition (BAS-3)[19]. A normative school sample was also assessed with the MABC-2. Children with scores at or below the 16th percentile were identified as having motor difficulties (MD group) and those scoring at or above the 25th percentile were included in the TD group. Any child scoring more than two standard deviations below the mean on the BAS-3 was excluded, as were any children in the DCD group with additional diagnoses of attention-deficit hyperactivity disorder or autism spectrum disorder, or any medical condition. Parents reported no diagnoses for any child in the TD and MD groups.

At Time 2 children were assigned to their original groups: TD (n=20), DCD (n=19) and MD (n=17). However, to confirm group membership and suitability for the study, participants were re-assessed on motor and cognitive ability. Five children were excluded from the sample because they no longer met criteria for their original group (2 DCD, 3 TD; see Supplementary Materials for further details). The final sample, therefore, included 51 children, 17 in each group (25 males; mean age: 8.9 years, SD: 1.1 years, range: 7.20–11.9). Background characteristics are presented for each group in Table 1, together with group comparisons on these measures.

--- Table 1 about here ---

Measures

A comprehensive EF assessment battery was administered, including a verbal and a nonverbal measure for each of the following EFs: executive-loaded working memory; fluency; response inhibition; planning; and cognitive flexibility (see Table 2 for a summary, and Supplementary Materials for further details). These measures were identical to those administered at Time 1 and reported in the previous study[13].

--- Table 2 about here ---

Procedure

Children who were seen at the research lab or in their home completed the assessment on the same day or over two to three sessions of 1.5 – 2 hours. Children who were tested in their school (66% at Time 1 and 48% at Time 2) completed five or six sessions of 45 minutes – one hour each. All children were assessed individually in a quiet room and sufficient breaks were given between tasks to maintain motivation. Task order was varied to suit the child's needs and offer maximum variety.

Statistical analysis

Hierarchical multiple regressions were conducted to explore any differences in EF performance between groups at each time point. Since participants in this follow-up were a subgroup of the original sample[10], regressions were conducted at both Time 1 and Time 2 in order to compare the same subgroup of participants across time. A multiple regression approach was taken so that the group differences in age and IQ (which are reported in Table 1, and are important for EF development) could be controlled at Step 1 of each regression, before examining whether there were group differences in EF performance at Step 2 using two dummy-coded Group variables. The reference group was always TD children, (i.e., TD vs. MD; TD vs. DCD). Bonferroni corrections were applied to the final models ($p \leq .005$).

A repeated measures MANOVA was used to test for differences in EF performance between the two time points and identify whether the group variable had an impact on these differences over time. Group was entered as the between-subjects factor (3 levels) and time as the within-subjects factor (2 levels), and all EF measures were entered as dependent variables[a].

Results

The means, standard deviations and ranges of scores for each of the 10 EF measures at both time points are presented in Table 3. The data met all assumptions for the following analyses (see Supplementary Materials).

--- Table 3 about here ----

Significant group differences at each time point in EF performance (RQ1) from the multiple regression analyses are reported in the text below. Summary details of Step 2 of each regression for all EF tasks are reported in Table 4.

--- Table 4 about here ----

On the *nonverbal ELWM* task, the MD and DCD groups performed significantly more poorly than the TD group at both time points.

On the *nonverbal fluency* task the final regression model at Time 1 became a non-significant trend (p=.007) after applying Bonferroni correction, whereas at Time 2 it remained significant. The MD and DCD groups performed more poorly than the TD group at both times.

On the *nonverbal response inhibition* and *nonverbal planning* tasks there was a significant group difference between the MD and TD groups at Time 1, which was not evident at Time 2. The DCD group performed more poorly than the TD group at both time points on both tasks.

[a]Age was not included because the analyses aimed to assess EF gains over time *irrespective* of age changes. Age was taken into account in the first set of analyses by entering it into Step 1 of the hierarchical multiple regressions.

On the *verbal fluency* and *nonverbal switching* tasks no differences between the MD and TD groups were identified. The DCD group performed significantly more poorly than the TD group at both time points on both tasks[b].

In summary, children with DCD obtained poorer scores than TD children on all nonverbal EF tasks, as well as on verbal fluency, at both time points. Children with MD at Time 1 performed more poorly than TD children in all nonverbal EF domains except switching; however, at Time 2, nonverbal planning and nonverbal inhibition differences were no longer evident and only nonverbal ELWM and nonverbal fluency differences persisted.

A repeated measures MANOVA addressed the second and third research questions investigating whether children with poor motor skills demonstrate gains in EFs and how these gains compare to those of TD children.

A significant effect of Time $F(1,45)=12.11$, $p<.001$, $\eta_p^2=.771$ was identified. Univariate tests indicated the effect of Time was significant for verbal ELWM $F(1,45)=32.42$, $p<.001$, $\eta_p^2=.419$, nonverbal ELWM $F(1,45)=11.25$, $p=.002$, $\eta_p^2=.200$, verbal fluency $F(1,45)=20.21$, $p<.001$, $\eta_p^2=.310$, nonverbal fluency $F(1,45)=34.10$, $p<.001$, $\eta_p^2=.431$, nonverbal planning $F(1,45)=6.76$, $p=.013$, $\eta_p^2=.131$, verbal switching $F(1,45)=13.12$, $p=.001$, $\eta_p^2=.226$, and nonverbal switching $F(1,45)=5.10$, $p=.029$, $\eta_p^2=.102$. The effect of time was non-significant for verbal inhibition $F(1,45)=.30$, $p=.59$, $\eta_p^2=.007$, nonverbal inhibition $F(1,45)=1.37$, $p=.25$, $\eta_p^2=.030$, and verbal planning $F(1,45)=.70$, $p=.79$, $\eta_p^2=.002$.

There was a main effect of Group $F(1,45)=3.17$, $p<.001$, $\eta_p^2=.462$. However, group differences have been assessed through the previous regression analyses and will not be discussed further.

The relevant result for RQ3 was the outcome of the interaction between Time and Group, which was non-significant $F(1,45)=.94$, $p=.54$, $\eta_p^2=.202$. Thus, EF performance changed in a similar way over time in each group.

Discussion

The current study investigated EF difficulties over two years in 7-11 year-old children with poor motor skills. As predicted, children with poor motor skills showed persistent EF difficulties at both time points, largely associated with nonverbal domains of EF. In particular, children with a diagnosis of DCD performed significantly more poorly than TD children at both time points on *all* nonverbal measures of EF, and also on verbal fluency. Children without a DCD diagnosis, but with equivalent motor difficulties (MD group), also demonstrated poorer performance at Time 1 on nonverbal EF tasks (all nonverbal EF tasks

---

[b]Additional regression analyses were conducted to directly compare children with DCD and MD across the 10 EF measures. The two groups differed significantly in *verbal fluency* at both time points (Final model Time 1, $F(4,45)=5.49$, Adj. $R^2=.27$, $p=.001$, DCD vs. MD: $B=7.72$, $SE\ B=2.80$, $p=.008$; Final model Time 2, $F(4,46)=6.09$, Adj. $R^2=.29$, $p=.001$, DCD vs. MD: $B=7.87$, $SE\ B=3.35$, $p=.023$), and in *nonverbal switching* at both time points (Final model Time 1, $F(4,46)=9.36$, Adj. $R^2=.40$, $p<.001$, DCD vs. MD: $B=-9.60$, $SE\ B=4.37$, $p=.033$; Final model Time 2, $F(4,46)=7.10$, Adj. $R^2=.33$, $p<.001$, DCD vs. MD: $B=-8.36$, $SE\ B=3.81$, $p=.033$.

except switching). However, at Time 2 only nonverbal fluency and nonverbal ELWM difficulties persisted in this group.

Also in accordance with predictions, significant improvements over time across all three groups were detected in many EF tasks: verbal and nonverbal ELWM, fluency and switching; and nonverbal planning. The fact that performance on the VIMI task did not improve over time is consistent with studies in typical populations suggesting that the ability to inhibit a prepotent response changes rapidly in early childhood but becomes more stable with age[11], and may develop earlier than other EF domains[24]. Critically, the interaction between time and group was non-significant across the EF domains. Therefore, no differences between groups were identified in the pattern of developmental change in EF over a period of two years. This result suggests that the gap in EF performance identified in children with DCD and MD compared to TD children, tends to remain stable during middle childhood.

Findings are consistent with longitudinal studies in younger populations of children with poor motor skills[8,9]. Furthermore, the fact that mainly nonverbal EF difficulties were identified at both time points in the MD and DCD groups supports recent findings that the links between motor and cognitive brain networks may lag behind those of TD controls during childhood[25].

Although the pattern of growth in EF abilities was similar between groups, some of the difficulties encountered by children with MD at Time 1 were not evident at Time 2 (nonverbal inhibition and nonverbal planning). Therefore, it is important to clarify with further longitudinal research whether specific EF domains reach typical levels of ability at a later stage during development, or whether impairments persist into adulthood. EF difficulties may have a growing impact on everyday life and academic achievement, given that the executive load of the environment is likely to increase with age while support decreases (e.g., transition to secondary school). Understanding which factors can lead to an improvement in EF will be vital in identifying those at most risk of falling behind[3].

Children with DCD demonstrated more pervasive EF difficulties over time than children with MD. The significant differences in nonverbal switching and verbal fluency performance between the MD and DCD groups cannot be attributed to an intermediate level of motor impairment in the MD group, because the range and mean of MABC-2 scores did not differ between these two groups. Perhaps given the relatively low awareness of DCD amongst parents, teachers, and clinicians[26], children with fewer or less obvious EF difficulties may be less likely to be flagged for clinical referral, despite similar levels of motor difficulty. Children with better EF may be able to deal with everyday tasks more effectively, and require less support. However, not all children with MD may show this EF profile over time, so it is important for future research to investigate this group and help to identify those that are in need of extra support.

An important finding was that children with poor motor skills did worse than TD children largely on *nonverbal* EF tasks. This suggests that EF difficulties in children with DCD and MD are primarily linked to their core impairments rather than to more domain general cognitive processing problems. The nonverbal EF tasks in the current study had either a motor or a visuo-spatial demand, and the strong links between areas of the brain associated with these functions and those involved in executive control goes someway to explaining the EF difficulties seen in DCD. Indeed, previous research has suggested atypical functioning of prefrontal and parietal cortices and the cerebellum[27], as well as atypical connectivity or coupling between these areas[25], in children with DCD. However, it should be noted that the DCD group also had difficulties with verbal fluency, and that everyday situations require the ability to master *both* verbal and nonverbal domains of EF simultaneously and adaptably. It remains important to focus not only on reducing nonverbal demands in everyday and school-related tasks for children with poor motor skills, but to consider the cognitive load of tasks overall in order to support these children effectively.

Although the current study was rigorous in its sampling and produced in-depth data from each child over developmental time, there are limitations that should be addressed in future research. First, the small sample size meant that more complex statistical techniques, such as multi-level modelling or a cross-sequential design, were not appropriate - hence, some more subtle group differences in age-related changes in EF ability may not have been captured. It might be expected that younger children would show a greater improvement over time than older children[10], so future research should collect larger age-stratified samples to address this issue. Second, although children with additional diagnoses were excluded from the DCD sample, subclinical symptoms could still have an impact on EF. This was tested in the original study[13], and these symptoms did not significantly predict performance for any EF measure. However, conducting further research with larger samples, including those with co-occurring disorders, will be important in order to provide a fuller picture of the individual differences in a representative clinical sample. Third, our study focused on standardised and experimental measures of EF, in which task demands are set by the experimenter and do not necessarily represent the demands of EF tasks in everyday life. More ecologically valid measures of EF assessing real-life situations and 'hot' EFs, including emotional and motivational aspects, might further contribute to understanding EF difficulties associated with poor motor skills[7].

In conclusion, children with poor motor skills, both with and without a DCD diagnosis, demonstrated a range of EF difficulties that persisted over two years. EF problems largely affected nonverbal domains and were less developmentally persistent in children with MD without a diagnosis of DCD. Both the MD and DCD groups showed significant gains in EFs over middle childhood that matched those of the TD group, indicating that EF progression over time was at the level expected.

Acknowledgements

This work was supported by The Waterloo Foundation under grant 920-2318, and by a City, University of London PhD Studentship to the first author. We would like to thank all of the children, parents, teachers, headteachers and classroom assistants who kindly assisted with this project.

References

[1]American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). Washington, DC: American Psychiatric Association, 2013.

[2]Tal-Saban M, Ornoy A, Parush S. Executive function and attention in young adults with and without Developmental Coordination Disorder–A comparative study. *Res Dev Disabil* 2014; **35**: 2644-50.

[3]Diamond A. Executive functions. *Annu Rev Psychol* 2013; **64**: 135-168.

[4]Gathercole SE, Pickering SJ, Knight C, Stegmann Z. Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Appl Cogn Psychol* 2004; **18**: 1-6.

[5]Moffitt TE, Arseneault L, Belsky D, Dickson N, Hancox RJ, Harrington H, Houts R, Poulton R, Roberts BW, Ross S, Sears MR. A gradient of childhood self-control predicts health, wealth, and public safety. *Proc Natl Acad USA* 2011; **108**: 2693-8.

[6]Wilson PH, Smits-Engelsman B, Caeyenberghs K, et al. Cognitive and neuroimaging findings in developmental coordination disorder: new insights from a systematic review of recent research. *Dev Med Child Neurol* 2017 Sep 5. doi: 10.1111/dmcn.13530 [E-pub ahead of print].

[7]Leonard HC, Hill, EL. Executive difficulties in developmental coordination disorder: methodological issues and future directions. *Curr Dev Disord Rep* 2015; **2**:141-149.

[8]Michel E, Roethlisberger M, Neuenschwander R, Roebers CM. Development of cognitive skills in children with motor coordination impairments at 12-month follow-up. *Child Neuropsychol* 2011; **17**:151-72.

[9]Michel E, Molitor S, Schneider W. Differential changes in the development of motor coordination and executive functions in children with motor coordination impairments. *Child Neuropsychol* 2016 Sep 13. doi: 10.1080/09297049.2016.1223282. [E-pub ahead of print].

[10]Zelazo, PD, Anderson, JE, Richler J, Wallner-Allen K, Beamont JL, and Weintraub S. NIH Toolbox Cognition Battery (CB): measuring executive function and attention. *Monogr Soc Res Child Dev* 2013; **78**: 16-33.

[11]Davidson MC, Amso D, Anderson LC, Diamond A. Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*. 2006; **44**: 2037-78.

[12]Huizinga M, Dolan CV, van der Molen MW. Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia* 2006; **44**: 2017-36.

[13]Leonard HC, Bernardi M, Hill EL, Henry LA. Executive functioning, motor difficulties, and developmental coordination disorder. *Dev Neuropsychol* 2015; **40**: 201-15.

[14]Miyake A, Friedman NP. The nature and organization of individual differences in executive functions four general conclusions. *Curr Dir Psychol Sci* 2012; **21**: 8-14.

[15]Levin HS, Fletcher JM, Kufera JA, Harward H, Lilly MA, Mendelsohn D, Bruce D, Eisenberg HM. Dimensions of cognition measured by the Tower of London and other cognitive tasks in head-injured children and adolescents. *Dev Neuropsychol* 1996; **12**: 17-34.

[16]Pennington BF, Ozonoff S. Executive functions and developmental psychopathology. *J Child Psychol Psychiatry* 1996; **37**: 51-87.

[17]Henry LA, Messer DJ, Nash G. Executive functioning in children with specific language impairment. *J Child Psychol Psychiatry* 2012; **53**: 37-45.

[18]Henderson SE, Sugden DA, Barnett AL. Movement assessment battery for children (Movement ABC-2) (2nd ed.). London: The Psychological Corporation, 2007.

[19]Elliot CD, Smith P. British Ability Scales – Third Edition (BAS-3). London: GL Assessment, 2011

[20]Pickering S, Gathercole SE. Working memory test battery for children (WMTB-C). London: The Psychological Corporation, 2001.

[21]Henry LA. How does the severity of a learning disability affect working memory performance? *Memory* 2001; **9**: 233-47.

[22]Delis DC, Kaplan E, Kramer JH. Delis–Kaplan Executive Function System (D-KEFS®). London: Psychological Corporation, 2001.

[23]Cambridge Cognition. The Cambridge Neuropsychological Test Automated Battery (CANTAB®): Test–retest reliability characteristics. Cambridge: Author, 2006

[24]Klenberg L, Korkman M, Lahti-Nuuttila P. Differential development of attention and executive functions in 3-to 12-year-old Finnish children. *Dev Neuropsychol* 2001; **20**: 407-28.

[25]Ruddock S, Caeyenberghs K, Piek J, Sugden D, Hyde C, Morris S, Rigoli D, Steenbergen B, Wilson P. Coupling of online control and inhibitory systems in children with atypical motor development: A growth curve modelling study. *Brain Cognition* 2016; **109**: 84-95.

[26]Kirby A, Davies R, Bryant A. Do teachers know more about specific learning difficulties than general practitioners? *Brit J Special Educ* 2005; **32**: 122-6.

[27]Debrabant J, Gheysen F, Caeyenberghs K, Van Waelvelde H, Vingerhoets G. Neural underpinnings of impaired predictive motor timing in children with Developmental Coordination Disorder. *Res Dev Disabil* 2013; **34**: 1478-87.

Table 1

Means, standard deviations (in parenthesis) and *ranges* of age and scores on motor and intellectual ability tasks in typically-developing children (TD), children screened for motor difficulties (MD) and children with a diagnosis of Developmental Coordination Disorder (DCD). One-way ANOVA Welch adjusted *F* values, degrees of freedom (in parenthesis) and effect sizes are reported for age, intellectual ability scores and motor skills.

Measure	**TD Group** (*n*=17;11 girls)	**MD group** (*n*=17; 9 girls)	**DCD group** (*n*=17; 4 girls)	**ANOVA** *Welch adjusted*
	Mean (SD) *Range*	Mean (SD) *Range*	Mean (SD) *Range*	*F*(df) η_p^2
Time1 – Chronological Age (Months)	109.14 (10.92) *90.33-128*	100.76 (7.37) *93.22-124.22*	118.82 (13.96) *97-143*	*11.91* (2,29.89)*** .320
Time2 – Chronological Age (Months)	135.01 (11.60) *116.22-157*	126.13 (6.91) *118-148*	144.18 (14.48) *121-169*	*11.97* (2,29.03)*** .306
Time1 – BAS3 General Conceptual Ability Score	108.47 (12.46) *92-138*	96.82 (17.02) *71-125*	98.88 (12.81) *78-119*	*3.50* (2,31.51)* .122
Time2 – BAS3 General Conceptual Ability Score	117.29 (17.42) *89-153*	99.47 (22.57) *70-136*	104.41 (12.08) *79-127*	*4.21* (2,30.04)* .158
Time1 – MABC-2 Percentile	58.82 (20.13) *25-95*	3.76 (2.68) *0.5-9*	5.71 (5.74) *0.1-16*	*61.08* (2,25.29)*** .823
Time2 – MABC-2 Percentile	51.06 (21) *25-84*	5.35 (4.01) *1-16*	2.22 (2.58) *0.1-9*	*46.32* (2,27.11)*** .774

Note. MABC-2 = Movement Assessment Battery for Children; BAS3 = British Abilities Scales. Children with DCD were significantly older than TD children at Time 1 (*p*=.037) and children with MD at both time points (*p*s<.001); TD children obtained significantly higher intellectual ability scores than the MD group at Time 2 (*p*=.015); TD children had higher motor ability than the DCD and MD groups at both time points (*p*s<.001).

*$p \leq .05$; ** $p \leq .01$; *** $p \leq .001$.

Table 2.
Description of tasks administered to assess Executive Functions.

EF Measured	Domain	Task	Description	Outcome Variable
Executive-Loaded Working Memory	Verbal	Listening Recall (Working Memory Test Battery for Children[20])	Participants recall the last word of a sentence after making a judgement as to whether the sentence was true or false, with the number of sentences increasing as the task continues.	Total correct trials
	Nonverbal	Odd-One-Out[21]	A nonverbal equivalent of the above task, in which participants recall the spatial location of a nonsense shape after making a judgement as to which of the shapes was the 'odd one out'.	Total correct trials
Fluency	Verbal	Verbal Fluency (D-KEFS[22])	Participants generate as many words as possible belonging to two different specific categories, within one minute.	Total correct responses
	Nonverbal	Design Fluency (D-KEFS[22])	Participants generate as many designs as possible, according to a series of particular criteria, within one minute.	Total correct responses
Inhibition	Verbal	VIMI[17] - verbal	Participants copy a word said by the experimenter, or provide another word (i.e., inhibit the copying response), depending on instructions.	Total errors
	Nonverbal	VIMI[17] - motor	Participants copy an action demonstrated by the experimenter, or provide another action (i.e., inhibit the copying response), depending on instructions.	Total errors
Planning	Verbal	Sorting (D-KEFS[22])	Participants sort two sets of six cards into two groups of three in as many ways as possible based on verbal features	Total correct verbal sorts
	Nonverbal	Sorting (D-KEFS[22])	Participants sort two sets of six cards into two groups of three in as many ways as possible based on perceptual features	Total correct perceptual sorts
Switching	Verbal	Trail Making Test (D-KEFS[22])	Participants have to draw a line between numbers and letters in sequence, switching between the two (e.g., 1-A-2-B, etc.)	Completion time switching cost
	Nonverbal	Intra/Extra Dimensional Shift (CANTAB[23])	Participants learn a rule through initial trial and error in relation to a shape and then have to switch to a different rule to continue achieving 'correct' answers.	Total errors

Table 3. Descriptive statistics for each EF measure at both time points.

EF Domain	EF measure		**TD** (*n*=17) Mean; *SD* (Range)	**MD** (*n*=17) Mean; *SD* (Range)	**DCD** (*n*=17) Mean; *SD* (Range)
Working Memory *Verbal*	WMTBC Listening Recall *Total Correct*	Time 1	14.24; *3.05* (8-21)	11.12; *3.86* (6-19)	13.88; *3.14* (10-23)
		Time 2	17.53; *4.99* (12-27)	14.35; *3.92* (8-24)	16.24; *4.09* (12-29)
Working Memory *Nonverbal*	Odd-One-Out *Total Correct*	Time 1	11.53; *3.20* (6-17)	6.88; *3.44* (3-14)	7.82; *3.19* (4-15)
		Time 2	13.18; *2.94* (7-18)	8.76; *3.31* (3-17)	9.88; *3.94* (4-16)
Fluency *Verbal*	D-KEFS Verbal Fluency *Total Correct*	Time 1	30.65;*8.08* (15-44)	26.24; *5.98* (16-39)	24.50; *7.79*[a] (3-38)
		Time 2	38.06; *9.46* (17-52)	30.41; *7.94* (18-51)	28.82; *8.83* (12-48)
Fluency *Nonverbal*	D-KEFS Design Fluency *Total Correct*	Time 1	14.76; *4.25* (7-22)	10.35; *4.44* (1-20)	12.12; *3.71* (5-21)
		Time 2	19.65; *5.56* (10-28)	14.24; *3.56* (10-22)	15.12; *4.48* (9-23)
Response Inhibition *Verbal*	VIMI Verbal *Total Errors*	Time 1	9.47; *6.50* (0-23)	12.35; *6.65* (5-29)	16.53; *9.96* (4-36)
		Time 2	8.53; *5.99* (0-24)	12.82; *6.52* (5-28)	14.82; *6.55* (6-27)
Response Inhibition *Nonverbal*	VIMI Motor *Total Errors*	Time 1	28.94; *14.17* (3-51)	43.53; *12.39* (21-61)	48.82; *16.62* (21-74)
		Time 2	26.71; *11.12* (8-48)	40.53; *13.85* (11-64)	43.71; *15.83* (14-71)
Planning *Verbal*	D-KEFS Verbal Sorting *Total Correct*	Time 1	2.24; *.97* (1-4)	2.00; *1.06* (0-3)	2.65; *1.06* (1-4)
		Time 2	2.65; *1.06* (1-4)	2.41; *1.0* (1-4)	2.35; *1.17* (0-4)
Planning *Nonverbal*	D-KEFS Perceptual Sorting *Total Correct*	Time 1	7.12; *1.65* (3-9)	4.41; *2.45* (0-7)	4.47; *2.24* (0-8)
		Time 2	7.47; *1.18* (6-10)	4.88; *2.74* (0-9)	6.06; *1.39* (3-9)
Cognitive Flexibility *Verbal*	D-KEFS Trail Making *Switching cost (sec.)*	Time 1	34.65; *41.16* (-8 – 162)	86.60; *87.09*[b] (-31 – 244)	24.81; *47.75*[c] (-101 – 102)
		Time 2	16.35; 33.*94* (-16 – 128)	22.88; *32.14* (-25 – 84)	9.18; *40.77* (-41 – 121)
Cognitive Flexibility *Nonverbal*	CANTAB IEDS *Total Errors*	Time 1	20.29; *12.90* (8-42)	29.53; *14.92* (8-56)	29.53; *11.59* (8-51)
		Time 2	16.94; *8.98* (7-35)	24.82; *10.76* (9-38)	23.35; *12.61* (9-54)

Note. EF=Executive Function; WMBTC=Working Memory Test Battery for Children; D-KEFS=Delis-Kaplan Executive Function System; VIMI=Verbal Inhibition, Motor Inhibition; CANTAB=Cambridge Neuropsychological Test Automated Battery; IEDS=Intra-/Extra-Dimensional Shift.

[a]1 Missing data point; [b]2 missing data points; [c]1 missing data point.

Table 4. Summary details of step 2 of the hierarchical multiple regression analyses predicting performance in all executive function measures.

EF Domain		Final Model *F*(df) Adj. R^2		*Age*	*IQ*	*TD Vs. MD*	*TD Vs. DCD*	ΔR^2 Step 2
				Details of Step 2 for each regression				
ELWM Verbal	Time 1	*10.47*(4,46) **.43***** *p*<.001	β *Unst.β* *SE*	.48*** *.13* *(.04)* *p=.001*	.37** *.09* *(.03)* *p=.002*	-.13 *-.99* *(1.01)* *p=.33*	-.11 *-.83* *(1.05)* *p=.43*	.01 *p=.56*
	Time 2	*8.24*(4,46) **.37***** *p*<.001	β *Unst.β* *SE*	.57*** *.19* *(.05)* *p<.001*	.42*** *.10* *(.03)* *p=.001*	.02 *.218* *(1.40)* *p=.87*	-.19 *-1.81* *(1.33)* *p=.18*	.03 *p=.31*
ELWM Nonverbal	Time 1	*7.90*(4,46) **.36***** *p*<.001	β *Unst.β* *SE*	.38** *.11* *(.04)* *p=.010*	.13 *.03* *(.03)* *p=.30*	-.42** *-3.37* *(1.14)* *p=.005*	-.57*** *-4.51* *(1.18)* *p<.001*	.22*** *p*=.001
	Time 2	*6.36*(4,46) **.30***** *p*<.001	β *Unst.β* *SE*	.16 *.05* *(.04)* *p*=.27	.36** *.07* *(.03)* *p*=.009	-.34* *-2.74* *(1.27)* *p*=.036	-.35* *-2.81* *(1.21)* *p*=.024	.10* *p*=.035
Fluency Verbal	Time 1	*6.25*(4,45) **.27***** *p*=.001	β *Unst.β* *SE*	.56*** *.53* *(.14)* *p*<.001	.17 *.14* *(.11)* *p*=.178	-.09 *-2.83* *(4.11)* *p*=.560	-.55*** *-2.55* *(4.26)* *p*=.001	.20** *p*=.003
	Time 2	*6.81*(4,46) **.29***** *p*=.001	β *Unst.β* *SE*	.44** *.10* *(.19)* *p*=.003	.22 *.06* *(.12)* *p*=.140	-.14 *-3.2* *(5.84)* *p*=.168	-.54*** *-3.0* *(5.54)* *p*=.001	.19** *p*=.003
Fluency Nonverbal	Time 1	*4.04*(4,46) .20** *p*=.007	β *Unst.β* *SE*	.29 .10 *(.05)* *p*=.085	.16 .05 *(.04)* *p*=.401	-.33* *-3.04* *(1.49)* *p*=.047	-.34* *-3.20* *(1.55)* *p*=.044	.10† *p*=.058
	Time 2	*5.28*(4,46) **.26***** *p*=.001	β *Unst.β* *SE*	.36* *.14* *(.06)* *p*=.018	.12 .03 *(.04)* *p*=.380	-.34* *-3.63* *(1.74)* *p*=.042	-.50** *-5.39* *(1.65)* *p*=.002	.17** *p*=.006
Response Inhibition Verbal	Time 1	*1.66*(4,46) .05 *p*=.175	β *Unst.β* *SE*	-.02 *-.01* *(.11)* *p*=.898	-.01 *-.01* *(.08)* *p*=.965	.16 *2.72* *(3.01)* *p*=.370	.41* *7.15* *(3.06)* *p*=.024	.10 *p*=.076
	Time 2	*2.96*(4,46) .14* *p*=.029	β *Unst.β* *SE*	-.22 *-.11* *(.08)* *p*=.165	-.16 *-.06* *(.05)* *p*=.265	.16 *2.24* *(2.48)* *p*=.373	.46** *6.54* *(2.34)* *p*=.008	.14* *p*=.027
Response Inhibition Nonverbal	Time 1	*4.60*(4,46) **.22**** *p*=.003	β *Unst.β* *SE*	-.14 *-.18* *(.19)* *p*=.365	-.08 *-.09* *(.15)* *p*=.547	.35* *12.04* *(5.46)* *p*=.032	.59*** *20.59* *(5.56)* *p*=.001	.22** *p*=.002

	Time 2	*4.86*(4,46) **.24**** *p*=.002	β *Unst.β* *SE*	-.29† *-.34* *(.17)* *p*=.055	-.09 *-.07* *(.11)* *p*=.515	.29 *9.52* *(5.30)* *p*=.079	.59*** *19.05* *(5.01)* *p*<.001	.22** *p*=.002
Planning Verbal	Time 1	*2.04*(4,46) .08 *p*=.104	β *Unst.β* *SE*	.22 *.02* *(.01)* *p*=.194	.21 *.02* *(.01)* *p*=.150	.04 *.08* *(.38)* *p*=.824	.18 *.39* *(.39)* *p*=.321	.02 *p*=.596
	Time 2	*.82*(4,46) -.02 *p*=.525	β *Unst.β* *SE*	-.21 *-.02* *(.01)* *p*=.221	-.18 *-.01* *(.01)* *p*=.267	.25 *-.56* *(.42)* *p*=.189	.12 *-.27* *(.42)* *p*=.498	.04 *p*=.414
Planning Nonverbal	Time 1	*7.79*(4,46) **.35***** *p*<.001	β *Unst.β* *SE*	.11 *.02* *(.03)* *p*=.441	.37** .06 *(.02)* *p*=.005	-.36* *-1.84* *(.74)* *p*=.017	-.44** *-2.27* *(.76)* *p*=.005	.14** *p*=.007
	Time 2	*13.84*(4,46) **.51***** *p*<.001	β *Unst.β* *SE*	.34** *.06* *(.02)* *p*=.006	.54*** .06 *(.01)* *p*<001.	-.23 *-1.02* *(.59)* *p*=.094	-.25† *-1.13* *(.56)* *p*=.051	.05 *p*=.094
Cognitive Flexibility Verbal	Time 1	*4.15*(4,43) .22** *p*=.006	β *Unst.β* *SE*	-.18 *-.90* *(.77)* *p*=.249	-.29* -1.32 *(.62)* *p*=.039	.22 *31.02* *(22.25)* *p*=.170	-.08 *-11.59* *(22.52)* *p*=.610	.05 *p*=.216
	Time 2	1.48(4,46) .04 *p*=.223	β *Unst.β* *SE*	-.27 *-.71* *(.44)* *p*=.115	-.24 -.44 *(.28)* *p*=.123	-.10 *-7.66* *(13.69)* *p*=.579	-.09 *-6.40* *(13.03)* *p*=.625	.01 *p*=.822
Cognitive Flexibility Nonverbal	Time 1	*8.84*(4,46) **.39***** *p*<.001	β *Unst.β* *SE*	-.45** *-.47* *(.14)* *p*=.002	-.40** *-.37* *(.11)* *p*=.002	.03 *.83* *(4.02)* *p*=.836	.34* *9.85* *(4.09)* *p*=.020	.08* *p*=.048
	Time 2	*7.10*(4,46) **.33***** *p*<.001	β *Unst.β* *SE*	-.63*** *-.53* *(.12)* *p*<.001	-.17 *-.10* *(.06)* *p*=.194	.06 *1.49* *(3.61)* *p*=.682	.42** *9.85* *(3.43)* *p*=.006	.12* *p*=.016

Note. For each regression the final model *F* values, degrees of freedom in parentheses, and adjusted R^2 are presented, along with the change in R^2 in Step 2 of the model. Standardized beta values, *unstandardized coefficients*, and *standard errors* (in parentheses) are reported for each predictor variable. Significant final regression models after Bonferroni corrections ($p \leq .005$) are indicated in boldface. ELWM: executive-loaded working memory. 1 missing data point for verbal fluency measures at Time 1 (DCD group). 3 missing data points for verbal cognitive flexibility measures at Time 1 (2 MD, 1 DCD).

*$p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; † $p \leq .06$ non-significant trend.

SUPPLEMENTARY MATERIALS

Recruitment procedures and participants. Participants with a diagnosis of Developmental Coordination Disorder (DCD) were recruited for the original study[1] through an advert placed with a charitable organisation, requesting children aged 7-11 with a diagnosis of DCD/dyspraxia to participate in research. Parents volunteered for the study by emailing the research team to receive more information, and to check eligibility. Children with a co-occurring diagnosis of autism spectrum disorder or attention deficit-hyperactivity disorder were excluded from participating due to the potential problems in executive functioning associated with these disorders. Reading and language difficulties, as well as intellectual disability, were assessed through standardised tests (see Materials), and any child demonstrating performance outside of the typical range on these measures was also excluded. The DCD diagnosis was corroborated by the research team using standardised measures and parent report (see Materials). The DCD group in the original study[1] consisted of 23 children (16 males; mean age: 10.0 years, SD: 1.1 years, range: 8.1–11.9). Of these 23 children, 19 agreed to participate in the follow-up study and were re-assessed to ensure that they continued to meet inclusion criteria for the DCD group, and that their diagnosis was stable across time points. Two children scored more than two standard deviations below the mean on the test of intellectual ability. These two children were excluded from the sample (see Table S1 for inclusion/exclusion criteria), because one of the criteria for a DCD diagnosis is that motor deficits are not better explained by intellectual disability (hence the diagnosis could not be corroborated), and because low intellectual ability was likely to impact on their ability to understand task instructions and rules. The final DCD group for the follow-up study consisted of 17 children (11 males; mean age at Time 2: 12.0 years, SD: 1.2 years, range: 10.1 – 14.1).

Children without a diagnosis of DCD were recruited through local schools: parents of 250 children aged 7-11 received information sheets about the study, and volunteered to take part by returning a signed consent form to the research team through the class teacher. Children who did not have any reported medical condition or neurodevelopmental disorder were assessed on the standardised assessments to ensure they met inclusion criteria (see Table S1). Children were included in the typically developing control (TD) group if they scored at or above the 25th percentile on the standardised motor assessment, had no parent-reported motor difficulties, and scored in the typical range on the standardised measures of reading, language and intellectual abilities. Children were identified as having motor difficulties (MD) if they scored at or below the 16th percentile on the standardised motor assessment, but scored in the typical range on the other standardised measures. The original sample[1] included 38 children in the TD group (17 males; mean age: 9.3 years, SD: 1.0 years, range: 7.2–11.1), and 30 children in the MD group (17 males, mean age: 8.9 years, SD: 1.2 years, range: 7.1–11.3). Of these 68 children, 37 were available for follow-up and were re-assessed to ensure they continued to meet inclusion criteria for their assigned group. One TD child performed on the 16th percentile of the MABC-2 and two TD children performed on the 9th percentile. As these children demonstrated some degree of motor difficulty at Time 2 they could no longer be included in the TD group and were therefore excluded from the sample. All children in the MD group continued to meet criteria for group membership, demonstrating persistent motor difficulties across the two time points. The final TD group consisted of 17 children (6 males; mean age at Time 2: 11.3 years, SD: 1.0 years, range: 9.7 – 13.1). The final MD group consisted of 17 children (8 males; mean age at Time 2: 10.5 years, SD: 0.6 years, range: 9.8 – 12.3).

---*Table S1 here*---

Materials. As outlined above, participants were assessed on several standardised measures to confirm their eligibility for the study. These tests are described first, followed by the executive functioning battery.

Movement Assessment Battery for Children (MABC-2) and Checklist. The MABC-2[2] is a standardised assessment of motor ability, comprising three components: manual dexterity, aiming and catching, and balance. Scores for each component can be summed to provide a total standard score (*M*=10, *SD*=3) and percentile ranks, based on UK norms. Children performing at or below the 16th percentile can be identified as having some motor difficulties. Test-retest reliability is reported as .80 for the total sum of the three component scores[2].

The MABC-2 Checklist[2] consists of 30 statements requiring parents to judge their child's level of motor competence in tasks involving movement in a static and/or predictable environment and in a dynamic and/or unpredictable environment, in comparison to other children of the same age. The Checklist is used to assess the impact of motor difficulties on daily life[3], which is central to the diagnostic criteria for DCD. Parents respond to the statements deciding how their child deals with the tasks on a scale from "Very well" to "Not close" (scoring 0–3 points), and a Total Score is calculated. These rating are summed to calculate a total score, which is mapped on three percentile bands, with scores below the 15th percentile representing a risk of motor difficulties and scores below the 5th percentile being indicative of motor difficulties affecting daily living. Test-retest reliability ranged between .77 to .91 in studies using the previous edition of the M-ABC[4], the content of which is highly overlapping with the more recent version.

British Abilities Scales (BAS-3). The BAS-3[5] is a standardised measure of intellectual abilities, comprising both verbal and nonverbal subtests. It was used to ensure that all children were functioning at an appropriate level in order to understand the instructions of the tasks, and to confirm that those in the DCD group did not have an intellectual disability. The Verbal Similarities and Word Definitions subtests were used to measure verbal reasoning, with the Matrices subtest used as a measure of nonverbal reasoning. Scores for each subtest were summed and converted to standard (*T*) scores, with the Matrices *T*-score first doubled to ensure that verbal and nonverbal abilities were equally weighted in the final score (as outlined in the BAS-3 manual). The average of the *T*-scores from the verbal subtests and the doubled nonverbal subtest was calculated and converted into a standard score (General Conceptual Ability [GCA]; *M*=100, *SD*=15). Children in all three groups were required to have a GCA score within two standard deviations of the mean (i.e., at or above 70) at both time points in order to be included in the study. Test-retest reliability is reported as .73 for the Matrices subtest, as .86 for the Word Definition subtest and .79 for the Verbal Similarities subtest[4].

Clinical Evaluation of Language Fundamentals 4th Edition (CELF-4-UK). The CELF-4-UK[6], a widely used assessment of receptive and expressive language abilities, was administered to ensure that children did not perform poorly on the verbal executive function measures due to problems with language skills[7], and to exclude children with very low scores indicative of language disorder. Those with scaled scores at or below two SD from the mean (of four or less; *M*=10, *SD*=3) on two core subtests, Formulated Sentences (expressive language), and Word Classes-Receptive (receptive language), were excluded. This ensured that children with clear evidence of language disorder did not take part in the study, and that the cut-off harmonised with that used for other study tests (i.e., 2 SD from the mean). Test-retest reliability for relevant ages ranged from .74 to .79 for the Formulated Sentences subtest, and from .83 to .91 for the Word Classes-Receptive subtest[6].

Test of Word Reading Efficiency (TOWRE). The TOWRE[8] was used to assess reading of words and non-words, to ensure that children did not have any reading problems indicative of dyslexia, a disorder that may affect performance on executive functioning tasks[9]. Children were timed when reading a list of words, followed by a list of non-words, and the total number of words read correctly within the time limit of 45 seconds was calculated. Total scores were converted to a standard score (M=100, SD=15). Children in all three groups were required to have a Total Standard Score within two standard deviations of the mean (i.e. above 70) in order to be included in the study. Test-retest reliability ranged from .82 to .97 for 6 to 9 year-old children[8].

Executive functioning battery. A verbal and a nonverbal test was completed for each of the following executive functions: executive-loaded working memory (ELWM); fluency; inhibition; planning; and cognitive flexibility / switching. A summary of the tasks is provided in Table 2 within the current paper.

For *verbal ELWM*, the Listening Recall test from the Working Memory Battery for Children[10] was completed. Sentences were presented to participants in blocks of six trials, beginning with a block of one-sentence trials, with an increasing number of sentences per trial in each subsequent block. Participants were asked to judge whether the sentence was true or false, and then to hold the last word in memory while providing judgements on the next sentences in the trial. At the end of each trial, children were asked to recall the last words of each sentence in order. The test was ended when three out of six trials within a block were incorrect. Total number of trials correct was scored rather than span, as this has been reported to be a more reliable measure of verbal working memory[11]. Test–retest reliabilities of .38–.83 are reported for relevant ages[10]. For *nonverbal ELWM*, an equivalent test of visuospatial ELWM was adopted from previous research, called the ‘Odd-One-Out’ test[12]. On each trial, the child was presented with a card depicting a set of three simple nonsense diagrams and asked to point to the ‘odd one out’. Participants were asked to hold the spatial location of the odd-one-out in memory while they provided judgements on the next set of diagrams in the trial. Sets of diagrams were presented in blocks of three, beginning with a block of one-set trials, with an increasing number of sets per trial in each subsequent block. At the end of each trial, children were asked to recall the spatial location of the odd-one-out for each card by pointing to the relevant location on a blank grid. The test was ended when two out of three trials within a block were incorrect. Total number of trials correct was scored. The span version of this task has a reliability of .80[12].

To assess fluency, the Delis-Kaplan Executive Functioning System (D-KEFS[13]) was used. For *verbal fluency*, children were required to generate as many words as possible within one minute that belonged to a specific category (i.e., animals and boys’ names). Total correct words (without repetitions) summed from the two categories was used as the measure of verbal fluency. Test-retest reliability is reported as .70 for category fluency[13]. For *nonverbal fluency* (‘Design Fluency’), children were provided with a grid in which there were either a number of filled dots (condition one), or a mixture of filled and empty dots (condition two), presented in each square of the grid. Children were required to use four connected straight lines to draw as many different designs as possible within one minute. In condition two, children were only allowed to connect the empty dots. Nonverbal fluency was calculated using the total correct designs (i.e. those following the rules) across the two conditions. Test-retest reliabilities are reported as .66 for filled dots and .43 for empty dots[13].

To assess inhibition, a test was adopted from previous research[7] called the Verbal Inhibition, Motor Inhibition (VIMI) test. For *verbal inhibition*, children were required to repeat words said by the experimenter (i.e., either ‘doll’ or ‘car’), which were presented in a pseudo-random order for 20 trials (‘copy’ block). For the next block of 20 trials (‘inhibit block’), participants were required to inhibit this copying response by responding with the opposite word (i.e., 'car' was the response to 'doll', and vice versa). The copy and inhibit blocks were then repeated once with the same words (Part A), followed by a set of four blocks following the same pattern but using different words ('bus' and 'drum'; Part B). Total number of errors across the full task provided the measure of verbal inhibition. Cronbach's alpha, based on total error scores, was .73[7]. For *nonverbal inhibition*, the test followed an identical format but used hand actions instead of words. Participants were required to copy the experimenter in presenting a pointed finger or a fist (Part A), or a flat horizontal hand or flat vertical hand (Part B). In the 'inhibit' blocks, participants again had to present the opposite hand action to the experimenter. Total number of errors across the full task provided the measure of nonverbal inhibition, and Cronbach's alpha for these error scores was .92[7].

To assess planning, the D-KEFS Sorting task[13] was used. Participants were presented with two sets of six cards and asked to sort them into two groups of three in as many different ways as they could. Categories could be created based on the words presented on the cards (*verbal planning*), or on the perceptual properties of the cards (*nonverbal planning*). There were three possible verbal sorts (e.g., transports *vs.* animals, things that fly *vs*. things that move along the ground) and five possible nonverbal sorts (e.g., small cards *vs.* large cards, straight edges *vs.* curved edges) in each card set. Total numbers of correct sorts were used as the measures of verbal and nonverbal planning, respectively. Test-retest reliability for the Sorting task is reported as .49[13].

To assess cognitive flexibility, two tasks were adopted from standardised batteries of executive functioning measures. For *verbal cognitive flexibility*, the D-KEFS Trail Making Test[13] was used. In the number-letter switching task, participants were required to connect letters and numbers in an alternating sequence (i.e., 1-A-2-B-3-C, etc., until 16-P) as quickly as possible. In order to ensure that reduced performance on this task was not caused by difficulties with sequencing numbers or letters, or due to motor speed or visual scanning abilities, component skills were also assessed. In the motor speed task, children were required to follow a line with their pencil between dots placed around the page (as in a 'dot-to-dot' game) as quickly as they could, thus removing any of the verbal element from the task. In the visual scanning task, children were asked to find all the number 3s on the page and cross them off as quickly as possible. The number sequencing task involved connecting the numbers from 1-16, and the letter sequencing task required connecting the letters from A-P. The measure of verbal cognitive flexibility was the total time for the number-letter switching task minus the total time for the number and letter sequencing tasks (i.e., 'switching cost'). Test-retest reliabilities for the component tasks are reported as .77 (number sequencing), .57 (letter sequencing) and .22 (letter-number switching)[13]. The fact that switching measures depend on difference scores can make reliability of these tasks somewhat low, but this is an inherent problem with these measures[14]. For *nonverbal cognitive flexibility*, the Intra-Extra Dimensional Set Shift test from the Cambridge Neuropsychological Test Automated Battery (CANTAB)[15] was used. Participants were first presented with two coloured shapes and asked to work out the rule by touching one of the two shapes on the screen and finding out whether they were 'correct'. Feedback was provided by the computer program, and participants were told that if they had found the correct shape, they should continue to touch this shape on subsequent trials until the rule changed (i.e., until they received feedback that their response was 'incorrect'). At this point children would need to switch rule, and choose the other shape instead. In the second part of the task, a white line was added to the stimuli, either adjacent to or overlaying the coloured shape, but the child continued to attend to the coloured shape to obtain correct responses ('intra-dimensional shift'). In the final part of the task, the rule changed again and children had to attend to the white line in order to obtain correct responses ('extra-dimensional shift'), ignoring the coloured shape to which they had previously been attending. Total number of errors across the task was used as the measure of nonverbal cognitive flexibility. Test-retest reliability for total errors is reported as .40[15].

Statistical Analyses. Statistical checks in each regression (e.g. Durbin-Watson, variance inflation factor statistics, standardised residuals, Cook's/Mahalanobis distances) revealed no evidence of multicollinearity and no outliers or influential cases[16].

Supplementary References

[1]Leonard HC, Bernardi M, Hill EL, Henry LA. Executive functioning, motor difficulties, and developmental coordination disorder. *Dev Neuropsychol* 2015; 40(4):201-15.

[2]Henderson SE, Sugden DA, Barnett AL. Movement assessment battery for children (Movement ABC-2) (2nd ed.). London: The Psychological Corporation, 2007.

[3]Schoemaker MM, Niemeijer AS, Flapper BC, & Smits-Engelsman BCM. Validity and reliability of the Movement Assessment Battery for Children-2 Checklist for children with and without motor impairments. *Developmental Medicine & Child Neurology* 2012; 54(4): 368-375.

[4]Henderson SE, Sugden, DA. The Movement Assessment Battery for Children (Movement ABC). London: The Psychological Corporation, 1992.

[5]Elliot CD, Smith P. British Ability Scales – Third Edition (BAS-3). London: GL Assessment, 2011.

[6]Semel E Wiig EH, Secord WA. Clinical Evaluation of Language Fundamentals CELF-4 UK (4th ed.). London, UK: Psychological Corporation, 2006.

[7]Henry LA, Messer DJ, Nash G. Executive functioning in children with specific language impairment. *J Child Psychol Psychiatry* 2012; 53(1):37-45.

[8]Torgensen J, Wagner R, Rashotte C. Test of Word Reading Efficiency (TOWRE). Austin, TX: Pro-Ed, 1999.

[9]Booth JN, Boyle JME, Kelly, SW. Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis. *Brit J Dev Psychol* 2010, *28*, 133-176.

[10]Pickering S, Gathercole SE. Working memory test battery for children (WMTB-C). London: The Psychological Corporation, 2001.

[11]Ferguson AN, Bowey JA, Tilley A. The association between auditory memory span and speech rate in children from kindergarten to sixth grade. *J Exp Child Psychol*, 2002; 81: 141–156.

[12]Henry LA. How does the severity of a learning disability affect working memory performance? *Memory* 2001; 9(4-6):233-47.

[13]Delis DC, Kaplan E, Kramer JH. Delis–Kaplan Executive Function System (D-KEFS®). London: Psychological Corporation, 2001.

[14]Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex ‘‘frontal lobe’’ tasks: A latent variable analysis. *Cogn Psychol*, 2000; 41: 49–100.

[15]Cambridge Cognition. The Cambridge Neuropsychological Test Automated Battery (CANTAB®): Test–retest reliability characteristics. Cambridge: Author, 2006

[16]Field AS. Discovering Statistics using IBM SPSS Statistics (4th ed.). London: Sage Publications Ltd, 2013.

Table S1
Inclusion criteria for group membership at Time 1 and Time 2

Inclusion Measure	TD group	MD group	DCD group
Movement Assessment Battery for Children (MABC-2) and Checklist	MABC-2 Total score ≥ 25^{th} %, Checklist > 15^{th} %	MABC-2 Total score ≤ 16^{th} %	MABC-2 Total score ≤ 16^{th} %, Checklist < 5^{th} %
British Abilities Scales (BAS-3)	Standard score ≥ 70	Standard score ≥ 70	Standard score ≥ 70
Clinical Evaluation of Language Fundamental (CELF-4-UK)	Scaled score ≥ 4 on Formulated Sentences and Word Classes-Receptive subtests	Scaled score ≥ 4 on Formulated Sentences and Word Classes-Receptive subtests	Scaled score ≥ 4 on Formulated Sentences and Word Classes-Receptive subtests
Test of Word Reading Efficiency (TOWRE)	Standard score ≥ 70	Standard score ≥ 70	Standard score ≥ 70
Parent reports of clinical diagnosis	No clinical diagnosis	No clinical diagnosis	Diagnosis of DCD only