

Development of Reasoning Style in ADHD: A Cross-Sectional Gaze Movement Study with Children and Adolescents

Jana Bretthauer

University of Cologne

Julia Schmidgen

University of Cologne

Julia Hertel

University of Cologne

Daniela Canu

University of Freiburg

Ulf Thiemann

LVR Hospital Bonn

Monica Biscaldi

University of Freiburg

Nikolaos Smyrnis

University General Hospital Attikon

Stephan Bender

University of Cologne

André Beauducel

University of Bonn

Christoph Klein

`christoph.klein@uk-koeln.de`

University of Cologne

Research Article

Keywords: ADHD, Raven Matrices, Intelligence, Eye Movements

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11219099/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Although Raven's Standard Progressive Matrices (RSPM) is an established measure of fluid intelligence, a core construct of cognitive development, little is known about how children and adolescents process these matrices towards solutions. Eye movement analyses map such processing in space and time. Regarding RPSM, two different *processing styles* have been established: *constructive matching* and *answer elimination*. Existing findings largely pertain to early primary school age and adulthood, but it remains unclear whether processing style in ADHD differs from that of typically-developing peers.

In the present study, N = 160 children and adolescents aged 9–17 (78 ADHD, 82 controls) took a nine-item RSPM version while eye-movements were recorded. Various eye-movement measures and age, intelligence, and pupil size were analysed with ANCOVA. With increasing difficulty, gaze patterns shifted towards constructive matching (CM; i.e., preferred exploration of the problem matrix). Age and intelligence both raised the level of the spatial aspects of CM. Participants with ADHD showed significantly reduced intelligence alongside significantly reduced CM. Statistically controlling for intelligence annihilated the significant group differences in CM. Adjunct factor analysis reduced the eye-movement measures to two correlated dimensions, spatial and temporal, consistently across groups.

Results suggest that RSPM processing style depends on age and intelligence. CM is reduced in ADHD if their IQ is reduced, plausibly reflecting shared underlying processes and direct influences of ability on processing style.

1. Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental condition that typically manifests in childhood and frequently continues into adult life [1]. Much of recent work on ADHD has turned from its behavioural symptoms to cognitive mechanisms that underlie them [2]. Rather than a uniform intellectual weakness, ADHD is better described by a differentiated profile in which some cognitive domains are more affected than others [3, 4]: meta-analyses report modestly lower average IQ scores than in controls [3, 5], and impaired executive functions, which become more pronounced as cognitive load increases [3, 6–8]. Such difficulties concern not only which operations can be performed, but also how stable this performance is [6, 9].

Accordingly, the way how individuals with ADHD achieve solutions, that is their *processing style (PS)*, has become a topic of interest. A task well suited to investigate PS is the Raven Standard Progressive Matrices (RSPM) [10] in combination with eye-tracking. Here, each item presents a three-by-three matrix of figural patterns with one cell left blank in the upper half and several answer alternatives in the lower part. The solver must infer the governing rules of the matrix to choose the missing pattern from the answer alternatives [11]. Two routes to a solution are commonly distinguished [11, 12]. With *constructive matching (CM)*, the figural patterns in the matrix are analysed and rules that govern it are inferred, leading to hypotheses regarding matching elements within the answer options [13, 14]. With *answer elimination (AE)*, subjects pick answer candidates from the answer options and check their fit within the

matrix [13, 14]. CM leans heavily on visuospatial working memory and sustained attentional control to mentally represent the increasing number of relations in progressively complex problems [11, 15]. Required mental resources provided, CM is the more efficient procedure to solve the matrices [11, 13].

The two strategies lead to distinct gaze patterns that can be captured using eye tracking, offering a direct window into the solution process [13, 16]. CM is reflected in prolonged and systematic scrutiny of the matrix before the options are consulted; AE in more frequent movement between matrix and answers [13, 14]. Such gazing behaviour entails both spatial and temporal aspects, i.e. spatial distribution of gaze between matrix and answers and temporal frequency and timing of attentional shifts. Whether these aspects reflect a single underlying processing style has not been examined.

Expectedly, strategy use undergoes profound developmental changes in neurotypical children and adolescents, showing strong increases in CM that mediate over half of the age-related variance in reasoning performance as well as adaptation of response times and strategy use to item difficulty with increasing age [17]. Eye-tracking studies have corroborated this trajectory at finer resolution for ages 6–9ys [18], and graded signs of adaptive strategy use for ages 7–8ys [19]. Whether such strategy development differs in children and adolescents with ADHD, is so far unknown.

Eye-tracking research in ADHD has documented altered gaze behaviour, including atypical fixation patterns and greater variability in how attention is allocated [20]. Overall, abnormalities are evident in a broad range of functions including saccade inhibition, initiation of antisaccades, spatial memory [21]; longer and more intra-individually variable saccadic reaction times [22] as well as evidence of differential developmental processes such as developmental lags and slowed development in ADHD participants aged 7–15ys [23]. Whether individuals with ADHD differ in terms of constructive matching and answer elimination was only recently addressed for the first time in adults [24].

Importantly, differences in performance cannot be explained by cognitive abilities alone; energetic factors must be considered also. The cognitive-energetic model [25, 26] expands upon cognitive processing steps by adding an energetic level with three components, including “effort”, defined as the energy required to meet task demands.

Pupil size provides a physiological index of mental effort mobilized during cognitive tasks [27, 28], with increasing memory load, both in children and in adults [29, 30]. While gaze movements and pupil size are increasingly examined together in reasoning research [31, 32], their relationship during the solving of matrix tasks has rarely been examined.

The present study builds upon [24], examining CM in children and adolescents with ADHD, deriving the following hypotheses and research questions. *First*, the CM pattern should become more pronounced with increasing age. *Second*, psychometric intelligence should make an age-independent contribution to CM. *Third*, we expected differences between children and adolescents with and without ADHD in CM to diminish with increasing age. *Fourth*, we determined the factorial structure of our seven gaze movements measures using exploratory factor analysis. *Fifth*, we examined whether pupil size provides

an independent contribution to CM processing style. Trial difficulty effects will be reported to show that our experimental set up functioned in the expected way.

2. Methods

2.1 Participants

The study protocol was approved by the Ethics Committee of the University of Freiburg (EK 124/17). Participants gave written informed consent; for minors, consent was additionally obtained from their legal guardians. Data were collected within a larger test battery [33] that also included a measure of intelligence, the Culture Fair Test (CFT-20-R; [34]) alongside the short form of Raven's Standard Progressive Matrices (RSPM; [10, 35]).

The final sample comprised 78 participants with ADHD (26 female) and 82 typically developing participants (42 female), aged 9-17ys (ADHD: $M = 13.08 \pm 2.53$; TD: $M = 13.18 \pm 2.37$), with CFT-assessed IQ ranging from 70 to 130 (ADHD: $M = 101.6 \pm 10.47$; TD: $M = 111.16 \pm 11.74$). The groups were comparable in age ($t(158) = 0.248$, $p = .805$), but differed in gender distribution ($\chi^2(1) = 5.233$, $p = .022$), with a higher proportion of male participants in the ADHD group. CFT-20-R scores ($t(158) = 5.423$, $p < .001$) and RSPM scores ($t(158) = 1.830$, $p = .069$) were lower in the ADHD group. All participants had normal or corrected-to-normal vision and reported no history of epilepsy or other neurological disorder.

ADHD participants were recruited through the departmental outpatient services and carried a confirmed clinical diagnosis, based on standard clinical practice combining rating scales (such as the German versions of the Conners' Parent and Teacher Rating Scales), structured parent-child interviews, and behavioural observation. Controls were recruited from the departmental database and through public advertisements; they were screened for ADHD symptoms with the Conners' Self- and Parent Rating Scales and were not enrolled if they or a first-degree relative had a psychiatric or neurological condition.

Various comorbid conditions served as exclusion criteria (see supplementary materials). Participants prescribed methylphenidate suspended medication 24hs before testing. Participants were excluded if their CFT-assessed IQ fell outside the range of 70 to 130 (TD: $n = 32$, ADHD: $n = 4$) or if insufficient Raven data were available (< 2 trials per difficulty level or < 8 trials overall; TD: $n = 5$, ADHD: $n = 8$).

2.2 Apparatus

Binocular gaze was recorded at 1,000Hz (spatial resolution: 0.01° , accuracy: $0.25\text{--}0.5^\circ$ of visual angle) with a desktop-mounted EyeLink 1000Plus (SR Research, Canada). Stimuli were presented with EyeLink Experiment Builder (Version 2.1.140) on a 24-inch LCD-monitor ($1,920 \times 1,080\text{px}$). Participants sat 90cm away from the monitor and 60cm from the tracker, with the head supported by a chin rest.

2.3 Procedure

All sessions took place in a dimly lit cabin with constant illumination (70–80lx). A 13-point calibration with subsequent validation was carried out at the start of each session to ascertain accuracy within 1° for both eyes. Drift correction preceded each trial (deviation < 0.5°). One experimenter gave the instructions inside the cabin, a second monitored the recording from an adjacent room.

The nine-item RSPM short form [10], including items 11, 24, 28, 36, 43, 48, 49, 53, and 55 of the original version reported to correlate by $r=.98$ with the full test [35] was presented following three practice trials. Each trial started with a fixation cross ($1.6^\circ \times 1^\circ$) presented for 1,000ms before the matrix appeared; trials had no time limit and ended once a response was entered on the keyboard. Participants were reimbursed with cinema or book vouchers (€7.50/hour).

2.4 Data Analysis

Eye-movement recordings were processed in EyeLink Data-Viewer (Version 3.1.97; SR Research) with the default event-detection algorithm, under which all non-saccadic, non-blink segments count as fixations. For every trial, two areas of interest (AOIs) were defined: the upper field matrix pattern and the lower field response alternatives. Seven eye movement measures were derived: Four *spatial* measures quantify the proportions of dwell time and fixations on the matrix and the answers, while three *temporal* measures quantify the switch rate (matrix–answer transitions per second), time to first fixation on the answers, and the total trial duration. The nine items were assigned to three difficulty levels on the basis of their empirical, present-sample solution frequency rather than on the original ordering [10, 35]. The proportion of trial time without pupil signal did not differ between groups (TD: $M = 4.90\% \pm 3.34\%$; ADHD: $M = 4.77\% \pm 4.06\%$; $t(158) = 0.21$, $p=.833$).

Pupillary measure. For each trial, a 500ms baseline was defined and aggregated into a person-level mean across trials that was independent of item difficulty, trends, and group status (see **Table S8**), and showed high odd-even reliability (Spearman–Brown-corrected: $r=.93$).

Covariates. Three person-level covariates were entered, each z-standardized across the full sample: (a) AGE in months; (b) age-residualized INTELLIGENCE raw scores (correlation of age and intelligence raw scores: $r=.76$) [36]; (c) PUPIL: baseline-corrected pupil size (see below). For each covariate, its interaction with TRIAL DIFFICULTY was additionally modelled. To retain the full sample, missing CFT scores (6 TD, 7 ADHD) were estimated from age and Raven scores ($R^2=.50$, $p<.001$); sensitivity analyses excluding these participants ($N = 147$) and adding gender as a covariate left all Holm-corrected effects unchanged (**Tables S6** and **S7**).

Statistics. Eye movements measures were analysed with univariate repeated-measures analysis of covariance (ANCOVA; GLM; Type III SS) in SPSS (Version 29; IBM Corp., Armonk, NY), with GROUP (ADHD vs. TD) as between-subjects factor and TRIAL DIFFICULTY as within-subjects factor with three levels (easy, medium, difficult). Greenhouse–Geisser–corrected degrees of freedom are reported whenever sphericity was violated (Mauchley). The aforementioned co-variables were part of all models

except INTELLIGENCE, which was omitted in additional ANCOVAs to determine the effects of reduced intelligence on the eye-movement measures.

The significance level was set at $\alpha = .05$, and partial eta squared (η^2) served as the effect-size measure. To save space, test statistics are reported in the main text only for effects reaching $F = 2.50$, which corresponds to $\eta^2 \approx .016$ and lies below the critical value for significance. Complete results are documented in **Tables S2** and **S3**.

Whenever *homogeneity* of variance was violated (Levene), all ANCOVAs were computed on Blom rank-normalized scores [37, 38] as the primary analyses. This normalization turned no non-significant result into a significant one. The main effect of TRIAL DIFFICULTY and its polynomial trends are reported from the untransformed model, all other effects from the Blom-normalized model.

Holm–Bonferroni corrections were applied separately within each set of effects; unless stated otherwise, all significant effects survived correction within their set.

Factor analysis

Exploratory factor analyses (principal axis factoring, promax rotation, Kaiser criterion, sampling adequacy assessed by the Kaiser-Meyer-Olkin criterion, KMO) were computed separately in each group on scores averaged across difficulty levels. Regression-method factor scores were entered simultaneously into multiple regressions predicting age-residualized CFT scores.

3. Results

Trial difficulty affected all seven eye-movement measures (Fig. 1; **Table S1**), but in two distinct ways. Regarding *spatial variables*, with harder items, more dwell time and fixations was spent on the matrix and less on the answer alternatives (all $F(2,304) \geq 171.12$, $p < .001$, $\eta^2 = .53-.60$; **Table S2**). Most of these changes occurred between easy and medium items and then levelled off, with both polynomial contrasts significant for all four measures (linear: $F(1,152) \geq 302.32$, $\eta^2 = .67-.73$; quadratic: $F(1,152) = 47.09-99.01$, $\eta^2 = .24-.39$; all $p < .001$). Regarding the *temporal variables*, switch rate decreased ($M = .33/.30/.24$ runs per second; $F(2,304) = 92.24$, $p < .001$, $\eta^2 = .38$), while time to first fixation (5,881 to 8,790ms; $F(1.91,290.79) = 37.61$, $p < .001$, $\eta^2 = .20$) and trial duration (14,082 – 25,127ms; $F(1.79,271.41) = 89.63$, $p < .001$, $\eta^2 = .37$) increased, with all changes being even across difficulty levels (linear: $F(1,152) = 66.59-201.87$, $\eta^2 = .31-.57$, all $p < .001$; quadratic: $F \leq 13.60$).

Increasing difficulty increased CM and processing time, indicating that our task elicited the intended variation in processing style and hence “worked” as required.

3.1 AGE

With increasing age, CM processing style and time to first fixation increased as well (age effects: spatial CM: $F(1,152) \geq 56.68$, $p < .001$, $\eta^2 = .27-.29$; time to first fixation: $F(1,152) = 9.39$, $p = .003$, $\eta^2 = .058$).

On the other two time-based measures the effect of AGE depended on TRIAL DIFFICULTY. For run rate ($F(2,304) = 16.36, p < .001, \eta^2 = .097$), the relationship reversed across levels: at easy items older participants switched more often than younger ones, at medium items age made almost no difference, and at difficult items older participants switched less often (linear contrast $F(1,152) = 30.12, p < .001, \eta^2 = .165$). Processing time turned the same way ($F(2,304) = 7.21, p < .001, \eta^2 = .045$; linear contrast: $F(1,152) = 15.92, p < .001, \eta^2 = .095$): older participants were faster at easy items and slower at difficult ones. For time to first fixation, the pattern was weaker and did not reverse ($F(2,304) = 3.13, p = .045, \eta^2 = .020$); this effect, however, did not survive Holm-Bonferroni correction.

Residual correlations analyses showed these effects to occur at the transition from medium to difficult items both groups (see supplementary materials).

Confirming hypothesis 1, age raised the overall level of the spatial measures but modulated the course of the temporal ones.

3.2 INTELLIGENCE

Participants with higher age-residualized CFT scores spent more dwell time and fixations on the matrix and less on the alternatives (all $F(1,152) \geq 15.82, p < .001, \eta^2 = .09-.12$), independently of group membership. Averaged across difficulty levels, CFT scores were unrelated to all three time-based measures (**Table S2**). The interaction between intelligence and trial difficulty on run rate ($F(2,304) = 3.19, p = .043, \eta^2 = .021$) did not survive Holm-Bonferroni correction; on the remaining six measures it was non-significant (**Table S3**).

Partially confirming hypothesis 2, higher intelligence was associated with more CM processing style regarding its spatial, but not its temporal aspects.

3.3 GROUP

If intelligence was statistically controlled for, TD and ADHD did not differ in how they allocated their gaze on most measures except run rate: averaged across difficulty levels, participants with ADHD switched more frequently than controls ($F(1,152) = 5.74, p = .018, \eta^2 = .036$; .31 versus .27 alternations per second), by about the same margin at every difficulty level. This effect, however, did not survive Holm-Bonferroni correction ($\alpha = .0071$). Groups, furthermore, did not differ in how they adapted to increasing task difficulty or interactions with age, CFT score or pupil size (see **Tables S2 and S3**).

Given that the ADHD group scored lower on the CFT than controls ($t(158) = 3.64, p < .001, d = 0.58$) and given that controlling for intelligence is still a controversial issue [39], we repeated all models without this covariate.

Group effects were substantially larger in these models as shown in **Fig. 2**. On all four spatial measures, the group effect sums of squares increased by a factor of 6–25: participants with ADHD spent less dwell time and fewer fixations on the matrix and more on the answer options than controls. The effect reached significance for dwell time ($F(1,154) = 6.28, p = .013, \eta^2 = .039$) and fixations on the matrix ($F(1,154) =$

5.23, $p=.024$, $\eta^2=.033$), and narrowly missed it for the two measures on the answer options ($p=.052$ and $.056$; $\eta^2=.024$). The group effect on the run rate increased further ($F(1,154) = 9.00$, $p=.003$, $\eta^2=.055$), while time to first fixation and processing time showed no group effect in either model.

Contrary to hypothesis 3, group differences did not diminish with age; they emerged on the spatial measures only when the CFT covariate was omitted, whereas the higher run rate was present in both models.

3.4 Factor structure of the eye-tracking measures

Factor analyses of our eye movement data revealed the same two dimensions in both groups ($KMO=.69$; 72.6% and 76.7% of the common variance). The four spatial measures loaded the first factor, the three time-based measures the second. These two correlated factors (controls: $r=.62$; ADHD: $r=.50$) thus represent the spatial and temporal aspects of CM processing style with high congruency between groups (**Table S4**; Tucker's $\phi = .996$ and $.972$); age-residualized scores yielded the same structure (**Table S5**). The spatial measures were largely redundant ($|r|\geq .92$, across groups) whereas the temporal ones cohered less closely ($|r|=.25-.50$ in TD and $.33-.60$ in ADHD).

Spatial factor scores were associated with age (controls: $r=.55$, ADHD: $r=.47$, both $p<.001$) and with CFT-assessed intelligence (controls: $r=.33$, $p=.003$, ADHD: $r=.25$, $p=.025$); associations with the temporal factor were smaller (controls: age: $r=.24$, $p=.033$; CFT: $r=.22$, $p=.048$; ADHD (age: $r=.20$, $p=.079$; CFT: $r=.10$, $p=.405$; Fig. 3). Pupil size was unrelated to both factors in both groups.

Both factor scores accounted for some variance in age-residualized CFT scores (controls: 10.8%; $F(2,79) = 4.76$, $p=.011$, ADHD: 6.7%, $F(2,75) = 2.68$, $p=.075$), mainly due to the first factor ($\beta = .332$, $p=.024$, and $\beta = .282$, $p=.035$), but not the second ($\beta=-.006$ and $-.053$, both $p\geq .69$).

3.5 PUPIL SIZE

Pupil size, controlled for age and age-residualized CFT, was unrelated to six eye-movement measures and unmodulated by trial difficulty (**Tables S2 and S3**). Participants with larger dilated pupils took longer to finish the items ($F(1,152) = 6.27$, $p=.013$, $\eta^2=.040$), an effect that did not meet the first Holm–Bonferroni threshold. Pupil size was uncorrelated with CFT scores in controls ($r = -.05$) and ADHD (.13, both $p\geq .26$), and correlated with age in ADHD ($r=.27$, $p=.017$), but not in controls ($r=.07$, $p=.562$).

4. Discussion

The present study set out to investigate how ADHD, age, and intelligence impact spatial and temporal aspects of CM processing style during matrix reasoning, using a nine-item RSPM version and employing eye tracking in participants with ADHD and healthy controls. Within the present experimental setup, we showed that increasing trial difficulty drew fixations towards the matrix, delayed the first fixation of the answers and reduced the number of switches between matrix and answers. Such effects of trial difficulty on gazing behaviour have been reported repeatedly [14, 18]. The replication of these effects

here indicates that our setup worked in the expected way and that the subsequently discussed findings can be meaningfully integrated into the pertinent literature.

4.1 AGE

Increasing age led to increasing CM processing style regarding its spatial aspects: older participants spent more time scanning the matrix and turned their attention to the answer options later, whereas among the temporal measures only time to first fixation increased with age. Our findings using the *original* item format, *conceptually* replicate the developmental increase in CM when solving an *experimentally modified* version of the RSPM in which matrix and answer fields were never presented simultaneously [17]. Other eye-tracking studies of matrix reasoning covered different age ranges ([40]: 5–8ys; [18]: 6- and 9-ys; [41]: 18–28ys; [14]: 17–54ys) than the present study. To our knowledge, the present study is the first to describe with gaze-movement analyses developmental changes in processing style across childhood and adolescence, using original RSPM items and demonstrating linear age effects primarily for the spatial aspects of CM processing style. Given that the core challenge of RSPM problems is the reconstruction of the constructive principles of the matrices by inductive reasoning, this study suggests that an improvement in the “strategic” understanding of the nature of the task is part of the development changes taking place during childhood and adolescence.

4.2 INTELLIGENCE

Age-residualized intelligence exerted an incremental positive effect on spatial CM processing style, but none on its temporal aspects, echoing the afore-mentioned age results in principle, albeit with smaller effect sizes. Given that age and intelligence did not interact statistically, this effect appears to be developmentally stable across most of childhood and adolescence. Unlike the age effects, however, the association between intelligence and CM was independent of trial difficulty. That Liu et al. [42] found the opposite pattern – task difficulty-dependent intelligence effects – in young adults might indicate that the principle facilitation of the understanding of task requirements through intelligence precedes a more fine-grained adaptation of CM processing style to varying task complexity. Associations between intelligence and specific aspects of CM have been found reported in various studies [14, 41, 42] with (young) adults. Importantly, however, in all of these studies gaze movement and intelligence variables were *derived from the same task*, leading to confounded effects. The present study thus increases the validity of this relationship by showing that it also holds when gaze movements and intelligence are assessed independently, using different tasks at different times.

4.3 GROUP DIFFERENCES

Group differences in CM processing style were large, but disappeared when intelligence was statistically controlled for. These effects held for all spatial variables and run rate, which exhibited the greatest group effect sizes with and without controlling for intelligence. Here, like in many previous studies, the ADHD group scored lower in intelligence than the control group, with the magnitude of difference in the

expected range: Frazier et al. [3] report a weighted $d=.61$ (full-scale IQ), our difference amounts to $d = 0.58$ (CFT).

Whether intelligence should be used as a covariate in ADHD studies has been debated. Frazier et al. [3] consider this approach methodologically questionable, since reduced general ability is a characteristic of the disorder. Dennis et al. [39] extend this objection to neurodevelopmental disorders in general: since these develop very early on, the associated level of intelligence is at least partially an outcome of disordered development and thus part of the disorder itself. Controlling for intelligence either by participant recruiting or statistically thus removes this disorder-related variance, which is obviously conceptually inappropriate. In our study we have investigated the influence of intelligence on processing style, with diagnosis and age as potential moderators; the CFT score is thus a predictor of interest and not a confounding variable and we have assumed that intelligence influences processing style, but not vice versa. Both may additionally rest on shared processes such as working memory capacity, making constructive matching harder to sustain [11, 15] and lower test performance more likely. Nevertheless, as long as this debate is undecided, we suggest reporting both models, as they speak to two sides of the same problem.

We found the greatest group effect sizes for the variable *switch rate*, which quantifies the frequency of redirecting the attentional focus from the matrix to the answers and back to the matrix. Such switching in the absence of any changes in the stimulus that elicits gaze movements is an endogenously generated behaviour and may arguably reflect an instance of intra-subject variability (ISV). ISV has been defined typically as the increased fluctuation of reaction times and is among the strongest and most consistent abnormalities in ADHD [43]. It has been found in all kinds of behavioural paradigms [22, 23], various electro-physiological tasks [44], and even in classroom behaviour, where those with ADHD show strong increases in the switch rates between on-task and off-task behaviour [45], underlining the ecological validity of this phenomenon. Furthermore, Brydges et al. [46] demonstrated that the relationship between ADHD symptom severity and fluid intelligence was partially mediated by reaction time variability. The present finding of significant group differences in switch rate even when intelligence was statistically controlled for, confirm such partial redundancy and indicates that whatever the group differences in switch rate represent cannot be fully explained by group differences in intelligence.

4.4 Factor analysis

We found robust factor analytic evidence for a spatial and a temporal factor of CM processing style during matrix reasoning, highly congruent across groups. These two factors differed in their correlations with age and with intelligence, both of which were higher for factor 1. This differential relevance was confirmed in the prediction of CFT performance: entered simultaneously, the two factor scores accounted for 10.8% and 6.7% of the variance in age-residualized CFT scores in controls and ADHD, respectively, carried entirely by factor 1 in both groups.

Regarding RSPM, various eye-movement measures are treated as indicators of the same strategic dimension and used interchangeably [17] to predict constructs like, e.g., test performance [14, 47],

strategy classification [42, 48] or item difficulty [18]. Whether these eye-movement measures share variance and capture the same construct, is not addressed in these studies. The present study provides for the first time dimensional analyses and reveals two robust inter-related factors of problem solving: spatial and temporal. The spatial factor seems to be more relevant and can itself be approximated for practical purposes by a single variable, matrix viewing time, correlating with it in ADHD and controls ($.97 \leq r \leq .98$)

4.5 Pupil size

Pupil size, indicating task-related processing load [29] and person-related effort [28, 49], was detached from all person-related parameters, namely group and intelligence, and its association with gaze measures did not vary with trial difficulty. Only age was positively correlated with pupil size: older participants showed a greater difficulty-related increase during reasoning, suggesting greater task-related effort allocation [28, 30]. This association was somewhat stronger in ADHD ($r=.27$) than in TD ($r=.07$). Because pupil size was measured relative to an immediately preceding resting phase [27], this age effect of both groups concerns reactive, task-related effort allocation rather than developmental changes in tonic pupil diameter.

A relationship between pupil size and processing style was found only for processing time, possibly indicating sustained engagement with the item as an instance of actually expended effort [49]. This preliminary interpretation (failed Holm-Bonferroni correction) aligns with the interpretation of the age-dependency of trial duration: when older children spend more time on difficult items, they also invest more effort in them.

4.6 Limitations and Implications

A major limitation of the present study is the use of a RSPM short-version, consisting of nine items only [35]. This precluded differentiating between correctly and incorrectly solved items, and limits the focus to the solution process independent of its outcome, and limits the generalizability of our findings to the original 60-item version. Also, rather than defining trial difficulty by the original item number, we grouped three items each into "easy", "medium" and "difficult" categories based on empirical solution rates. While this data-driven grouping again diminishes generalizability to the original version, it represents the most straightforward way of relating task difficulty and gazing behaviour and ameliorates the unwanted confounding of item difficulty and item position that is inherent in Raven's matrices [50].

Our study suggests that processing style differences between participants with ADHD and controls covary with their lower intelligence, except for "run rate" as a putative indicator of ISV, that stands controlling for intelligence. Furthermore, we show that controlling for intelligence may dramatically impact comparisons between participants with ADHD and controls. The conclusion here is to control for group differences in potentially disorder-related variables only by the "reversible" statistical way (rather than by sample recruitment) and report results with and without statistical control.

Declarations

Author Contribution

J.B. and C.K. wrote the main manuscript text. J.B. and J.S. prepared figures 1-3. J.H. contributed to the manuscript text. J.B. and A.B. performed the data analysis. D.C. and M.B. were involved in data collection. D.C. designed the experimental setup. S.B., N.S., A.B. and C.K. conceptualised the study. C.K. is the principal investigator. All authors reviewed and approved the final manuscript.

References

1. American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, 5th edn. American Psychiatric Publishing, Washington DC
2. Woods SP, Lovejoy DW, Ball JD (2002) Neuropsychological characteristics of adults with ADHD: a comprehensive review of initial studies. *Clin Neuropsychol* 16:12–34. <https://doi.org/10.1076/clin.16.1.12.8336>
3. Frazier TW, Demaree HA, Youngstrom EA (2004) Meta-analysis of intellectual and neuropsychological test performance in attention-deficit/hyperactivity disorder. *Neuropsychology* 18:543–555. <https://doi.org/10.1037/0894-4105.18.3.543>
4. Claesdotter E, Cervin M, Åkerlund S, Råstam M, Lindvall M (2018) The effects of ADHD on cognitive performance. *Nord J Psychiatry* 72:158–163. <https://doi.org/10.1080/08039488.2017.1402951>
5. Agnew-Blais JC, Polanczyk GV, Danese A, Wertz J, Moffitt TE, Arseneault L (2020) Are changes in ADHD course reflected in differences in IQ and executive functioning from childhood to young adulthood? *Psychol Med* 50:2799–2808. <https://doi.org/10.1017/S0033291719003015>
6. Sonuga-Barke EJ (2005) Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biol Psychiatry* 57:1231–1238. <https://doi.org/10.1016/j.biopsych.2004.09.008>
7. Schulz-Zhecheva Y, Voelkle MC, Beauducel A, Biscaldi M, Klein C (2016) Predicting fluid intelligence by components of reaction time distributions from simple choice reaction time tasks. *J Intell* 4:8. <https://doi.org/10.3390/jintelligence4030008>
8. Schulz-Zhecheva Y, Voelkle MC, Biscaldi M, Beauducel A, Klein C (2024) On the relationships between processing speed, intra-subject variability, working memory, and fluid intelligence — a cross-sectional study. *Intelligence* 105:101836. <https://doi.org/10.1016/j.intell.2024.101836>
9. Willcutt EG, Doyle AE, Nigg JT et al (2005) Validity of the executive function theory of ADHD: a meta-analytic review. *Biol Psychiatry* 57:1336–1346. <https://doi.org/10.1016/j.biopsych.2005.02.006>
10. Raven J, Raven JC, Court JH (1998) Manual for Raven's Progressive Matrices and Vocabulary Scales. Oxford Psychologists, Oxford
11. Carpenter PA, Just MA, Shell P (1990) What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices test. *Psychol Rev* 97:404–431. <https://doi.org/10.1037/0033-295X.97.3.404>

12. Snow RE (1980) Aptitude processes. In: Snow RE, Frederick PA, Montague WE (eds) *Aptitude, learning, and instruction: cognitive process analysis of aptitude*. Erlbaum, Hillsdale, pp 27–63
13. Bethell-Fox CE, Lohman DF, Snow RE (1984) Adaptive reasoning: componential and eye movement analysis of geometric analogy performance. *Intelligence* 8:205–238. [https://doi.org/10.1016/0160-2896\(84\)90009-6](https://doi.org/10.1016/0160-2896(84)90009-6)
14. Vigneau F, Caissie AF, Bors DA (2006) Eye-movement analysis demonstrates strategic influences on intelligence. *Intelligence* 34:261–272. <https://doi.org/10.1016/j.intell.2005.11.003>
15. Gonthier C, Roulin JL (2020) Intraindividual strategy shifts in Raven's matrices, and their dependence on working memory capacity and need for cognition. *J Exp Psychol Gen* 149:564–579. <https://doi.org/10.1037/xge0000660>
16. Klein C, Ettinger U (2019) *Eye movement research*. Springer International Publishing, Cham
17. Gonthier C, Harma K, Gavornikova-Baligand Z (2024) Development of reasoning performance in Raven's matrices is grounded in the development of effective strategy use. *J Exp Psychol Gen* 153:689–705. <https://doi.org/10.1037/xge0001518>
18. Niebaum J, Munakata Y (2023) The development of relational reasoning: an eyetracking analysis of strategy use and adaptation in children and adults performing matrix completion. *Open Mind* 7:197–220. https://doi.org/10.1162/opmi_a_00078
19. Hannon BA (2025) Sources of individual differences in children's matrix problem-solving abilities: evidence from eye movements. *J Educ Train Stud* 14:253. <https://doi.org/10.11114/jets.v14i1.8056>
20. Andreou G, Argatzopoulou A (2025) Unraveling ADHD through eye-tracking procedures: a scoping review. *J Atten Disord* 29:977–988. <https://doi.org/10.1177/10870547251344731>
21. Maron DN, Bowe SJ, Spencer-Smith M et al (2021) Oculomotor deficits in attention deficit hyperactivity disorder (ADHD): a systematic review and comprehensive meta-analysis. *Neurosci Biobehav Rev* 131:1198–1213. <https://doi.org/10.1016/j.neubiorev.2021.10.012>
22. Munoz DP, Armstrong IT, Hampton KA, Moore KD (2003) Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. *J Neurophysiol* 90:503–514. <https://doi.org/10.1152/jn.00192.2003>
23. Klein CH, Raschke A, Brandenbusch A (2003) Development of pro- and antisaccades in children with attention-deficit hyperactivity disorder (ADHD) and healthy controls. *Psychophysiology* 40:17–28. <https://doi.org/10.1111/1469-8986.00003>
24. Bretthauer J, Canu D, Thiemann U et al (2026) Watching reasoning at work: eye movement patterns of participants with ADHD solving Raven's matrices. <https://doi.org/10.1007/s00787-026-03038-y>. *Eur Child Adolesc Psychiatry*
25. Sergeant JA (2000) The cognitive-energetic model: an empirical approach to attention-deficit hyperactivity disorder. *Neurosci Biobehav Rev* 24:7–12. [https://doi.org/10.1016/S0149-7634\(99\)00060-3](https://doi.org/10.1016/S0149-7634(99)00060-3)
26. Sergeant JA (2005) Modeling attention-deficit/hyperactivity disorder: a critical appraisal of the cognitive-energetic model. *Biol Psychiatry* 57:1248–1255.

<https://doi.org/10.1016/j.biopsycho.2004.09.010>

27. Eckstein MK, Guerra-Carrillo B, Miller Singley AT, Bunge SA (2017) Beyond eye gaze: what else can eyetracking reveal about cognition and cognitive development? *Dev Cogn Neurosci* 25:69–91. <https://doi.org/10.1016/j.dcn.2016.11.001>
28. van der Wel P, van Steenbergen H (2018) Pupil dilation as an index of effort in cognitive control tasks: a review. *Psychon Bull Rev* 25:2005–2015. <https://doi.org/10.3758/s13423-018-1432-y>
29. Kahneman D, Beatty J (1966) Pupil diameter and load on memory. *Science* 154:1583–1585. <https://doi.org/10.1126/science.154.3756.1583>
30. Johnson EL, Miller Singley AT, Peckham AD, Johnson SL, Bunge SA (2014) Task-evoked pupillometry provides a window into the development of short-term memory capacity. *Front Psychol* 5:218. <https://doi.org/10.3389/fpsyg.2014.00218>
31. Purcell ZA, Roberts AJ, Handley SJ, Howarth S (2023) Eye movements, pupil dilation, and conflict detection in reasoning: exploring the evidence for intuitive logic. *Cogn Sci* 47:e13293. <https://doi.org/10.1111/cogs.13293>
32. Toth AJ, Campbell MJ (2019) Investigating sex differences, cognitive effort, strategy, and performance on a computerized version of the mental rotations test via eye tracking. *Sci Rep* 9:19430
33. Canu D, Ioannou C, Müller K et al (2022) Evidence towards a continuum of impairment across neurodevelopmental disorders from basic ocular-motor tasks. *Sci Rep* 12:16521. <https://doi.org/10.1038/s41598-022-19661-z>
34. Weiß RH, Weiß B (2008) CFT 20-R mit WS/ZF-R: Grundintelligenztest Skala 2-Revision (CFT 20-R). Hogrefe, Göttingen
35. Bilker WB, Hansen JA, Brensinger CM, Richard J, Gur RE, Gur RC (2012) Development of abbreviated nine-item forms of the Raven's standard progressive matrices test. *Assessment* 19:354–369. <https://doi.org/10.1177/1073191112446655>
36. Delaney HD, Maxwell SE (1981) On using analysis of covariance in repeated measures designs. *Multivar Behav Res* 16:105–123. https://doi.org/10.1207/s15327906mbr1601_6
37. Blom G (1958) Statistical estimates and transformed beta-variables. Wiley, New York
38. Beasley TM, Erickson S, Allison DB (2009) Rank-based inverse normal transformations are increasingly used, but are they merited? *Behav Genet* 39:580–595
39. Dennis M, Francis DJ, Cirino PT, Schachar R, Barnes MA, Fletcher JM (2009) Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *J Int Neuropsychol Soc* 15:331–343. <https://doi.org/10.1017/S1355617709090481>
40. Chen Z, Honomichl R, Kennedy D, Tan E (2016) Aiming to complete the matrix: eye-movement analysis of processing strategies in children's relational thinking. *Dev Psychol* 52:867–878
41. Laurence PG, Mecca TP, Serpa A, Martin R, Macedo EC (2018) Eye movements and cognitive strategy in a fluid intelligence test: item type analysis. *Front Psychol* 9:380.

<https://doi.org/10.3389/fpsyg.2018.00380>

42. Liu Y, Zhan P, Fu Y et al (2023) Using a multi-strategy eye-tracking psychometric model to measure intelligence and identify cognitive strategy in Raven's Advanced Progressive Matrices. *Intelligence* 100:101782. <https://doi.org/10.1016/j.intell.2023.101782>
43. Klein C, Wendling K, Huettner P, Ruder H, Peper M (2006) Intra-subject variability in attention-deficit hyperactivity disorder. *Biol Psychiatry* <https://doi.org/10.1016/j.biopsych.2006.04.003>
44. Saville CW, Feige B, Kluckert C et al (2015) Increased reaction time variability in attention-deficit hyperactivity disorder as a response-related phenomenon: evidence from single-trial event-related potentials. *J Child Psychol Psychiatry* 56:801–813. <https://doi.org/10.1111/jcpp.12348>
45. Kofler MJ, Rapport MD, Sarver DE et al (2013) Reaction time variability in ADHD: a meta-analytic review of 319 studies. *Clin Psychol Rev* 33:795–811. <https://doi.org/10.1016/j.cpr.2013.06.001>
46. Brydges CR, Ozolnieks KL, Roberts G (2021) Working memory and intraindividual variability in response time mediate fluid intelligence deficits associated with ADHD symptomatology. *J Atten Disord* 25:63–72. <https://doi.org/10.1177/1087054718772143>
47. Laurence PG, Jana TA, Bunge SA, Macedo EC (2023) Eye gaze patterns during reasoning provide insights regarding individual differences in underlying cognitive abilities. *J Intell* 11:75. <https://doi.org/10.3390/jintelligence11040075>
48. Liu Y, He K, Man K, Zhan P (2025) Exploring critical eye-tracking metrics for identifying cognitive strategies in Raven's Advanced Progressive Matrices: a data-driven perspective. *J Intell* 13:14. <https://doi.org/10.3390/jintelligence13020014>
49. da Silva Castanheira K, LoParco S, Otto AR (2021) Task-evoked pupillary responses track effort exertion: evidence from task-switching. *Cogn Affect Behav Neurosci* 21:592–606. <https://doi.org/10.3758/s13415-020-00843-z>
50. Sun S, Schweizer K, Ren X (2019) Item-position effect in Raven's matrices: a developmental perspective. *J Cogn Dev* 20:370–379. <https://doi.org/10.1080/15248372.2019.1581205>

Figures

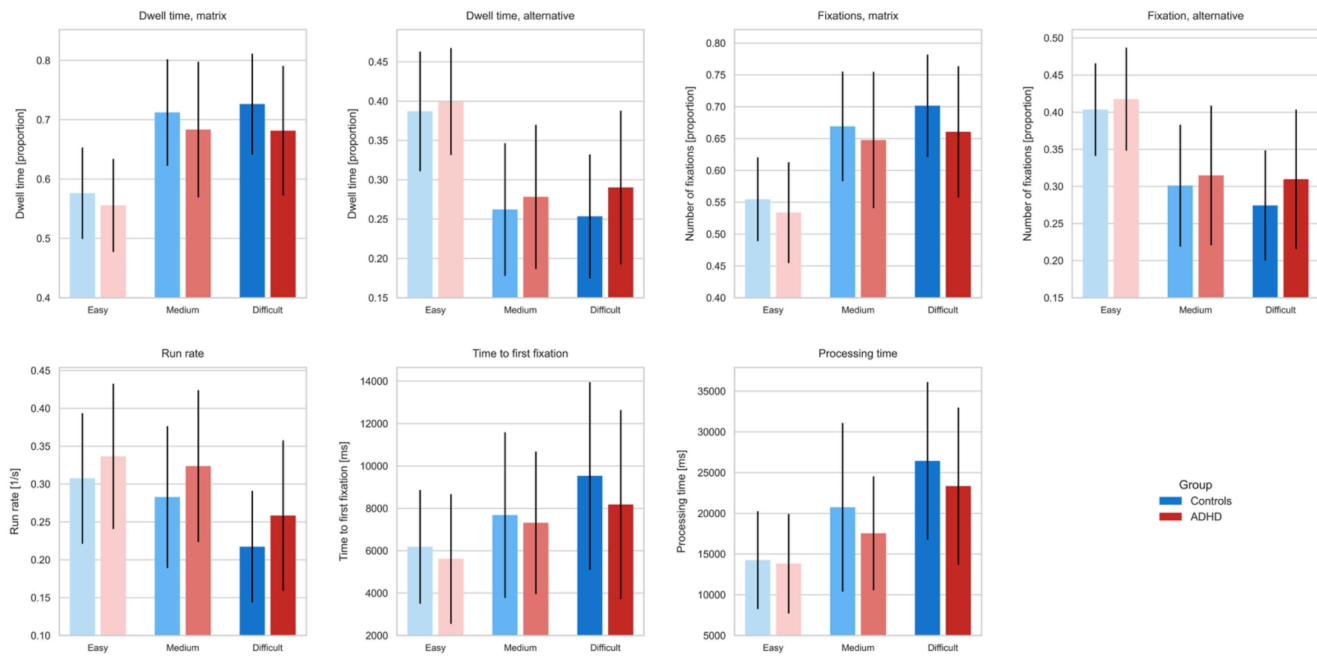


Figure 1

Eye-movement measures by group (TD/ADHD) and difficulty level (easy/medium/difficult). Bars show untransformed means; error bars=+SD. Top row: spatial measures. Bottom row: temporal measures. N=160 (ADHD=78; controls= 82).

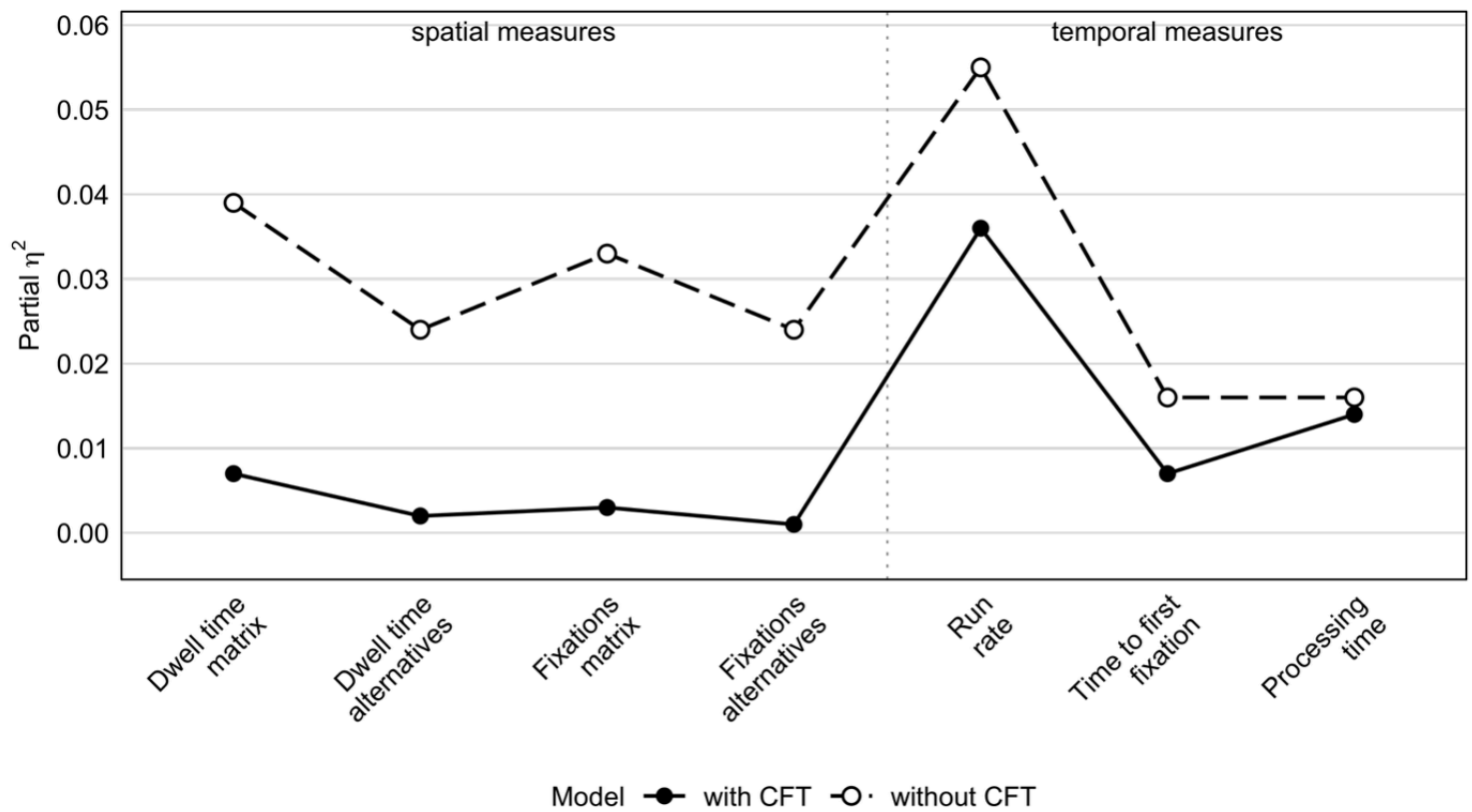


Figure 2

Partial η^2 of the group main effect on the seven eye-movement measures, in models with and without CFT as a covariate. Based on RM-ANCOVAs with Blom-transformed values; both models include age and baseline-corrected pupil size as covariates. Filled circles connected by solid lines model with CFT; open circles connected by dashed lines model without CFT. The dotted vertical line separates spatial from temporal measures.

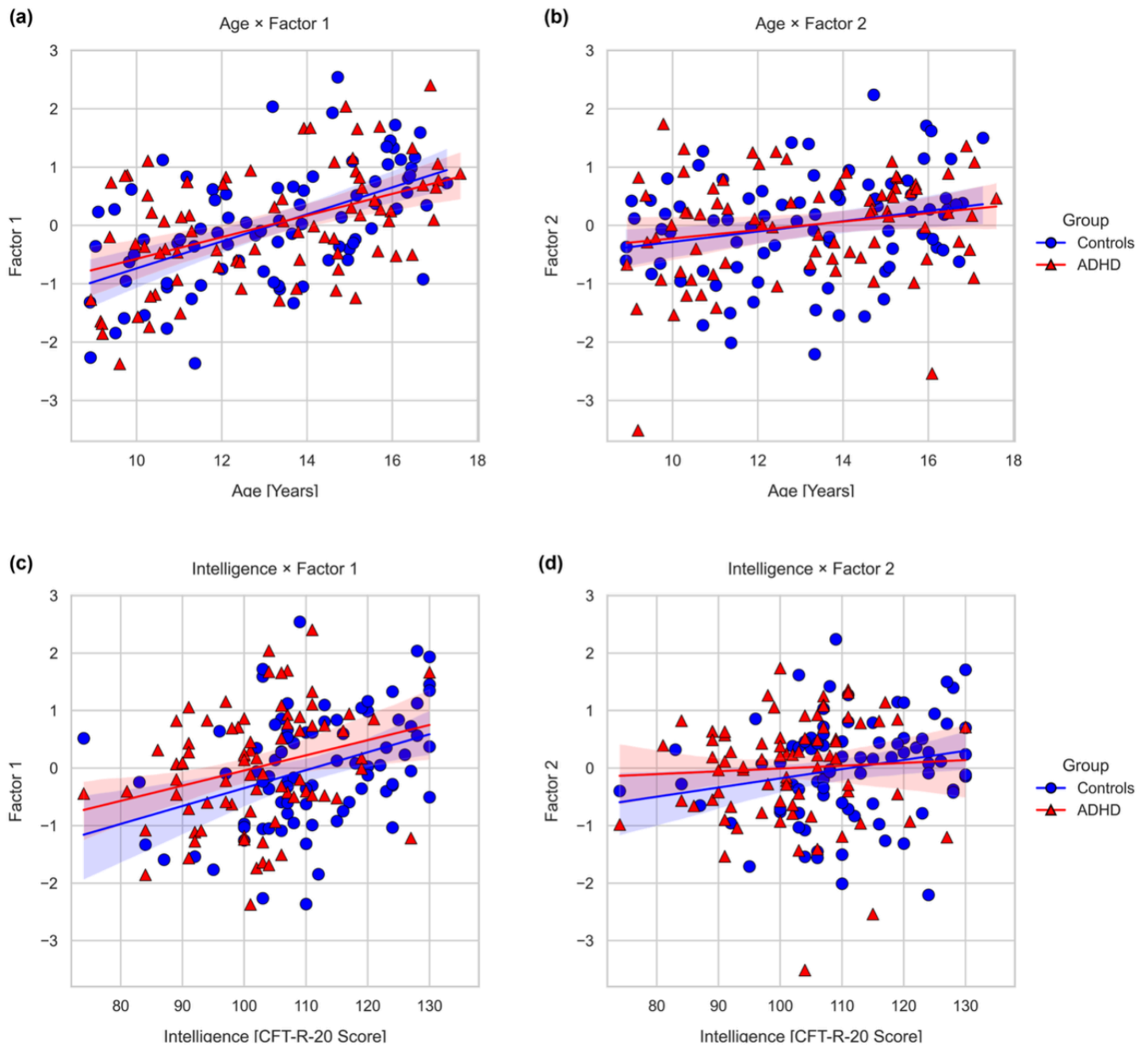


Figure 3

Factor scores as a function of age (a, b) and CFT (c, d). Factor 1: spatial attentional distribution; higher values indicate stronger weighting of the matrix. Factor 2: temporal processing dynamics. Factor scores were derived from the two-factor solution estimated separately for each group (Table S4).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)