Questioning cognitive heterogeneity and intellectual functioning in fetal alcohol spectrum disorders from the Wechsler intelligence scale for children

Eliot Kerdreux, Justine Fraize, Pauline Garzón, Esther Chalain, Léa Etchebarren, Delphine Sitbon, Anna Maruani, Odile Boespflug-Tanguy, Lucie Hertz-Pannier, Marion Noulhiane, et al.

Routledge
Taylor & Francis Group

The Clinical Neuropsychologist

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ntcn20

Questioning cognitive heterogeneity and intellectual functioning in fetal alcohol spectrum disorders from the Wechsler intelligence scale for children

Eliot Kerdreux, Justine Fraize, Pauline Garzón, Esther Chalain, Léa Etchebarren, Delphine Sitbon, Anna Maruani, Odile Boespflug-Tanguy, Lucie Hertz-Pannier, Marion Noulhiane, Charlotte Pinabiaux & David Germanaud

To cite this article: Eliot Kerdreux, Justine Fraize, Pauline Garzón, Esther Chalain, Léa Etchebarren, Delphine Sitbon, Anna Maruani, Odile Boespflug-Tanguy, Lucie Hertz-Pannier, Marion Noulhiane, Charlotte Pinabiaux & David Germanaud (16 Nov 2023): Questioning cognitive heterogeneity and intellectual functioning in fetal alcohol spectrum disorders from the Wechsler intelligence scale for children, The Clinical Neuropsychologist, DOI: 10.1080/13854046.2023.2281703

To link to this article: https://doi.org/10.1080/13854046.2023.2281703

THE CLINICAL NEUROPSYCHOLOGIST
https://doi.org/10.1080/13854046.2023.2281703

Routledge
Taylor & Francis Group

Questioning cognitive heterogeneity and intellectual functioning in fetal alcohol spectrum disorders from the Wechsler intelligence scale for children

Eliot Kerdreux[a,b], Justine Fraize[a,b], Pauline Garzón[a,b], Esther Chalain[c], Léa Etchebarren[d], Delphine Sitbon[d], Anna Maruani[e], Odile Boespflug-Tanguy[f], Lucie Hertz-Pannier[a,b], Marion Noulhiane[a,b], Charlotte Pinabiaux[g] and David Germanaud[a,b,c]

[a]CEA Paris-Saclay, Frederic Joliot Institute, NeuroSpin, UNIACT, Gif-sur-Yvette, France; [b]Université Paris Cité, Inserm, NeuroDiderot, inDEV, Paris, France; [c]Genetics Department, Centre of Excellence InovAND, Robert-Debré Hospital, AP-HP, Paris, France; [d]Pediatric Neurology Department, Centre of Excellence InovAND, Robert-Debré Hospital, AP-HP, Paris, France; [e]Child and Adolescent Psychiatry Department, Centre of Excellence InovAND, Robert-Debré Hospital, AP-HP, Paris, France; [f]Université Paris Cité, Inserm, NeuroDiderot, NeuroDEV, Paris, France; [g]Université Paris Nanterre, Department of Psychology, DysCo Lab, Nanterre, France

ABSTRACT

Introduction: Fetal Alcohol Spectrum Disorders (FASD) are characterized by a variety of multiple cognitive and behavioral impairments, with intellectual, attentional, and executive impairments being the most commonly reported. In populations with multiple neurodevelopmental disorders, the Full Scale Intelligence Quotient (FSIQ) may not be a proper measure of intellectual abilities, rarely interpreted in FASD clinical practice because the heterogeneity of the cognitive profile is deemed too strong. We propose a quantitative characterization of this heterogeneity, of the strengths and weaknesses profile, and a differential analysis between global cognitive (FSIQ) and elementary reasoning abilities in a large retrospective monocentric FASD sample. **Methods**: Using clinical and cognitive data (Wechsler Intelligence Scale for Children) from 107 children with FASD, we characterized subject heterogeneity (variance and scatter of scaled/composite scores), searched for strengths and weaknesses, and specified intellectual functioning in terms of FSIQ and elementary reasoning (General Abilities Index, Highest Reasoning Scaled Score), in comparison with standardization norms and a Monte-Carlo-simulated sample from normalization data. **Results:** Performance of children with FASD was lower on all subtests, with a significant weakness in working memory and processing speed. We found no increase in the variance and scatter of the scores, but a discordance between the assessment of global cognitive functioning (28% borderline, 23% deficient) and that of global and elementary reasoning abilities (23–9% borderline, 15–14% deficient). **Conclusion**: Our results question the notion of WISC profile heterogeneity in FASD and point to working memory and processing speed over-impairment, with global repercussions but most often preserved elementary reasoning abilities.

ARTICLE HISTORY

Received 22 March 2023
Accepted 2 November 2023
Published online 16 November 2023

KEYWORDS

Fetal alcohol spectrum disorders; prenatal alcohol exposure; intellectual abilities; WISC profile heterogeneity; elementary reasoning abilities

CONTACT Eliot Kerdreux eliot.kerdreux@inserm.fr; David Germanaud david.germanaud@cea.fr CEA Paris-Saclay, Frederic Joliot Institute, NeuroSpin, UNIACT, Gif-sur-Yvette, 91191, France

Supplemental data for this article can be accessed online at https://doi.org/10.1080/13854046.2023.2281703.

Introduction

Alcohol consumption during pregnancy is a major cause of neurodevelopmental disorders (Flak et al., 2014; Tsang et al., 2016). The clinical consequences of prenatal alcohol exposure are grouped under the diagnosis continuum of Fetal Alcohol Spectrum Disorder (FASD) (Riley & McGee, 2005). Several international diagnostic guidelines converge on common key clinical criteria (Astley & Clarren, 2000; Cook et al., 2016; Hoyme et al., 2016). They distinguish fetal alcohol syndrome (FAS), with its specific association of physical features, from non-specific non-syndromic FASD (NS-FASD), when these physical criteria are absent or incomplete but a probabilistic causal link can be assumed between cognitive and behavioral impairment and significant prenatal alcohol exposure.

Primary studies have reported a group-level cognitive profile in FASD characterized by poor intelligence or intellectual functioning. The majority of children with FASD showed borderline to mildly deficient Intellectual Quotient (IQ) (Mattson et al., 1997, 2011; Streissguth et al., 1996), but IQ score showed a weak correlation with adaptive functioning in this population (Boseck et al., 2015; Kautz-Turnbull & Petrenko, 2021). A typical FASD cognitive profile dominated by poorer performance on executive and perceptual-motor tasks than verbal ones has also been reported (Sampson et al., 1989; Streissguth et al., 1989, 1996). However, this profile has not been fully or consistently replicated and the current consensus insists instead on inter-individual variability in the cognitive profile of individuals with FASD (Astley et al., 2009; Kodituwakku, 2007). Indeed, more recent studies of the cognitive phenotype in FASD highlight impairments in cognitive domains that vary from study to study but collectively cover the full range of cognition: impairments in executive functions, working memory, attention, learning, language, problem-solving, intelligence, and fluid reasoning (Astley et al., 2009; Green et al., 2009; J. L. Jacobson et al., 2021; Lewis et al., 2015; Rasmussen et al., 2013; Streissguth et al., 1989; Vaurio et al., 2008). Yet, within this variability of profiles, the executive deficit seems to be the most consistent and common, leading some authors to conceptualize it as the core of cognitive symptomatology (e.g. Fuglestad et al., 2015). In clinical practice, multiple neurodevelopmental diagnoses have been given to patients with prenatal alcohol exposure reflecting this variability and multiplicity of profiles. These diagnoses are in decreasing order of frequency: Attention Deficit Disorder with or without Hyperactivity (ADHD), mild to borderline Intellectual Development Disorder (IDD), various Specific Learning Disorders (SLD), and Developmental Coordination Disorder (DCD) (Chasnoff et al., 2015; Geier & Geier, 2022; Popova et al., 2016; Weyrauch et al., 2017). This diversity and multiplicity of functional diagnoses is also reflected in international diagnostic guidelines, where cognitive and behavioral impairment in at least two domains is a required criterion for the diagnosis of FASD (Astley & Clarren, 2000; Cook et al., 2016; Hoyme et al., 2016).

The question of how multiple impairments, particularly attentional and executive ones, influence the assessment of intellectual functioning in children with FASD has implications for the diagnosis of Intellectual Developmental Disorder (IDD) in this population, but the issue has not been specifically studied. According to the International Classification of Diseases 11th revision (ICD-11), IDD is characterized by

intellectual and adaptive functioning below the 2.5th percentile of the general population (World Health Organization, 2019). Intelligence is most often seen as an overall measure of cognitive abilities, a latent variable that explains most of the variability in cognitive performance (Deary et al., 2010). The Wechsler Intelligence Scales for Children (WISC) are the most widely used scales for intellectual assessment in the pediatric population. Based on multiple specific cognitive tasks, a hierarchical factorial model supports the calculation of several domain-specific composite scores and the Full Scale Intellectual Quotient (FSIQ), a proxy for the most general factor (Wechsler, 2016). However, the use of the FSIQ to assess intellectual functioning may be misleading in some individuals with neurodevelopmental disorders because of strong cognitive dissociation (Fiorello et al., 2007; Flanagan & Kaufman, 2004). Hence, earlier versions of the WISC cautioned against the relevance of the FSIQ if there was "too much" heterogeneity in the cognitive profile, at least in the series of standard notes or scores hereafter referred to as the "Wechsler or WISC profile" (Wechsler, 1991). This claim is no longer present in more recent editions (e.g. Wechsler, 2005b, 2016) but is still current in daily clinical practice, which can make intellectual assessment somewhat complicated or equivocal. Indeed, the Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5) definition of IDD tries to specify the clinical meaning of intellectual functioning as a reflection of abilities in" *reasoning, problem-solving, planning, abstract thinking, judgment, academic learning, and learning from experience*" (American Psychiatric Association, 2013). However, it still mixes different levels of integration, and to a certain extent, reasoning and executive functioning. To obtain a composite measure of intellectual functioning that better disentangles elementary reasoning abilities from more integrative, executive, or procedural ones, the WISC proposes the General Ability Index (GAI) that focuses on reasoning subtests and excludes working memory and processing speed ones (Lecerf et al., 2010). Nevertheless, it remains sensitive to instrumental dysfunction such as verbal, gestural, and visual-spatial impairments or any source of domain variability in reasoning. Alongside the FSIQ and GAI, a third level of analysis may then be proposed at the scale of each subtest specifically built upon an elementary reasoning task—for instance inductive, sequential, quantitative, and categorical (McGrew, 2009)—to unveil any reasoning ability of the subject, may it be preserved or accessible in only one modality.

This study aims to question the analysis of the WISC-based first step of cognitive investigation in children with FASD to determine what can be expected in terms of the WISC profile (internal heterogeneity and inter-subject regularity) and how to interpret intellectual functioning outcomes. As a first hypothesis, we assumed the WISC profile to be more heterogeneous in the FASD than in the normalization sample, reflecting the variability and multiplicity of cognitive impairments previously described in the literature. To test this hypothesis, we measured individual variance in inter-scaled and inter-composite scores as an adjunct to the classical scatter analysis. Secondly, we assumed sufficient regularity in the WISC profile among children with FASD to identify a mean profile of strengths and weaknesses at the group level. Finally, we expected that in the FASD sample, an assessment of intellectual functioning grounded in reasoning abilities with GAI would lead to the classification of fewer patients within the deficit or borderline range compared to the utilization of a global cognitive functioning indicator such as FSIQ. We introduced the Highest Reasoning Scaled Score

(HRSS) as a potentially more differentiating measure than the GAI, as a new metric that would specify intellectual assessment by identifying and quantifying the preservation of elementary reasoning across at least one modality. We then compared the range of intellectual functioning within our FASD sample using the HRSS, the GAI, or the FSIQ for classification.

Material and methods

Participants

Participants were retrospectively included from a large clinical series of patients attending the NDD-dedicated child neurology consultation at the Robert-Debré University Hospital (AP-HP Paris, France) between 2012 and 2020. The diagnosis of FASD was established according to the 4-Digit Diagnostic Code or 4-DDC (Astley & Clarren, 2000) and the revised guidelines of the Institute of Medicine (Hoyme et al., 2016) and a complete differential diagnosis work-up had to be performed, including systematic brain MRI and genetic testing (see supplementary material Table 1 for details of the diagnostic procedure). We assessed the concordance between FASD diagnostic assignments using one or the other guidelines.

Inclusion required a fully documented diagnosis of FASD by a specialized child neurologist with reference expertise in FASD (D.G.) and the availability of a cognitive assessment using the WISC 4th or 5th edition, *de facto* restricting the age range to between 6 and 16 years of age. Exclusion criteria were prenatal exposure to another major developmental toxic agent such as sodium valproate (see supplementary material Table 1 for details) or explicit refusal to participate in the study. A total of 108 children met the inclusion criteria, one patient with concomitant exposure to sodium valproate was excluded, leaving a total of 107 children (39 girls, 68 boys). Three previous neuroimaging studies have described part of the cohort (see Fraize, Convert, et al., 2023; Fraize, Garzón, et al., 2023; Fraize, Fischer, et al. 2023).

This study was not a clinical trial, and the use of care-related data was approved by the appropriate ethics committee (CER-Paris-Saclay 2020-094) ensuring proper consent of participants.

Dataset and metrics

Primary variables and cognitive profiles

From the medical records of each patient, we collected several clinical and demographic variables (age, sex, family status, maternal professional status, FASD diagnosis, functional diagnosis in terms of NDD diagnosis (DSM-5), methylphenidate treatment at time of assessment, site of assessment), as well as WISC scaled scores, composite scores and version of the scale (IV or V). All cognitive assessments were reviewed by a neuropsychologist of the research team (E.K.).

WISC profile heterogeneity and strengths/weaknesses

To examine the within-subject variability, we calculated the scatter (i.e. the difference between the maximum and minimum score) of the scaled and composite scores,

as well as the variance between scaled scores and the variance between composite scores.

To examine the profile of patients' strengths and weaknesses, we computed the differences between each composite score and the Mean of the Composite Scores (MCS), and the differences between each scaled score and the Mean of the Scaled Scores used to calculate the composite scores (MSS). In the following, we will refer to these differences from the reference simply as strengths and weaknesses when the statistical tests for inference are significant.

Intellectual functioning and elementary reasoning abilities

We calculated the General Abilities Index (GAI) based on the French norms published by Lecerf et al. (2010) for the WISC-IV subgroup, as it was not available in the French Administration and Scoring Manual (Wechsler, 2005a). We selected the Highest Reasoning Scaled Score (HRSS) among the elementary reasoning subtests (Similarities, Comprehension, Block Design, Matrix Reasoning, Figure Weight, Visual Puzzles, and Picture Concepts) to obtain a single score reflecting the best elementary reasoning irrespective of the modality.

We categorized FSIQ, GAI, and HRSS into three categories:

1. Normal range: FSIQ $\geq$ 80; GAI $\geq$ 80; HRSS $>$ 8
2. Borderline range: 70 $\leq$ FSIQ $\leq$ 79; 70 $\leq$ GAI $\leq$ 79; HRSS = 8
3. Deficit range: FSIQ $<$ 70; GAI $<$ 70; HRSS $\leq$ 7

The choice of the deficit range thresholds for the classification based on the FSIQ and the GAI was made in accordance with the threshold reported in the DSM-V and ICD (American Psychiatric Association, 2013; World Health Organization, 2019). For the borderline range, since DSM-V no longer specifies an FSIQ score range, we selected the threshold corresponding to the "below average score" threshold of the American Academy of Clinical Neuropsychology consensus (Guilmette et al., 2020). The thresholds for the classification based on the HRSS were chosen for both fitting the distribution of the highest score among the five elementary reasoning subtests and being relevant in terms of clinical interpretation of the highest-scoring subtest. Indeed, we computed the HRSS cumulative distribution according to the WISC-IV and WISC-V standardization samples (see below the *Simulated Comparison Group* section). The 2.5th percentile was between an HRSS of 7 and 8 in both cases and the 10th between 8 and 9 for the WISC-V while being a bit above 9 for the WISC-IV (see supplementary material Table 2), consistent with our choice to limit the borderline range to an HRSS of 8. At the level of the highest scoring subtest, these choices remained in agreement with the classical clinical interpretation that states $\leq$ 7 to be weak and 8 to be the first value of the average range (Sattler & Dumont, 2004), hence arguably limit, particularly for the highest scoring.

Statistics

All analyses were performed using R statistical Software (v4.1.2; R Core Team, 2021). We used the Benjamini and Hochberg False Discovery Rate (FDR) method (Benjamini

& Hochberg, 1995) for *p*-value correction to account for the increased risk of type I error in multiple comparisons.

Accounting for covariables and test version

We assessed the feasibility of pooling cognitive data from the two versions of the WISC. First, we tested whether the WISC-IV and WISC-V subgroups differed on clinical, cognitive, and demographic variables (Student's t-test for continuous dependent variables and Chi-2 tests for categorical variables). Next, we assessed the effect of the WISC version on the scaled and composite scores in a unifactorial analysis (Student's t-test). We then performed multifactorial analysis (ANOVAs) to disentangle the influence of clinical, sociodemographic, and version factors. To account for missing data in the sociodemographic variables, we estimated the overall percentage of missing data, then checked whether these missing data were randomly distributed using Little's method (Little, 1988) from the *naniar* package (Tierney & Cook, 2018), and finally, we performed an imputation of missing data using mixed data factorial analysis from the *FactoMineR* package (Lê et al., 2008).

Simulated comparison group (SCG). Since the norms available in the administration and scoring manuals (Wechsler, 2005a, 2016) did not cover all the metrics we wanted to analyze or the norms of the metrics used were not calculated in the same way for both versions of the scales, we performed two Monte-Carlo simulations based on the method fully described in Aubry and Bourdin (2018) to construct a simulated comparison group (SCG). This method consists in generating a large sample ($n = 10{,}000$) that follows a normal distribution while respecting the constraints of the correlations between the subtests of the standardization sample. This procedure was used twice to simulate two SCGs based on the version of the scale (WISC-IV or WISC-V). To compare the pooled FASD data (WISC-IV and WISC-V), we created a pooled SCG by randomly selecting the same percentage of WISC-IV and V SCGs as in the FASD sample.

Inferences. As a preliminary to our heterogeneity analyses, we tested whether there was a relationship between the magnitude of the FSIQ and the proposed heterogeneity measures in the standardization sample, i.e. whether lower IQ or higher IQ was associated with lower or higher WISC profile heterogeneity. To answer this preliminary question, we tested a linear predictive model of scatter and of variance by FSIQ. Then, we compared the mean variance and scatter between the subgroups of children with FASD and the SCG using two-sample unpaired Student's t-tests, we also compared the proportions of children with FASD with abnormal variance and scatter (greater than or equal to the 90th percentile of the SCG distribution) using Fisher's exact tests, and finally we compared the means of the scatter of the FASD subgroups with those of the standardization samples using one-sample Student's t-tests. As for the profile itself, we compared the means of the FASD and pooled subgroups with the SCG for each scaled score and composite score using unpaired two-sample

Student's t-tests. The procedure was the same for weaknesses and strengths, but we also compared the proportions of children with FASD who had a weakness at or below the 10th percentile of the SCG distribution using Fisher's exact tests. Finally, we compared the proportions of children with FASD and SCG subjects classified in the deficient, borderline, or normal range using Fisher's exact tests and associated odds ratio (OR). We also conducted Bootstrap resampling comparisons ($N_{samples}$ = 1000) to assess whether there were differences in the OR of being classified in in the borderline-to-deficient range between the different metrics. For the latter analysis, the sample of children with FASD was reduced in size (N = 88) because the analysis required sufficiently complete data to calculate the FSIQ, GAI, and the HRSS across all reasoning subtests. Post-hoc multiple linear regression analyses were performed with the functional diagnoses and their interaction, or the sociodemographic variables, as predictor variables to examine their potential contribution to WISC profile heterogeneity.

The search for an abnormal proportion of subjects above or below the 10th and 90th percentile limit values (known as normative analysis) was considered complementary to comparing means in the case of potentially unequal or heterogeneous distributions among patients, and because it would show the extent of potential FASD specificity at the individual level.

Results

All included children had received a cognitive assessment using the WISC-IV (62.62%) or WISC-V (37.38%). This assessment was performed on-site (Robert-Debré Hospital) for 55.14% and at a mean age of 10.14 years (SD = 2.28). Descriptive statistics of the 4-DDC criteria in the FASD sample are presented in supplementary material Tables 3 and 4, proportions of the main associated NDD diagnoses in the FASD sample are represented in supplementary material Figure 1. Interrater reliability between the 4-DDC and the IOM diagnostic guidelines was high (K = .79, $p < .001$).

Parameters affecting WISC performance

Unifactorial effect of WISC version

Descriptive statistics of the sociodemographic and clinical variables of the FASD sample and the comparison between the two subgroups according to the WISC version are presented in Table 1. The FASD subgroups did not differ in clinical and sociodemographic variables except for methylphenidate treatment at the time of assessment with a slightly higher proportion of treated subjects in the WISC-V subgroup (N = 15 *vs*. 11 for the WISC-IV subgroup) but the p-value was not significant after correction for multiple comparisons ($\chi^2(1) = 4.68$, $p_{adj} = .49$). Mean performances on the scaled scores and composite scores were not different between the two groups except for the Similarities subtest ($t(102) = 3.03$, $p_{adj} = .02$, cohen's $d = .59$), with a higher mean in the WISC-IV subgroup (Table 2).

Table 1. Clinical and demographic characterization of the fetal alcohol spectrum disorders sample and comparisons between Wechsler intelligence scale for children 4th and 5th editions groups.

			Missing values	*Group comparison*
Variable	*N*	%	*N*	*p*
Age *M, SD*	10.14	2.28	0	.93
Sex				
Boys	68	63.55	0	1.00
Girls	39	36.45		
Family status				
Adopted	72	67.29	0	1.00
Foster-cared	21	19.63		1.00
Professional status				
Retired	6	5.61	14	.90
Family assistant	17	15.89		1.00
Employee	13	12.15		1.00
Worker	1	0.93		1.00
Unemployed	8	7.48		1.00
Intermediate	17	15.89		.83
Manager and intellectual	31	28.97		1.00
FASD diagnosis*				
Fetal alcohol syndrome	61	57.01	0	1.00
Non-syndromic FASD	46	42.99		
Functional diagnosis				
Attention deficit/hyperactivity disorders	93	86.92	0	1.00
Language disorders	40	37.38		.83
Developmental coordination disorder	22	20.56		1.00
Autism spectrum disorders	5	4.67		1.00
Methylphenidate treatment	26	24.30	10	.49

Note. * With revised institute of medicine criteria: 62 children with fetal alcohol syndrome, 39 with non-syndromic fetal alcohol spectrum disorders, and 6 with another diagnosis.
M = Mean; SD = Standard Deviation; FASD = Fetal Alcohol Spectrum Disorders; Group comparison = false discovery rate corrected *p*-value from Chi squared tests or Student's t-tests between *Wechsler Intelligence Scale for Children 4th and 5th editions* FASD groups.

Multifactorial clinical and sociodemographic effects

Missing data on the 7 subtests common to both scales and the 7 clinical and socio-demographic variables accounted for 3.60% of the data and were completely randomly distributed ($\chi^2(227) = 199$, $p = .08$), allowing their imputation for multifactorial ANOVA. We found only two parameters affecting WISC performances in FASD: a sex effect on the Coding subtest (girls > boys, $F(1,96) = 8.38$, $p = .005$, $p_{adj} = .04$, $\eta^2 = .08$) and an FASD diagnostic effect on the Symbol Search subtest (NS-FASD > FAS, $F(1,96) = 8.22$, $p = .005$, $p_{adj} = .04$, $\eta^2 = .08$). The tendency for a WISC version effect on the Similarities subtest ($p_{adj} = .15$), Socio-economic status and methylphenidate treatment effects on the Vocabulary subtest ($p_{adj} = .19$, $p_{adj} = .07$ resp.) and a sex effect on the Symbol Search subtest ($p_{adj} = .14$) did not last after correction for multiple comparisons (Table 3).

We performed further analyses on both the separate WISC subgroups and the pooled data, except for the Similarities subtest and the Verbal Comprehension Index for which pooling seemed too disputable.

WISC profiles

Heterogeneity

Preliminary analyses found no significant relationship between FSIQ categories and the variances or the scatters of the composite and scaled scores (all $p_{adj} > .05$,

Table 2. Scaled and composite score mean comparisons between fetal alcohol spectrum disorders groups.

	WISC-IV $N = 67$		WISC-V $N = 40$			
Scale	*n*	*M* (*SD*)	*n*	*M* (*SD*)	*t*	*d*
Indexes						
FSIQ	54	81.91 (16.89)	39	80.97 (14.95)	0.28*ns*	0.06
Verbal comprehension	64	89.34 (18.81)	40	85.97 (15.00)	1.01*ns*	0.20
Perceptual reasoning	64	83.33 (15.73)	–	–	–	–
Visual spatial	–	–	40	86.97 (12.40)	–	–
Fluid reasoning	–	–	40	87.27 (15.07)	–	–
Working memory	62	79.29 (17.41)	38	80.05 (14.68)	−0.23*ns*	−0.05
Processing speed	64	82.73 (17.23)	39	84.46 (19.14)	−0.46*ns*	−0.09
Subtests						
Similarities	66	9.31 (3.93)	39	7.31 (2.84)	3.03*	0.59
Vocabulary	65	7.57 (3.69)	39	7.49 (3.11)	0.12*ns*	0.02
Comprehension	63	8.19 (2.62)	–	–	–	–
Block design	62	7.02 (3.14)	39	7.59 (2.61)	−0.99*ns*	−0.20
Visual puzzles	–	–	39	7.64 (2.43)	–	–
Picture concept	64	8.05 (3.52)	–	–	–	–
Matrix reasoning	65	7.37 (3.20)	39	7.87 (3.10)	−0.79*ns*	−0.16
Figure weight	–	–	39	7.69 (3.00)		
Digit span	65	6.95 (3.58)	39	5.77 (3.10)	1.78*ns*	0.35
Letter-number seq.	56	6.37 (3.24)	–	–	–	–
Picture span	–	–	37	6.92 (2.43)	–	–
Coding	64	6.05 (3.60)	39	6.51 (3.62)	−0.63*ns*	−0.13
Symbol search	62	7.50 (3.52)	38	7.92 (3.42)	−0.59*ns*	−0.12

Note. FSIQ = Full Scale Intelligence Quotient; M = Mean; SD = Standard Deviation; ; t = t statistic from student's t-test; d = Cohen's d.
False discovery rate adjusted *p*-value: *ns* $p_{adj} > .05$; * $p_{adj} \leq .05$; ** $p_{adj} < .01$; *** $p_{adj} < .001$.

Table 3. Socio-demographic and clinical factors on common subtest performances between the two versions of Wechsler intelligence scales for children in fetal alcohol spectrum disorders sample.

		Analysis of variance factors						
Subtests	*N*	Sex	Mother's prof. status	Family status	FASD diagnosis	Scale version	Site of test	Methylphenidate treatment*
Similarities	105	.95	.40	.64	.49	*.15*	.98	.47
Vocabulary	104	.95	*.19*	.40	.80	.74	.98	*.07*
Block design	101	.95	.40	.40	.49	.43	.98	.56
Matrix reasoning	104	.95	.73	.68	.49	.43	.98	.56
Digit span	104	.95	.40	.32	.49	.43	.98	.56
Coding	103	**.04**	.40	.32	.49	.43	.98	.56
Symbol search	100	*.14*	.40	.32	**.04**	.43	.98	.98

Note. Before this analysis we checked that the missing data (3.60%) were completely randomly distributed and imputed them. prof. = professional; FASD = Fetal Alcohol Spectrum Disorders; * = at the time of testing.
False Discovery Rate (FDR) adjusted *p*-value. Italic *p*-value = *p*-value ≤ 0.05 before FDR adjustment. Bold *p*-value = *p*-value ≤ .05 after FDR adjustment.

supplementary material Table 5), allowing direct comparison of children with FASD and SCG on these measures.

There were no significant mean differences in heterogeneity measures (scatter and variance) between the FASD subgroups and the SCG all $p_{adj} > .05$ (Table 4). The results were similar for normative analyses (supplementary material Table 6) or comparisons with the standardization sample, except for a single difference on scaled scores scatter between the WISC-IV subgroup and the standardization sample ($p_{adj} = .04$, $d = 0.35$) (supplementary material Table 7). There were no significant associations between

measures of heterogeneity and sociodemographic or clinical factors such as associated NDD diagnoses (all $p_{adj} > .05$) (supplementary material Tables 8 and 9**)**.

General population contrast

There were significant differences between the FASD and SCG groups for all scaled and composite scores (all $p_{adj} < .05$), except for the Similarities subtest in the WISC-IV subgroup ($p_{adj} = .17$). Effect sizes were large (Cohen's $d > 1$) for the FSIQ, Working Memory Index, Verbal Working memory subtests (Letter-Number Sequencing for the WISC-IV subgroup, and Digit Span for WISC-V and pooled groups), and Coding, with Cohen's *d* ranging from −1.09 to −1.35 (Table 5). Results were similar when comparing children with FASD to standardization norms or with normative analysis (supplementary material Table 10).

Intra-subject contrast: strengths and weaknesses

We found significant strengths on the Similarities, Picture Concept subtests, and the Verbal Comprehension Index only in the WISC-IV subgroup. Conversely we found significant weaknesses on the Letter-Number Sequencing and Coding subtests in the WISC-IV subgroup, on the Digit Span subtest in the WISC-V subgroup and pooled group, and on the Working Memory Index for both sub- and pooled groups (all $p_{adj} < .05$) (Table 6). Consistently, normative analyses showed significant weaknesses on the Coding subtest and the Working Memory Index in the pooled group (supplementary material Table 11).

Global intellectual functioning vs. elementary reasoning abilities

Descriptively, children with FASD were 1.4 times more likely to be classified within the borderline to deficient ranges on the FSIQ than on the GAI, and 2.25 times more likely than on the HRSS. The classification (deficient, borderline, normal) was the same for all criteria (FSIQ, GAI, and HRSS) in 51 subjects (58%). Of the 37 (42%) remaining subjects, 12 (13.6%) had an FSIQ in the borderline range but a GAI and an HRSS in

Table 4. Scatter and variance profiles mean comparisons between fetal alcohol spectrum disorders and simulated comparison groups.

	WISC-IV					WISC-V				
Scale	SCG *N* = 10,000	FASD *N* = 67				SCG *N* = 10,000	FASD *N* = 40			
Scatter	*M* (*SD*)	*n*	*M* (*SD*)	*t*	*d*	*M* (*SD*)	*n*	*M* (*SD*)	*t*	*d*
Index	24.28 (10.49)	60	25.12 (11.54)	0.56*ns*	0.08	25.12 (9.72)	38	23.84 (9.92)	−0.79*ns*	−0.13
Subtest Variance	7.61 (2.07)	50	8.40 (2.67)	2.07*ns*	0.33	7.17 (1.97)	36	7.08 (1.56)	−0.32*ns*	−0.05
Index	139.29 (115.16)	60	152.52 (142.00)	0.72*ns*	0.10	117.50 (87.62)	38	109.28 (83.32)	−0.61*ns*	−0.10
Subtest	6.31 (3.20)	50	7.80 (5.50)	1.92*ns*	0.33	5.60 (2.88)	36	5.58 (2.82)	−1.06*ns*	−0.17

Note. WISC-IV = Wechsler Intelligence Scale for Children 4th edition groups; WISC-V = WISC 5th edition groups; SCG = Simulated Comparison Groups; FASD = Fetal Alcohol Spectrum Disorders; M = Mean; SD = Standard Deviation; t = t statistic from student's t-test; d = cohen's d.
False discovery rate adjusted *p*-value: *ns* $p_{adj} > .05$; * $p_{adj} \leq .05$; ** $p_{adj} < .01$; *** $p_{adj} < .001$.

Table 5. Scaled and composite scores mean comparisons between fetal alcohol spectrum disorders and simulated comparison groups.

	WISC-IV N_{FASD} = 67			WISC-V N_{FASD} = 40			WISC-IV & V N_{FASD} = 107		
Scale	*n*	*t*	*d*	*n*	*t*	*d*	*n*	*t*	*d*
Indexes									
FSIQ	54	−7.85***	−1.13	39	−7.86***	−1.26	93	−11.12***	−1.20
Verbal comprehension	64	−4.55***	−0.63	40	−5.79***	−0.92	–	–	–
Perceptual reasoning	64	−8.45***	−1.08	–	–	–	–	–	–
Visual spatial	–	–	–	40	−6.64***	−0.95	–	–	–
Fluid reasoning	–	–	–	40	−5.22***	−0.83	–	–	–
Working memory	62	−9.32***	−1.27	38	−8.34***	−1.35	100	−12.45***	−1.30
Processing speed	64	−7.99***	−1.07	39	−5.06***	−0.90	103	−9.38***	−1.01
Subtests									
Similarities	66	−1.38*ns*	−0.19	39	−5.82***	−0.90	–	–	–
Vocabulary	65	−5.36***	−0.73	39	−4.82***	−0.80	104	−7.18***	−0.76
Comprehension	63	−5.49***	−0.65	–	–	–	–	–	–
Block design	62	−7.41***	−0.97	39	−5.81***	−0.86	101	−9.38***	−0.93
Visual puzzles	–	–	–	39	−5.98***	−0.86	–	–	–
Picture concept	64	−4.48***	−0.60	–	–	–	–	–	–
Matrix reasoning	65	−6.59***	−0.84	39	−4.22***	−0.69	104	−7.95***	−0.80
Figure weight	–	–	–	39	−4.66***	−0.74	–	–	–
Digit span	65	−6.86***	−0.92	39	−8.49***	−1.38	104	−10.36***	−1.09
Letter-number seq.	56	−8.30***	−1.15	–	–	–	–	–	–
Picture span	–	–	–	37	−7.68***	−1.13	–	–	–
Coding	64	−8.81***	−1.20	39	−6.01***	−1.05	103	−10.63***	−1.14
Symbol search	62	−5.54***	−0.76	38	−3.72***	−0.64	100	−6.70***	−0.72

Note. FSIQ = Full Scale Intelligence Quotient; seq. = sequencing; WISC-IV = Wechsler Intelligence Scale for Children 4th edition; WISC-V = WISC 5th edition; FASD = Fetal Alcohol Spectrum Disorders; t = t statistic from Student's t-test; d = Cohen's d; *SCG* = Simulated Comparison Group; . N_{SCG} = 10,000 for each group.
False discovery rate adjusted *p*-value: *ns* $p_{adj} > .05$; * $p_{adj} \leq .05$; ** $p_{adj} < .01$; *** $p_{adj} < .001$.

the normal range and 9 (10.23%) had an FSIQ and a GAI in the borderline to deficit range but an HRSS in the normal range, and 3 (3.4%) had an FSIQ in the deficit range, a GAI in the borderline range and an HRSS in the normal range (Table 7).

We then compared the proportions of subjects in the borderline, deficit, and borderline-to-deficit ranges between the SCG and children with FASD according to the different intellectual functioning criteria. The odds ratio (OR) of children with FASD classified in each range depending on the FSIQ, HRSS, or GAI were all > 1 (OR ranging from 3.68 to 12.71, all p_{adj} <.01) (Table 8). The OR of being classified as borderline-to-deficient was significantly different (approximately 1.7 times higher) between the FSIQ and GAI ($p_{adj} < .001$), and between the FSIQ and HRSS ($p_{adj} = .03$). Nonetheless, even though the HRSS reference was more conservative in terms of functioning (fewer patients were classified as borderline-to-deficient), this OR was not different between the GAI and HRSS ($p_{adj} = .81$). See supplementary material Figure 2 for an example of the profiles of six children with FASD, who were either consistently or inconsistently categorized using these three metrics.

Discussion

Our study found that children with FASD may not have more heterogeneous WISC profiles than the general population, contrary to what we first hypothesized. Yet it

Table 6. Strengths and weaknesses: means comparisons of differences between composite scores and mean composite score and of differences between scaled scores and mean scaled scores between fetal alcohol spectrum disorders and simulated comparison groups.

	WISC-IV				WISC-V				WISC-IV & V			
	SCG	FASD			SCG	FASD			SCG	FASD		
MCS & MSS	*M*(*SD*)	*M*(*SD*)	*t*	*d*	*M*(*SD*)	*M*(*SD*)	*t*	*d*	*M*(*SD*)	*M*(*SD*)	*t*	*d*
Indexes												
Verbal comprehension	−0.06 (9.66)	5.99 (10.62)	4.32***	0.58	−0.16 (9.91)	1.12 (9.43)	0.83*ns*	0.13	–	–	–	–
Perceptual reasoning	−0.01 (9.44)	0.26 (8.47)	0.25*ns*	0.03	–	–	–	–	–	–	–	–
Visual spatial	–	–	–	–	0.14 (8.74)	1.62 (7.77)	1.17*ns*	0.18	–	–	–	–
Fluid reasoning	–	–	–	–	−0.14 (8.76)	2.73 (9.51)	1.86*ns*	0.31	–	–	–	–
Working memory	−0.04 (10.31)	−5.39 (9.09)	−4.54***	−0.55	0.08 (8.97)	−4.69 (7.76)	−3.78**	−0.57	−0.06 (9.81)	−5.12 (8.56)	−5.82***	−0.55
Processing speed	0.00 (11.37)	−0.86 (11.47)	−0.58*ns*	−0.08	0.09 (11.75)	−0.77 (10.69)	−0.50*ns*	−0.08	0.01 (11.41)	−0.83 (11.11)	−0.74*ns*	−0.07
Subtests												
Similarities	−0.01 (2.12)	1.98 (2.58)	5.56***	0.84	−0.02 (2.12)	0.00 (1.87)	0.08*ns*	0.01	–	–	–	–
Vocabulary	0.03 (2.10)	0.40 (2.29)	1.15*ns*	0.17	−0.03 (2.32)	0.39 (2.16)	1.16*ns*	0.19	−0.01 (2.18)	0.40 (2.22)	1.67*ns*	0.18
Comprehension	0.01 (2.36)	0.42 (2.12)	1.36*ns*	0.18	–	–	–	–	–	–	–	–
Block design	−0.02 (2.40)	−0.24 (2.51)	−0.62*ns*	−0.09	0.05 (2.11)	0.17 (1.90)	0.39*ns*	0.06	0.04 (2.28)	−0.07 (2.27)	−0.44*ns*	−0.05
Visual puzzles	–	–	–	–	0.00 (2.00)	0.36 (1.65)	1.32*ns*	0.20	–	–	–	–
Picture concept	0.02 (2.50)	0.92 (2.52)	2.52*	0.36	–	–	–	–	–	–	–	–
Matrix reasoning	−0.01 (2.34)	0.02 (2.10)	0.10*ns*	0.01	0.00 (2.08)	0.59 (2.39)	1.48*ns*	0.26	0.01 (2.25)	0.26 (2.23)	1.02*ns*	0.11
Figure weight	–	–	–	–	−0.04 (2.23)	0.59 (2.22)	1.17*ns*	0.28	–	–	–	–
Digit span	0.00 (2.49)	−0.44 (2.27)	−1.38*ns*	−0.19	0.01 (2.19)	−1.53 (1.99)	−4.61***	−0.73	0.01 (2.38)	−0.89 (2.21)	−3.75***	−0.39
Letter-number seq.	−0.02 (2.35)	−1.28 (2.16)	−4.11***	−0.56	–	–	–	–	–	–	–	–
Picture span	–	–	–	–	0.02 (2.25)	−0.30 (1.67)	−1.14*ns*	−0.16	–	–	–	–
Coding	0.02 (2.61)	−1.72 (2.91)	−4.21***	−0.63	0.02 (2.56)	−0.86 (2.32)	−2.27*ns*	−0.36	0.00 (2.59)	−1.36 (2.70)	−4.66***	−0.52
Symbol search	−0.01 (2.51)	−0.06 (3.12)	−0.11*ns*	−0.02	0.01 (2.49)	0.59 (2.35)	1.48*ns*	0.24	0.00 (2.49)	0.21 (2.83)	0.68*ns*	0.08

Note. MCS = differences between composite scores and Mean Composite Score ; MSS = differences between scaled scores and Mean Scaled Scores; *seq.* = *Sequencing;* WISC-IV = Wechsler Intelligence Scale for Children 4th edition group; WISC-V = WISC 5th edition group; SCG = Simulated Comparison Group; FASD = Fetal Alcohol Spectrum Disorders; M = Mean; SD = Standard Deviation; d = Cohen's d; t = t statistic from student's t-test.

N_{SCG} = 10,000 for each group and analysis; N_{FASD} for the WISC-IV group = 60 for indexes and 50 for subtests; N_{FASD} for the WISC-*V* = 38 for indexes and 36 for subtests; N_{FASD} for the WISC-IV and WISC-V group = 98 for indexes and 86 for subtests.

False discovery rate adjusted *p*-value: *ns* $p_{adj} > .05$; * $p_{adj} \leq .05$; ** $p_{adj} < .01$; *** $p_{adj} < .001$.

confirmed the second assumption that these profiles would show a certain level of regularity among globally lower mean performance, subtests, and composite scores of working memory and processing speed emerging as significant weaknesses both at the mean level and in an excessive number of patients. The possibility of better verbal functioning was suggested by the assessment with the WISC-IV but not confirmed at the whole sample level. In line with our last proposition, our results showed that the proportion of children with FASD classified in the borderline to deficient ranges was approximately twice as large using the FSIQ compared to the GAI or the Highest Reasoning Scaled Score, with little group difference between the last two metrics, supporting that the choice of a clinically relevant measure of intellectual functioning is not only a theoretical but also a practical issue in this population.

No indication of excessive cognitive heterogeneity in FASD WISC profile

The absence of significant excess heterogeneity of the WISC profile in our FASD sample may be surprising given the description of FASD in the literature as highly heterogeneous in terms of cognitive impairment and associated functional diagnoses (Astley et al., 2009; Kodituwakku, 2007). In addition, a few studies have demonstrated heterogeneity in Wechsler profiles in samples of children with specific learning disorders using scaled scores scatter or discrepancies between composite scores (Poletti, 2016; Watkins, 2005). As we checked at the outset, this lack of significant heterogeneity cannot be directly explained by a shrinkage related to an overall lower level of functioning in the population. However, this overall lower level of functioning may still have masked the heterogeneity that a significant proportion of language and coordination developmental disorders would be expected to produce. Another non-exclusive and plausible explanation is that the Wechsler Intelligence Scales do not allow for a sufficiently accurate assessment of cognitive heterogeneity, since the range of cognitive domains is limited by tasks that focus mainly on multi-modal reasoning, working memory, and processing speed. Moreover, FASD-related executive dysfunction may induce variability in performance across WISC subtests, but not in a systematized or even systematic way (over time) which would be consistent with the Wechsler Intelligence Scales not being designed to be sensitive to or analytic in executive functioning. In any case, our result challenges the idea that the FSIQ score should not be used in cases of children with FASD because of an excessive heterogeneity of their WISC profile. In fact, the construct and prognostic validity of the FSIQ is not questioned by the profile heterogeneity (Daniel, 2007; Freberg et al., 2008; Watkins et al., 2007), but it is rather its diagnostic and functional significance that should be reconsidered in these clinical contexts.

Strengths and weaknesses revealed by FASD WISC profiles

In regard to group-level cognitive profiles, the strength observed on the Similarities subtest is consistent with a previous study conducted on a large cohort of children with FASD that identified this same group-level characteristic (Streissguth et al., 1996) but not strictly with another study conducted by the same team that identified the

Table 7. Differential analysis of global intellectual functioning, broad and elementary reasoning abilities: classification as a function of general ability index, highest reasoning scaled score, and full scale intelligence quotient.

		GAI			HRSS		
	N	Deficit	Borderline	Normal	Deficit	Borderline	Normal
FSIQ	Deficit	**12**	8	0	**10**	5	5
	Borderline	1A	**12**	12	1B	**3**	21
	Normal	0	0	**43**	1C	0	**42**
			HRSS				
	N	Deficit	Borderline	Normal			
GAI	Deficit	**7**	4	2			
	Borderline	4	**4**	12			
	Normal	1C	0	**54**			

Note. FSIQ = Full Scale Intelligence Quotient; GAI = General Ability Index; HRSS = Highest Reasoning Scaled Score.
A Subject scores: GAI = 66, FSIQ = 71, Similarities (SI) = 5, Vocabulary (VC) = 3, Comprehension (CO) = 5, Block design (BD) = 5, Matrix reasoning (MR) = 5, Picture concept (PC) = 8, Digit span (DS) = 9, Letter-number sequencing (LN) = 9, Coding (CD) = 4, Symbol search (SS) = 8;B Subject scores: FSIQ = 71, GAI = 75, SI = 7, VC = 10, BD = 5, Visual puzzles (VP) = 7, MR = 5, Figure weight (FW) = 4, DS = 7, Picture span (PS) = 8, CD = 2, SS = 5;C Subject scores: FSIQ = 83, GAI = 83, SI = 7, VC = 11, BD = 7, VP = 5, MR = 7, FW = 6, DS = 5, PS = 7, CD = 11, SS = 6.
Bold values on diagonals in cross-tabulations represent numbers of congruently classified subjects based on two intelligence metrics.

Table 8. Differential analysis of global intellectual functioning, broad and elementary reasoning abilities: comparison of proportions.

	FASD *N* = 88	SCG *N* = 10,000	
	n (%)	*n* (%)	$OR^{p\text{-}value}$
FSIQ			
< 70	20 (22.73)	226 (2.26)	12.71***
70 ≤ FSIQ ≤ 79	25 (28.41)	633 (6.33)	5.87***
≥ 80	43 (48.86)	9141 (91.41)	1/OR = 11.13***
GAI			
< 70	13 (14.77)	228 (2.28)	7.43***
70 ≤ GAI ≤ 79	20 (22.73)	631 (6.31)	3.68***
≥ 80	55 (62.50)	9141 (91.41)	1/OR = 6.38***
HRSS			
≤ 7	12 (13.64)	160 (1.60)	9.70***
= 8	8 (9.09)	255 (2.55)	3.82**
> 8	68 (77.27)	9585 (95.85)	1/OR = 6.79***

Note. 1/OR represents the odds ratio (OR) of being classified in the borderline-to-deficit range.
FSIQ = Full Scale Intellectual Quotient; GAI = General Ability Index; HRSS = Highest Reasoning Scaled Score; FASD = Fetal Alcohol Spectrum Disorders; SCG = Simulated Comparison Group.
False discovery rate adjusted *p*-value from Fisher's exact tests: *ns* $p_{adj} > .05$; *$p_{adj} \leq .05$; **$p_{adj} < .01$; ***$p_{adj} < .001$.

Vocabulary and Comprehension subtests instead (Sampson et al., 1989), and many subsequent studies that found no verbal and nonverbal dissociation (Kodituwakku, 2007 for a review). Furthermore, we did not find this effect in the WISC-V subgroup without finding any explanation other than the test version. The significant weaknesses found in working memory and processing speed were consistent with previous work showing that the Arithmetic and Digit Span subtests were more impaired (Sampson et al., 1989; Streissguth et al., 1996), as well as the processing speed composite score (Dalen et al., 2009). The large proportion of children with ADHD in our sample was consistent with the pattern of cognitive weakness previously reported in FASD-unrelated ADHD samples (Thaler et al., 2013; Theiling & Petermann, 2016). Interestingly, the population with FASD-related ADHD was reported to have more impairment in the

working memory and processing speed domains than the unrelated population. Furthermore, their impairment in the processing speed domain was mostly due to the Coding subtest (Raldiris et al., 2018), which calls on executive functions, working memory, graphomotor skills (L. A. Jacobson et al., 2011), and procedural learning (Beljan et al., 2022). Together, these findings are consistent with the proposition that the executive deficit is a key feature of the cognitive profile of children with FASD (Fuglestad et al., 2015). Identifying a specific neurobehavioral profile for individuals with FASD has been a major area of research (Mattson & Riley, 2011). One goal of this quest is to aid the diagnosis of non-syndromic forms of FASD, where the causal relationship of the diagnosis remains probabilistic in the absence of key physical features of FAS. Despite the extensive research, a clear characterization of individuals with FASD has yet to be established. Of course, our findings, which are based solely on the Wechsler Intelligence Scales as opposed to a variety of neuropsychological tasks, do not provide a comprehensive answer to this ongoing debate, nor do they suggest any specificity of the WISC profile. However, they show that, on a WISC-based initial neuropsychological assessment, children with FASD are expected to present with a significant and suggestive impairment in working memory and processing speed.

Global cognitive functioning vs. elementary reasoning abilities in FASD

We found that all three selected measures, or at least proxies, of intellectual functioning consistently classified more than half the children with FASD in our sample, and that low scores were restricted to the mild to borderline range of severity. The remaining children, with few exceptions, were as expected classified in a higher functional level category when this intellectual categorization was based on their elementary reasoning abilities (HRSS and GAI) compared to the FSIQ. Similarly, the odds of being classified in the borderline and deficit ranges for a child with FASD compared to the general population were larger based on the FSIQ than on the GAI or HRSS but not between these last two metrics. Indeed, there were fewer children classified as borderline on the HRSS. Nonetheless, it can be argued that this range of functioning was narrower by construction due to the raw granularity of this metric. However, the HRSS showed relevant in distinguishing individuals with borderline and even deficient composite reasoning abilities (GAI) and well-preserved reasoning ability in at least one modality (see supplementary material Figure 2). Taken together, these findings strongly suggest that intellectual impairment defined as a global deficit in cognitive functioning may not be uncommon in FASD, but that a generalized deficit in elementary reasoning abilities is much less common in this population.

Several studies have questioned the relevance of FSIQ as a good proxy of intellectual functioning across the whole spectrum of neurodevelopmental disorders. In specific learning disorders, categories of neurodevelopmental disorders without intellectual impairment by definition, poorer performance in working memory and processing speed were shown to be associated with a decrease in the FSIQ and dissociation between the Cognitive Proficiency Index CPI (an index including only the working memory and processing speed subtests) and the GAI (Toffalini et al., 2017). These findings were consistent with the dissociation observed in a sample of individuals with ADHD between the GAI and the FSIQ (Theiling & Petermann, 2016) and the

underestimation of the intellectual functioning of children with ASD with FSIQ when compared to fluid reasoning ability measured by Raven's Progressive Matrices task (Dawson et al., 2007; Nader et al., 2016). Indeed, the debate suffers from a certain circularity since the populations where the different proxies of intellectual functioning were tested were often defined in the first place as with or without intellectual impairment, or occasionally characterized by measures of adaptive functioning which can be strongly affected even in the absence of intellectual impairment (Cornoldi et al., 2014; Koriakin et al., 2013). The relevance of the distinction between a global cognitive measure and more elementary reasoning ones to assess intellectual functioning seems to be restricted to mild or borderline situations, but they are unfortunately frequent. In the end, aside from theoretical considerations on intelligence, it appears clearly that what is at stake is the ability to characterize elementary reasoning abilities independently of instrumental or executive ones, and that this may require the use of at least several proxies of intellectual functioning, not alternatively of one or the other, nor a single one that is applied to all situations (Lanfranchi, 2013).

Clinical considerations

The first clinical lesson of our results is that the measurement and interpretation of heterogeneity should not be made a preliminary question to the analysis of a psychometric profile in children with FASD, since the intra-subject WISC variances and scatters failed to capture and anticipate both the weakness in working memory and processing speed and the dissociation between FSIQ and more elementary reasoning proxies. It is all the more debatable since the construct of the scatter makes it a rather gross measure of heterogeneity and normative data are not easily available for the variances. Second, we consider that our results support a bottom-up approach to these WISC profiles, in which some characteristics at the lower level (scaled scores or composite scores) condition not the validity but the interpretability (for instance functional or diagnostic meaning) of higher-level constructs such as the FSIQ.

Regarding intellectual functioning, it means that the identification of an HRSS, in the first instance, in the preserved or borderline range of functioning should prompt the computation of the GAI among other ancillary indices before any interpretation of the FSIQ in terms of intellectual impairment or intellectual development disorder. In the case of a dissociation between the FSIQ and the GAI or HRSS, one should first consider only the most parsimonious interpretation of these 3 potential proxies of the patient's intellectual functioning: FSIQ should be seen as a measure of global cognitive abilities, GAI as a measure of multimodal reasoning abilities, and HRSS as a measure of the best elementary reasoning ability within a specific modality (conceptual, inductive, sequential, visuo-spatial, quantitative). The aim of this approach is to seek for intellectual preservation even in a context of multiple instrumental, procedural or executive (attentional) impairment that prevents more comprehensive and regular access to intellectual resources. Of course, the latter interpretation should be taken very cautiously given the low robustness of scaled score measurements taken individually (McDermott et al., 1990; Watkins, 2003), but it is at least supported by the rather good consistency between the odds of being considered intellectually

impaired with either HRSS or GAI in this population. In the end, FSIQ, GAI, and/or HRSS within different ranges of functioning should of course prompt the search for multiple specific cognitive impairments by means of complementary neuropsychological investigations.

Regarding more specific cognitive domains such as instrumental, procedural, or working memory, strengths and weaknesses should be systematically searched for in both scaled and composite scores, with at least a 10th (90th resp.) percentile threshold, and then related to the patient's behavioral and adaptive complaints. The relevance of the scaled score level of analysis, and not only the composite score, is highlighted in our results by the difference between the Coding and Symbol Search subtests that may be more sensitive to executive or attentional impairment than the composite Processing Speed Index, advocating for more in depth investigation during further neuropsychological assessment, especially for executive and attentional domains.

That said, this sequential procedure does not question the value of the FSIQ construct nor its well established adaptive or prognostic correlates (Daniel, 2007; Freberg et al., 2008; Watkins et al., 2007). It can and perhaps should be calculated and interpreted in any case, but its interpretation as a proxy of intellectual functioning that may lead to a diagnosis of intellectual impairment or IDD must be taken cautiously within the multiple impairment or multiple NDD context of FASD.

Limitations

Our study shows several limitations due to the retrospective nature of the data collection. The first is the lack of a proper typically developing control group. The WISC 4th and 5th editions benefit from French standardization upon a large normative sample, and we were able to overcome the limitation of these norms by extending them to the metrics we wanted to use, thanks to computational simulation (Aubry & Bourdin, 2018), replicating a statistically sound strategy previously reported in several publications (Brooks, 2010, 2011; Brooks & Iverson, 2010; Crawford et al., 2007). Nevertheless, such a simulated comparison group is still an external reference, not fully controlled for confounding factors (e.g. socioeconomic status) nor assessed under similar conditions.

A second limitation of this study is the difference in mean scores that we observed between our two WISC version subgroups, for the Similarity subtest. Our analysis of the effects influencing performance on the different subtests did not enable us to explain this difference, except for an effect of the version of the scale. It should be noted that this task presents differences between the two French versions only in the items that compose it: only 9 out of 23 items are kept between the two versions, the semantic category that groups the word pairs remains the same for 14 items, but the principle of the task and the instructions remain the same. To our knowledge, a single study found a similar effect on the same subtest in a sample of children with ASD (Kuehnel et al., 2019). This difference cannot be explained by the Flynn effect either, because the difference between the groups is on a single subtest, and also because the VCI appears to capture the dimension (i.e. crystallized intelligence) that is the least sensitive to this effect (Pietschnig & Voracek, 2015). The most plausible explanation, even if unexpected considering the sample size, remains a difference in

the recruitment of these two samples on some parameter not captured by the variables we collected. In any case, our results still raise the question, leaving it open, of some degree of dissociation in favor of verbal functioning, at least at the group level, in the FASD population.

A third limitation to the scope of our results may be the choice of the 4-DDC as the principal diagnostic guideline (Astley & Clarren, 2000). Several studies have shown differences in diagnostic assignment between diagnostic systems (Coles et al., 2016, 2023). It is difficult to infer how different the results would have been using other guidelines. Some being more conservative (Coles et al., 2023) might have resulted in a more severe sample where the cognitive dissociation and the heterogeneity of the WISC profiles might have been greater. Yet we carefully assessed the diagnosis concordance with the revised Institute of Medicine guideline (Hoyme et al., 2016), which was high in our sample when diagnoses were grouped into two classes (FAS and non-syndromic FASD). This high consistency may have partly resulted from our medical practice of applying a minimal threshold of clinical relevance for prenatal alcohol exposure, as proposed by Hoyme et al. (2016) even if we otherwise followed the 4-DDC criteria. In any case, we consider that both this consistency analysis and the description we provide of the diagnostic procedure and clinical features make it possible to envisage the interest of our results beyond the choice of guidelines.

Finally, the lack of systematization of the neuropsychological scales used in our care settings did not allow us to collect systematic data on functioning in other specific domains, such as executive functions or adaptive functioning, which would have been very interesting to consider in our differential analysis of global cognitive functioning and broad and elementary reasoning abilities. In addition, even if the WISC is widely used in clinical settings as a first step in neuropsychological assessment, our findings are limited to this particular tool and step. The use of a different intelligence scale, and the addition of complementary tests covering a broader range of cognitive domains, might have shed a possibly different and at least more precise light on the issue of cognitive heterogeneity and intellectual functioning in FASD.

Conclusion

Our study provided a detailed analysis of the WISC-based cognitive profile in a large sample of children with FASD. WISC profile heterogeneity was not greater than that of the standardization sample. Performance was reduced along the whole cognitive profile but more so in the working memory and processing speed tasks. Intellectual assessment based on FSIQ led to a much higher frequency of classification in the deficit range than when more specific measures of reasoning were used, suggesting that at least in the FASD population, FSIQ should be considered as a measure of global cognitive functioning that may underestimate specific aspects of intellectual functioning, We therefore advocate the more systematic use of ancillary indices such as the GAI or the search for at least one preserved modality of elementary reasoning that reveals access to conceptualization and abstraction.

Disclosure statement

The authors have no competing interests to declare.

Ethics approval

The Ethical Board of Paris-Saclay University (CER-Paris-Saclay-2020-094) granted ethical approval for this study.

Funding

This study was supported by the French National Agency for Research under Grant ANR-19-CE17-0028-01; French National Institute for Public Health Research under Grant IRESP-19-ADDICTIONS-08.

ORCID

Eliot Kerdreux http://orcid.org/0000-0002-4071-2008
Justine Fraize http://orcid.org/0000-0001-6434-7992
Odile Boespflug-Tanguy http://orcid.org/0000-0003-2389-1232
Lucie Hertz-Pannier http://orcid.org/0000-0003-4952-7009
Marion Noulhiane http://orcid.org/0000-0003-2832-0332
Charlotte Pinabiaux http://orcid.org/0000-0002-4430-6410
David Germanaud http://orcid.org/0000-0001-5055-4624

Data availability statement

The data that support the findings of this study are available from the corresponding authors, E.K. or D.G., upon reasonable request.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5TM* (5th ed., p. xliv, 947). American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596

Astley, S. J., & Clarren, S. K. (2000). Diagnosing the full spectrum of fetal alcohol-exposed individuals: Introducting the 4-digit diagnostic code. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, *35*(4), 400–410. https://doi.org/10.1093/alcalc/35.4.400

Astley, S. J., Olson, H. C., Kerns, K., Brooks, A., Aylward, E. H., Coggins, T. E., Davies, J., Dorn, S., Gendler, B., Jirikowic, T., Kraegel, P., MaraviLla, K., & Richards, T. (2009). Neuropyschological and behavioral outcomes from a comprehensive magnetic resonance study of children with fetal alcohol spectrum disorders. *Canadian Journal of Clinical Pharmacology = Journal Canadien de Pharmacologie Clinique*, *16*(1), e178–e201.

Aubry, A., & Bourdin, B. (2018). Short forms of Wechsler scales assessing the intellectually gifted children using simulation data. *Frontiers in Psychology*, *9*, 830. https://doi.org/10.3389/fpsyg.2018.00830

Beljan, P., Bree, K. D., Gardner, J. M., Bald, S. C., & Davis, A. S. (2022). Understanding mathematical learning disorder in regard to executive and cerebellar functioning: A failure of procedural consolidation. *Journal of Pediatric Neuropsychology*, *8*(3), 116–122. https://doi.org/10.1007/s40817-022-00127-z

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Boseck, J. J., Davis, A. S., Cassady, J. C., Finch, W. H., & Gelder, B. C. (2015). Cognitive and adaptive skill profile differences in children with attention-deficit hyperactivity disorder with and without comorbid fetal alcohol spectrum disorder. *Applied Neuropsychology Child*, *4*(4), 230–236. https://doi.org/10.1080/21622965.2013.877392

Brooks, B. L. (2010). Seeing the forest for the trees: Prevalence of low scores on the Wechsler intelligence scale for children, fourth edition (WISC-IV). *Psychological Assessment*, *22*(3), 650–656. https://doi.org/10.1037/a0019781

Brooks, B. L. (2011). A study of low scores in Canadian children and adolescents on the Wechsler intelligence scale for children, fourth edition (WISC-IV). *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *17*(3), 281–289. https://doi.org/10.1080/09297049.2010.537255

Brooks, B. L., & Iverson, G. L. (2010). Comparing actual to estimated base rates of "abnormal" scores on neuropsychological test batteries: Implications for interpretation. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *25*(1), 14–21. https://doi.org/10.1093/arclin/acp100

Chasnoff, I. J., Wells, A. M., & King, L. (2015). Misdiagnosis and missed diagnoses in foster and adopted children with prenatal alcohol exposure. *Pediatrics*, *135*(2), 264–270. https://doi.org/10.1542/peds.2014-2171

Coles, C. D., Bandoli, G., Kable, J. A., del Campo, M., Suttie, M., & Chambers, C. D. (2023). Comparison of three systems for the diagnosis of fetal alcohol spectrum disorders in a community sample. *Alcohol, Clinical & Experimental Research*, *47*(2), 370–381. https://doi.org/10.1111/acer.14999

Coles, C. D., Gailey, A. R., Mulle, J. G., Kable, J. A., Lynch, M. E., & Jones, K. L. (2016). A comparison among 5 methods for the clinical diagnosis of fetal alcohol spectrum disorders. *Alcoholism, Clinical and Experimental Research*, *40*(5), 1000–1009. https://doi.org/10.1111/acer.13032

Cook, J. L., Green, C. R., Lilley, C. M., Anderson, S. M., Baldwin, M. E., Chudley, A. E., Conry, J. L., LeBlanc, N., Loock, C. A., Lutke, J., Mallon, B. F., McFarlane, A. A., Temple, V. K., & Rosales, T, Canada Fetal Alcohol Spectrum Disorder Research Network. (2016). Fetal alcohol spectrum disorder: A guideline for diagnosis across the lifespan. *CMAJ: Canadian Medical Association Journal = Journal de L'Association Medicale Canadienne*, *188*(3), 191–197. https://doi.org/10.1503/cmaj.141593

Cornoldi, C., Giofrè, D., Orsini, A., & Pezzuti, L. (2014). Differences in the intellectual profile of children with intellectual vs. Learning disability. *Research in Developmental Disabilities*, *35*(9), 2224–2230. https://doi.org/10.1016/j.ridd.2014.05.013

Crawford, J. R., Garthwaite, P. H., & Gault, C. B. (2007). Estimating the percentage of the population with abnormally low scores (or abnormally large score differences) on standardized neuropsychological test batteries: A generic method with applications. *Neuropsychology*, *21*(4), 419–430. https://doi.org/10.1037/0894-4105.21.4.419

Dalen, K., Bruarøy, S., Wentzel-Larsen, T., & Laegreid, L. M. (2009). Cognitive functioning in children prenatally exposed to alcohol and psychotropic drugs. *Neuropediatrics*, *40*(4), 162–167. https://doi.org/10.1055/s-0029-1243176

Daniel, M. H. (2007). "Scatter" and the construct validity of FSIQ: Comment on Fiorello. *Applied Neuropsychology*, *14*(4), 291–295. https://doi.org/10.1080/09084280701719401

Dawson, M., Soulières, I., Ann Gernsbacher, M., & Mottron, L. (2007). The level and nature of autistic intelligence. *Psychological Science*, *18*(8), 657–662. https://doi.org/10.1111/j.1467-9280.2007.01954.x

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews. Neuroscience*, *11*(3), 201–211. https://doi.org/10.1038/nrn2793

Fiorello, C. A., Hale, J. B., Holdnack, J. A., Kavanagh, J. A., Terrell, J., & Long, L. (2007). Interpreting intelligence test results for children with disabilities: Is global intelligence relevant? *Applied Neuropsychology*, *14*(1), 2–12. https://doi.org/10.1080/09084280701280338

Flak, A. L., Su, S., Bertrand, J., Denny, C. H., Kesmodel, U. S., & Cogswell, M. E. (2014). The association of mild, moderate, and binge prenatal alcohol exposure and child neuropsychological outcomes: A meta-analysis. *Alcoholism, Clinical and Experimental Research*, *38*(1), 214–226. https://doi.org/10.1111/acer.12214

Flanagan, D. P., & Kaufman, A. S. (2004). *Essentials of WISC-IV assessment*. John Wiley & Sons.

Fraize, J., Convert, G., Leprince, Y., Sylvestre-Marconville, F., Kerdreux, E., Auzias, G., Lefèvre, J., Delorme, R., Elmaleh-Bergès, M., Hertz-Pannier, L., & Germanaud, D. (2023). Mapping corpus callosum surface reduction in fetal alcohol spectrum disorders with sulci and connectivity-based parcellation. *Frontiers in Neuroscience*, *17*, 1188367. https://doi.org/10.3389/fnins.2023.1188367

Fraize, J., Fischer, C., Elmaleh-Bergès, M., Kerdreux, E., Beggiato, A., Ntorkou, A., Duchesnay, E., Bekha, D., Boespflug-Tanguy, O., Delorme, R., Hertz-Pannier, L., & Germanaud, D. (2023). Enhancing fetal alcohol spectrum disorders diagnosis with a classifier based on the intracerebellar gradient of volumetric undersizing. *Human Brain Mapping*, *44*(11), 4321–4336. https://doi.org/10.1002/hbm.26348

Fraize, J., Garzón, P., Ntorkou, A., Kerdreux, E., Boespflug-Tanguy, O., Beggiato, A., Delorme, R., Hertz-Pannier, L., Elmaleh-Berges, M., & Germanaud, D. (2023). Combining neuroanatomical features to support diagnosis of fetal alcohol spectrum disorders. *Developmental Medicine and Child Neurology*, *65*(4), 551–562. https://doi.org/10.1111/dmcn.15411

Freberg, M. E., Vandiver, B. J., Watkins, M. W., & Canivez, G. L. (2008). Significant factor score variability and the validity of the WISC-III Full Scale IQ in predicting later academic achievement. *Applied Neuropsychology*, *15*(2), 131–139. https://doi.org/10.1080/09084280802084010

Fuglestad, A. J., Whitley, M. L., Carlson, S. M., Boys, C. J., Eckerle, J. K., Fink, B. A., & Wozniak, J. R. (2015). Executive functioning deficits in preschool children with fetal alcohol spectrum disorders. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *21*(6), 716–731. https://doi.org/10.1080/09297049.2014.933792

Geier, D. A., & Geier, M. R. (2022). Fetal alcohol syndrome and the risk of neurodevelopmental disorders: A longitudinal cohort study. *Brain & Development*, *44*(10), 706–714. https://doi.org/10.1016/j.braindev.2022.08.002

Green, C. R., Mihic, A. M., Nikkel, S. M., Stade, B. C., Rasmussen, C., Munoz, D. P., & Reynolds, J. N. (2009). Executive function deficits in children with fetal alcohol spectrum disorders (FASD) measured using the Cambridge neuropsychological tests automated battery (CANTAB). *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *50*(6), 688–697. https://doi.org/10.1111/j.1469-7610.2008.01990.x

Guilmette, T. J., Sweet, J. J., Hebben, N., Koltai, D., Mahone, E. M., Spiegler, B. J., Stucky, K., & Westerveld, M, Conference Participants. (2020). American academy of clinical neuropsychology consensus conference statement on uniform labeling of performance test scores. *Clinical Neuropsychologist*, *34*(3), 437–453. https://doi.org/10.1080/13854046.2020.1722244

Hoyme, H. E., Kalberg, W. O., Elliott, A. J., Blankenship, J., Buckley, D., Marais, A.-S., Abdul-Rahman, O., Jewett, T., Coles, C. D., Chambers, C., Jones, K. L., Adnams, C. M., Shah, P. E., Riley, E. P., Charness, M. E., Warren, K. R., & May, P. A. (2016). *Updated clinical guidelines for diagnosing fetal alcohol spectrum disorders*, *Pediatrics, 138*(2), 20.

Jacobson, J. L., Akkaya-Hocagil, T., Ryan, L. M., Dodge, N. C., Richardson, G. A., Olson, H. C., Coles, C. D., Day, N. L., Cook, R. J., & Jacobson, S. W. (2021). Effects of prenatal alcohol exposure on cognitive and behavioral development: Findings from a hierarchical meta-analysis of data from six prospective longitudinal U.S. cohorts. *Alcoholism, Clinical and Experimental Research*, *45*(10), 2040–2058. https://doi.org/10.1111/acer.14686

Jacobson, L. A., Ryan, M., Martin, R. B., Ewen, J., Mostofsky, S. H., Denckla, M. B., & Mahone, E. M. (2011). Working memory influences processing speed and reading fluency in ADHD. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *17*(3), 209–224. https://doi.org/10.1080/09297049.2010.532204

Kautz-Turnbull, C., & Petrenko, C. L. M. (2021). A meta-analytic review of adaptive functioning in fetal alcohol spectrum disorders, and the effect of IQ, executive functioning, and age. *Alcoholism, Clinical and Experimental Research*, *45*(12), 2430–2447. https://doi.org/10.1111/acer.14728

Kodituwakku, P. W. (2007). Defining the behavioral phenotype in children with fetal alcohol spectrum disorders: A review. *Neuroscience and Biobehavioral Reviews*, *31*(2), 192–201. https://doi.org/10.1016/j.neubiorev.2006.06.020

Koriakin, T. A., Mccurdy, M. D., Papazoglou, A., Pritchard, A. E., Zabel, T. A., Mahone, E. M., & Jacobson, L. A. (2013). Classification of intellectual disability using the Wechsler intelligence scale for children: Full scale IQ or general abilities index? *Developmental Medicine and Child Neurology*, *55*(9), 840–845. https://doi.org/10.1111/dmcn.12201

Kuehnel, C. A., Castro, R., & Furey, W. M. (2019). A comparison of WISC-IV and WISC-V verbal comprehension index scores for children with autism spectrum disorder. *Clinical Neuropsychologist*, *33*(6), 1127–1137. https://doi.org/10.1080/13854046.2018.1503721

Lanfranchi, S. (2013). Is the WISC-IV general ability index a useful tool for identifying intellectual disability? *Developmental Medicine and Child Neurology*, *55*(9), 782–783. https://doi.org/10.1111/dmcn.12210

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 18. https://doi.org/10.18637/jss.v025.i01

Lecerf, T., Reverte, I., Coleaux, L., Favez, N., & Rossier, J. (2010). Indice d'aptitude général pour le WISC-IV: Normes francophones. *Pratiques Psychologiques*, *16*(1), 109–121. https://doi.org/10.1016/j.prps.2009.04.001

Lewis, C. E., Thomas, K. G. F., Dodge, N. C., Molteno, C. D., Meintjes, E. M., Jacobson, J. L., & Jacobson, S. W. (2015). Verbal learning and memory impairment in children with fetal alcohol spectrum disorders. *Alcoholism, Clinical and Experimental Research*, *39*(4), 724–732. https://doi.org/10.1111/acer.12671

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.2307/2290157

Mattson, S. N., Crocker, N., & Nguyen, T. T. (2011). Fetal alcohol spectrum disorders: Neuropsychological and behavioral features. *Neuropsychology Review*, *21*(2), 81–101. https://doi.org/10.1007/s11065-011-9167-9

Mattson, S. N., & Riley, E. P. (2011). The quest for a neurobehavioral profile of heavy prenatal alcohol exposure. *Alcohol Research & Health: The Journal of the National Institute on Alcohol Abuse and Alcoholism*, *34*(1), 51–55.

Mattson, S. N., Riley, E. P., Gramling, L., Delis, D. C., & Jones, K. L. (1997). Heavy prenatal alcohol exposure with or without physical features of fetal alcohol syndrome leads to IQ deficits. *Journal of Pediatrics*, *131*(5), 718–721. https://doi.org/10.1016/S0022-3476(97)70099-4

McDermott, P. A., Fantuzzo, J. W., & Glutting, J. J. (1990). Just say no to subtest analysis: A critique on Wechsler theory and practice. *Journal of Psychoeducational Assessment*, *8*(3), 290–302. https://doi.org/10.1177/073428299000800307

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

Nader, A.-M., Courchesne, V., Dawson, M., & Soulières, I. (2016). Does WISC-IV underestimate the intelligence of autistic children? *Journal of Autism and Developmental Disorders*, *46*(5), 1582–1589. https://doi.org/10.1007/s10803-014-2270-z

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909-2013). *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *10*(3), 282–306. https://doi.org/10.1177/1745691615577701

Poletti, M. (2016). WISC-IV intellectual profiles in italian children with specific learning disorder and related impairments in reading, written expression, and mathematics. *Journal of Learning Disabilities*, *49*(3), 320–335. https://doi.org/10.1177/0022219414555416

Popova, S., Lange, S., Shield, K., Mihic, A., Chudley, A. E., Mukherjee, R. A. S., Bekmuradov, D., & Rehm, J. (2016). Comorbidity of fetal alcohol spectrum disorder: A systematic review and meta-analysis. *Lancet (London, England)*, *387*(10022), 978–987. https://doi.org/10.1016/S0140-6736(15)01345-8

R Core Team. (2021). *R: A language and environment for statistical computing [Logiciel]*. R Foundation for Statistical Computing. https://www.R-project.org/

Raldiris, T. L., Bowers, T. G., & Towsey, C. (2018). Comparisons of intelligence and behavior in children with fetal alcohol spectrum disorder and ADHD. *Journal of Attention Disorders*, *22*(10), 959–970. https://doi.org/10.1177/1087054714563792

Rasmussen, C., Tamana, S., Baugh, L., Andrew, G., Tough, S., & Zwaigenbaum, L. (2013). Neuropsychological impairments on the NEPSY-II among children with FASD. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *19*(4), 337–349. https://doi.org/10.1080/09297049.2012.658768

Riley, E. P., & McGee, C. L. (2005). Fetal alcohol spectrum disorders: An overview with emphasis on changes in brain and behavior. *Experimental Biology and Medicine (Maywood, N.J.)*, *230*(6), 357–365. https://doi.org/10.1177/15353702-0323006-03

Sampson, P. D., Streissguth, A. P., Barr, H. M., & Bookstein, F. L. (1989). Neurobehavioral effects of prenatal alcohol: Part II. Partial Least Squares analysis. http://deepblue.lib.umich.edu/handle/2027.42/27790

Sattler, J., & Dumont, R. (2004). *Assessment of children: WISC-IV and WPPSI-III Supplement.* Jerome M. Sattler, Publisher.

Streissguth, A. P., Barr, H. M., Kogan, J., & Bookstein, F. (1996). *Understanding the occurrence of secondary disabilities in clients with fetal alcohol syndrome (FAS) and fetal alcohol effects (FAE): Final report*. University of Washington School of Medicine, Dept. of Psychiatry and Behavioral Sciences, Fetal Alcohol and Drug Unit. https://search.library.wisc.edu/catalog/9910018828302121

Streissguth, A. P., Barr, H. M., Sampson, P. D., Darby, B. L., & Martin, D. C. (1989). IQ at age 4 in relation to maternal alcohol use and smoking during pregnancy. *Developmental Psychology*, *25*(1), 3–11. https://doi.org/10.1037/0012-1649.25.1.3

Streissguth, A. P., Bookstein, F. L., Sampson, P. D., & Barr, H. M. (1989). Neurobehavioral effects of prenatal alcohol: Part III. PLS analyses of ne*uropsychologic tests*. http://deepblue.lib.umich.edu/handle/2027.42/27791

Thaler, N. S., Bello, D. T., & Etcoff, L. M. (2013). WISC-IV profiles are associated with differences in symptomatology and outcome in children with ADHD. *Journal of Attention Disorders*, *17*(4), 291–301. https://doi.org/10.1177/1087054711428806

Theiling, J., & Petermann, F. (2016). Neuropsychological profiles on the WAIS-IV of adults with ADHD. *Journal of Attention Disorders*, *20*(11), 913–924. https://doi.org/10.1177/1087054713518241

Tierney, N., & Cook, D. (2018). Expanding tidy data principles to facilitate missing data exploration, visualization and assessment of imputations. *Monash Econometrics and Business Statistics Working Papers*, Article 14/18. https://ideas.repec.org//p/msh/ebswps/2018-14.html

Toffalini, E., Giofrè, D., & Cornoldi, C. (2017). Strengths and weaknesses in the intellectual profile of different subtypes of specific learning disorder: A study on 1,049 diagnosed children. *Clinical Psychological Science*, *5*(2), 402–409. https://doi.org/10.1177/216770261667203

Tsang, T. W., Lucas, B. R., Carmichael Olson, H., Pinto, R. Z., & Elliott, E. J. (2016). Prenatal alcohol exposure, FASD, and child behavior: A meta-analysis. *Pediatrics*, *137*(3), e20152542. https://doi.org/10.1542/peds.2015-2542

Vaurio, L., Riley, E. P., & Mattson, S. N. (2008). Differences in executive functioning in children with heavy prenatal alcohol exposure or attention-deficit/hyperactivity disorder. *Journal of the International Neuropsychological Society: JINS*, *14*(1), 119–129. https://doi.org/10.1017/S1355617708080144

Watkins, M. W. (2003). IQ Subtest Analysis: Clinical Acumen or Clinical Illusion? The Scientific Review of Mental Health Practice: Objective Investigations of Controversial and Unorthodox Claims in Clinical Psychology. *Psychiatry, and Social Work*, *2*(2), 118–141.

Watkins, M. W. (2005). Diagnostic validity of Wechsler substest scatter. *Learning Disabilities: A Contemporary Journal*, *3*(2), 18–27.

Watkins, M. W., Glutting, J. J., & Lei, P.-W. (2007). Validity of the full-scale IQ when there is significant variability among WISC-III and WISC-IV factor scores. *Applied Neuropsychology*, *14*(1), 13–20. https://doi.org/10.1080/09084280701280353

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—Manual* (3th ed.). Psychological Corporation.
Wechsler, D. (2005a). *WISC-IV échelle d'intelligence de Wechsler pour enfants—Manuel d'administration et de cotation* (4th ed.). ECPA/Pearson.
Wechsler, D. (2005b). *WISC-IV échelle d'Intelligence de Wechsler pour enfants—Manuel d'interprétation* (4th ed.). ECPA/Pearson.
Wechsler, D. (2016). *WISC-V échelle d'intelligence de Wechsler pour enfants—Manuel d'interprétation* (5th ed.). ECPA/Pearson.
Weyrauch, D., Schwartz, M., Hart, B., Klug, M. G., & Burd, L. (2017). Comorbid mental disorders in fetal alcohol spectrum disorders: A systematic review. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *38*(4), 283–291. https://doi.org/10.1097/DBP.0000000000000440
World Health Organization. (2019). *international statistical classification of diseases and related health problems* (11th ed.).