Zulueta, Aitziber; Díaz-Orueta, Unai; Crespo-Eguilaz, Nerea; Torrano, Fermín
Virtual Reality-based Assessment and Rating Scales in ADHD Diagnosis

Psicología Educativa

Virtual Reality-based Assessment and Rating Scales in ADHD Diagnosis

Aitziber Zulueta[a], Unai Díaz-Orueta[b], Nerea Crespo-Eguilaz[c], and Fermín Torrano[a]

[a]Universidad Internacional de La Rioja, Logroño, Spain; [b]Maynooth University, Maynooth, Ireland; [c]Navarra University Clinic, Pamplona, Spain

ABSTRACT

The diagnosis of Attention Deficit Hyperactivity Disorder (ADHD) is mainly based on structured scales, such as Conners' ADHD Rating Scale (EDAH in the Spanish version) and DSM interviews. The use of technologies in neuropsychological assessment, such as the AULA virtual reality based test leads to more accurate diagnosis. The current study presents findings from analyzing the external validity of AULA and its contribution to the diagnosis of ADHD. Four hundred and seven children (272 girls and 135 boys) from 6 to 16 years old (213 with ADHD diagnosis, 105 inattentive children, 108 combined-type, and 194 controls) were evaluated. First, a factor analysis of AULA variables was conducted in order to reduce data to factor and 5 factors or components that account for 82.37% of the total variance were obtained from 407 subjects, namely, sustained attention, impulsivity control, processing speed, response variability, and control of motor activity. Second, a discriminant analysis was then performed on data obtained by participants from whom the five factors were obtained, showing that AULA presents moderate levels of both specificity and sensitivity. Finally, in order to study whether AULA adds relevant information in the diagnosis of ADHD, a cluster analysis was carried out, showing 4 clusters in the analysis of conglomerates with the control group and 6 groups of clusters in the ADHD group. In summary, AULA test shows adequate external validity, allows correct classification of children with and without attentional problems, and confirms and provides additional ADHD diagnostic information that it is essential for the design of interventions.

La evaluación mediante realidad virtual y las escalas de valoración en el diagnóstico del TDAH

RESUMEN

El diagnóstico del trastorno por déficit de atención y/o hiperactividad (TDAH) se basa principalmente en escalas estructuradas, como la escala de Conners (validada en España con el nombre de EDAH), y entrevistas basadas en los criterios recogidos en el DSM. El uso de las nuevas tecnologías en el proceso de evaluación neuropsicológica, como el Test de Atención de Realidad Virtual AULA, proporciona información rigurosa que conduce a diagnósticos más precisos. En este contexto, el objetivo del presente estudio es proporcionar evidencias acerca de la validez externa del test AULA, un test que mejora la precisión diagnóstica del TDAH. La muestra estuvo formada por 407 niños (272 niños y 135 niñas) de entre 6 y 16 años del norte de España (213 niños tenían un diagnóstico de TDAH, 105 desatentos, 108 de tipo combinado y 194 fueron niños controles). En primer lugar se realizó un análisis factorial exploratorio, que puso de manifiesto una estructura formada por cinco factores (atención sostenida, control de la impulsividad, velocidad de procesamiento, variabilidad de respuesta y control de la actividad motora), que son capaces de explicar el 82.37% de la varianza. En segundo lugar, se realizó un análisis discriminante, de acuerdo con el rendimiento obtenido por los participantes en los cinco factores obtenidos en el test AULA, que mostró niveles moderados de especificidad y sensibilidad. Finalmente, se llevó a cabo un análisis de conglomerados con el grupo control y seis grupos de clúster dentro del grupo de niños con TDAH. Los resultados muestran una adecuada validez externa del instrumento AULA, que permite realizar clasificaciones correctas de sujetos sanos y patológicos, confirmándose que proporciona información adicional para el diagnóstico de TDAH, esencial para el posterior diseño de intervenciones clínicas eficaces.

In recent years, research on the Attention Deficit Hyperactivity Disorder (ADHD) diagnosis is certainly 'in vogue'. Many popular books have been published, several companies are marketing new drugs for ADHD, and it constitutes a commonly used term among parents and teachers. Specialized services, including extra time for standardized assessment and psychoeducational intervention, are

Correspondence: aitziber.zulueta@unir.net (A. Zulueta).

afforded to children and adolescents diagnosed with ADHD, a fact that improves the likelihood of a more accurate ADHD diagnosis (Gualtieri & Johnson, 2005).

Attention Deficit Hyperactivity Disorder (ADHD) is defined by the DSM-5 as a persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development, presents symptoms in two or more settings (e.g., at home, school, or work; with friends or relatives; in other activities), and has a negative direct impact on social, academic, or occupational functioning. Several symptoms must be present before the age of 12. Depending on diagnostic criteria and procedures that are used, the percentage of children affected by ADHD worldwide differs critically. According to the meta-analysis performed by Willcutt (2012), the overall prevalence for ADHD ranged from 4% to 13.3% depending on the specific procedures used to integrate information from multiple raters and to measure functional impairment. These results clearly illustrate how prevalence estimates are sensitive to methodological differences. Moreover, the lack of clear guidelines or specific directives when it comes to perform an ADHD diagnosis may also explain these differences in prevalence, not only between countries but also between individual clinicians.

Diagnostic criteria are properly defined and specified both in the latest versions of the Diagnostic and Statistical Manual for Mental Disorders (DSM-IV, IV-TR, and 5) (APA, 1996, 2000, 2013) and in the International Classification of Diseases (ICD-10) (World Health Organization, 1992). In general terms, the diagnosis of ADHD is mainly based on clinical interviews with the patient and third parties' observations by means of structured scales such as the ADHD Rating Scale-IV (DuPaul, Power, Anastopoulos, & Reid, 1998) or EDAH for teachers (Farré & Narbona, 2010) as well as DSM-based interviews with parents. Rating scales are systematic and quantitative but, unfortunately, relying on reported symptoms or a third party observation is by definition not fully objective. Thus, while rating scales are seen as a necessary component of the diagnostic process, they are not sufficient for an accurate diagnosis. Faced with an ambiguous clinical picture with mixed ratings from parents and teachers, the practitioner is often tempted to refer children to specialists, usually clinical child psychologists that are well trained to diagnose ADHD and provide a full cognitive assessment (Gualtieri & Johnson, 2005).

Rating Scales

Clinicians may not perform daily life behavioral observations such as those collected by means of validated questionnaires. Hence, that information is not directly available for them and has to be provided by third parties who have this kind of interaction with the child. Thus, parents and teachers are the main information sources in relation to the child's behavioral disorders and their reports are necessary to establish an accurate ADHD diagnosis. Rating scales and clinical interviews are complementary instruments for ADHD diagnosis and their joint utilization allows assessing frequency and intensity of symptoms.

However, studies about the usefulness of the information provided by these informants about subjects with and without ADHD are not conclusive. Variability of these results seems to be related to specific diagnostic criteria, measures, and even to the geographical area in which the study was carried out (Polanczyk, De Lima, Horta, Biederman, & Rohde, 2007). Other factors associated to these differences are the type of sample (community-dwelling vs. clinical sample), sociocultural features (Blázquez-Almería et al. 2005), and the information source (parents, teachers, or children).

With regards to the type of sample and based on the ADHD subtype or clinical presentation, the frequency for the inattentive presentation or subtype has been reportedly higher in community samples, while combined presentation or subtype is more frequent in clinical samples. According to Capdevila-Brophy, Navarro-Pastor, Artigas-Pallarés, and Obiols-Llandrich (2007), boys are overrepresented in both subtypes, while girls, though being a minority, have more representation in the inattentive group. Additionally, children in the combined group are identified and treated earlier than those in the inattentive group. In this sense, a misdiagnosis bias has been reported based on gender (Bruchmüller, Margraf, & Schneider, 2012). While both over- and underdiagnosis of ADHD can result in harmful consequences for the patient, it is important to strive for an accurate diagnosis, which is based on accepted diagnostic criteria and not biased by other factors. In relation to this, there is an ongoing scientific and public debate in relation to the potential overdiagnosis of ADHD in children. Only a few empirical studies have addressed this important issue and, broadly, they have found a trend for potential overdiagnosis. Thus, it appears that therapists do not adhere strictly to diagnostic criteria, and that not all therapists seem to follow DSM or ICD requirements to base their diagnosis on a thorough evaluation of the relevant diagnostic criteria (Bruchmüller, 2012). The great problem with over- or underdiagnosis extends to the fact that mental disorders are often associated with negative perceptions, and these perceptions may have pernicious consequences (Eisenberg & Schneider, 2007; Narad et. al. 2015). Consequences of stigma may include lower self-esteem, increased feelings of hopelessness and isolation, and a reduced likelihood of seeking and obtaining care. Negative perceptions also represent a known concern in the specific context of ADHD in the childhood. Children's self-esteem and self-confidence have been shown to be highly sensitive to perceptions maintained regarding the disorder by family, teachers, and peers (Roe, 1998; Wheeler & Carlson, 1994). Furthermore, parents' and teachers' perceptions and expectations affect their interactions with other children, which can in turn affect children's behavior and academic success (Chi & Hinshaw, 2002). Several studies reveal teachers' and parents' expectations may create a self-fulfilling prophecy and even affect the level of post-secondary education a child would attain (Madon, Guyll, & Spoth, 2004).

In addition, discrepancies between information sources may add more confusion to the achievement of a proper ADHD diagnosis. For example, as Eisenberg and Schneider (2007) suggest, there is a variety of non-mutually exclusive reasons that explain why parents' perceptions of children diagnosed with ADHD are more negative, on average, compared to judgments performed by teachers. Parents may be less familiar with ADHD, and that can be one reason to truly hold more negative views about the disorder. Alternatively, children with a diagnosis of ADHD may behave more disruptively and perform assignments more poorly at home than at school, especially when medication effects are restricted to the time they spend at the educational environment. Finally, teachers are probably less likely to be aware of ADHD diagnoses (Cáceres & Herrero, 2011; Wolraich et al. Worley, 2004) although they tend to identify the disorder more often than parents (Sandberg, 2002).

However, each reporter provides unique and valid clinical information related to ADHD symptoms presentation that should not be ignored (Narad et al. 2015). For that reason, clinical utility of behavioral rating scales, in contrast to structured interviews, relies on a dimensional methodology that provides an assessment of behavior and emotional difficulties along a continuum from normal to abnormal, without a clear delimitation between the presence or absence of the disorder (Achenbach & McConaughy, 1996).

Neuropsychological Testing

Neuropsychological evaluation has progressively become a part of the protocol for an efficient approach to the understanding of ADHD (Holmes et al., 2010). Conventional "psychological" tests can indeed

capture the elements of inattention and impulsivity that characterize patients with ADHD. This kind of testing also identifies present cognitive disabilities that make the management of the disorder more complex. However, there are two major limitations with neuropsychological testing as a routine approach to ADHD diagnosis: one is related to expenses and the other is related to the fact that ordinary "psychoeducational" batteries reflect only an indirect measure of ADHD. Attention deficits, locomotor hyperactivity, and cognitive impulsivity may or may not be evidenced when a child is in a small room, face to face with a single adult (Gualtieri & Johnson, 2005). As a consequence, this type of evaluation by means of conventional neuropsychological batteries has suffered considerable criticism due to a low predictive value that leads to numerous false negatives (i.e., patients with ADHD wrongly discarded as not suffering the disorder, thus preventing them from receiving appropriate treatment), a lack of correlation between ADHD symptoms and neuropsychological deficits (Bolea-Almagnac et al. 2014; Epstein et al., 2003), a limited clinical use confined to individual descriptions (Lange et al. 2014), and the lack of ecological validity (Bolea-Almagnac et al. 2014; Gualtieri & Johnson, 2005; Parsons, 2016).

These criticisms have opened the door to new neuropsychological evaluation methods, such as computerized tests of attention. According to several authors (Bioulac et al., 2012; Díaz-Orueta et al., 2014; Etchepareborda, Paiva-Barón, & Abad, 2009), computerized tests increase a patient's interest and motivation for the task, making the assessment more entertaining and enjoyable. They allow a more precise control of time (both regarding the stimuli presentation and response) and the "experimenter effect" is reduced to its minimum. Moreover, they save time, significantly improve accuracy and speed of scoring, and allow registering a series of variables and scores that exceed human capacities (Conners, Epstein, Angold, & Klaric, 2003). All these features also tend to increase the opportunities for research (Etchepareborda et al. 2009).

In this sense, as many researchers remark (Epstein et al. 2011; Negut, Jurma, & David, 2017; Nolin et al., 2016; Parsey & Schmitter-Edgecombe, 2013) clinical application of Virtual Reality (VR) certainly provides new opportunities for assessment, such as customization for target populations, specific cognitive domains, and unique settings. VR scenarios allow for measurement of simulated everyday tasks in a safe and controlled environment and offer the opportunity to assess the influence of environmental stimuli (e.g., distractions, interruptions) on cognitive performance, which may provide a more ecologically valid assessment of everyday skills (Díaz-Orueta, 2017; Parsey & Schmitter-Edgecombe, 2013). With regard to ecological validity, there is a need to move beyond the limited generalizability of results found in many developed neuropsychology batteries to measures that more closely approach real-world functioning (Parsons, Carlew, Magtoto, & Stonecipher, 2017; Parsons, 2016). As these authors suggest, a more ecological approach to neuropsychological assessment implies moving from construct-driven assessments to tests that are really "representative" of real-world "functions" and provide with results that are "generalizable" for prediction of functional performance across a range of situations (function-led approach), namely, to capture the complexity of the response required in many multistep tasks found in everyday activities.

When it comes to VR applied to the Continuous Performance Test (CPT) (considered the most accurate tool to measure attentional processes) (Epstein et al., 2011; Negut et al., 2017; Negut, Matu, Sava, & David, 2017; Nolin et. al. 2016; Tarantino, Cutini, Mogentale, & Bisiacchi, 2013) it can be said that VR platforms promote the collection of additional cognitive and behavioral information about neuropsychological testing performance beyond data obtained through traditional assessments. CPTs conducted through VR seem to be better at detecting problems experienced by users because they emphasize the ability to be representative of people's functioning in everyday situations (Nolin et al. 2016).

However, a recent meta-analysis conducted by Negut et al. (2017) pointed out that VR measures are more complex and difficult because they replicate conditions similar to everyday life and, as a consequence, performance obtained on these tests is usually poorer than on classical measures. Thus, tasks embedded in VR may have an increased level of complexity and difficulty that requires additional cognitive resources. This could be one of the reasons to explain resistance to adopt additional measures that incorporate technology components to existing tests, rating scales, and questionnaires (Parsey & Schmitter-Edgecombe, 2013). The field of neuropsychology has made lower progress than what would be expected in comparison to other fields in terms of adopting technology and, hence, the potential for an effective use of technology has not yet been realized. A higher awareness with this regards would imply recognizing an adequate adoption of new technologies, which could result in a more comprehensive assessment of cognitive dysfunction and, in turn, the achievement of better informed diagnoses and treatments (Diaz-Orueta, 2017).

Objectives

The goal of the current study is double: 1) to test the external validity of the AULA test by means of identifying the core factors of the test and whether it provides valid measures of inattention, impulsivity, and hyperactivity and 2) to analyze the accuracy of AULA test to support diagnosis of ADHD in addition to subjective reports provided by third parties (i.e., parents and teachers). First, it is hypothesized that AULA will show clear factors associated to measures of inattention, impulsivity, processing speed, sustained attention or vigilance, and hyperactivity. Second, it is expected that the information provided by AULA will show to be more accurate in correctly identifying ADHD cases and discarding non-ADHD individuals than subjective reports provided by third parties (i.e., parents and teachers).

Method

Sample and Diagnostic Procedures

The final study sample comprised 407 children (272 boys and 135 girls) between 6 and 16 years old. Among them, 213 had an ADHD diagnosis: 105 were inattentive (72 boys and 32 girls) and 108 were combined-type (83 boys and 25 girls), while 194 were typically developing controls (116 boys and 78 girls). They were recruited from two private primary and secondary schools, a neuropsychology clinic, and a pediatric neurology clinic in northern Spain (see Table 1).

The clinical sample comprises 213 patients with an ADHD diagnosis that fulfill the following inclusion criteria: a) to attend outpatient services in a neuropediatric unit or at the psychopedagogical services, where a neuropsychological assessment is performed as part of the diagnostic process; b) to show an IQ within the normal limits (IQ > 80); and c) to show consent to participate in this study. Ethics approval was obtained for the development of the study and parents provided written informed consent on behalf of their children to let them participate.

ADHD diagnosis was made by a clinical diagnostic team who considered both data from parents' (DSM-IV-TR criteria) and teachers' ratings (EDAH for teachers) (Farré & Narbona, 2010), as well as clinical interviews with children and their parents. No significance was obtained between the information gathered from parents and and the information from teachers (McNemar = .125, p = .05). Children were considered to meet the criteria for a symptom domain or clinical presentation (inattention and/or hyperactivity/impulsivity) if the parent and teacher rating scale reported at least 6 non-overlapping symptoms in a particular domain. Children who met this criterion

Table 1. Descriptive Statistics and Significant Differences between Subtypes of ADHD

| | | | Control
n = 194 | ADHD n = 213 | | t Student
ADHD-C
ADHD-I |
				ADHD-C n = 108	ADHD-I n = 105	
Age (6-16)		M (SD)	9.08 (2.66)	9.78 (2.65)	10.62 (2.79)	ns
Sex - male		n (%)	116 (59.8%)	83 (76.9%)	73 (69.5%)	
Sex - female		n (%)	78 (40.2%)	25 (23.1%)	32 (30.5%)	
Total IQ		M (SD)	101.44 (10.55)	101.46 (10.77)	98.78 (10.16)	ns
		rank	83-128	80-132	80-125	
DSM5	DA frequency	M (SD)	2.71 (1.71)	6.68 (1.32)	6.54 (1.12)	ns
		rank	0-5	5-9	5-9	
	H/I frequency	M (SD)	1.88 (1.89)	5.96 (1.76)	1.98 (1.56)	p = .000
		rank	0-7	0-9	0-5	
EDAH	DA raw score	M (SD)	4.2 (3.33)	10.81 (1.93)	10.48 (2.13)	ns
		rank	0-13	4-15	2-15	
	H raw score	M (SD)	2.91 (3.38)	9.75 (2.14)	4.5 (2.89)	p = .000[1]
		rank	0-14	4-15	0-14	
	TC raw score	M (SD)	3.04 (4.5)	7.86 (4.94)	5.6 (4.5)	p = .001[2]
		rank	0-23	1-26	0-22	

Note. ADHD-C (combined subtype); ADHD-I (inattentive subtype); *ns* = no significant differences. [1]In items 1, 3 y 5; [2]In items 6, 15 y 16.

for both domains were inserted in the combined-type group, while children who met symptoms only for inattention were enrolled in the inattentive group. Those children with ADHD taking stimulant medications were included in the study but were required to be off medication for 48 hours prior to testing.

There were no statistically significant differences according to age in relation to the symptomatology observed by parents (χ^2 = 16.546, p = .085) and teachers (χ^2 = 15.437, p = .117). Moreover, there were no differences according to gender in the symptomatology objectified either by parents (χ^2 = 1.452, p = .228) or teachers (χ^2 = 499, p = .480). Hence, no analysis was performed using age or gender as covariates.

The gender ratio for the ADHD group of boys to girls is almost 3:1 (more specifically, 2.7 boys per girl). Epidemiological studies held in schools using gender- and age-validated questionnaires have shown a similar incidence (Cardó, Servera-Barceló, & Llobera, 2007).

Controls showed minimal symptoms of ADHD reported on the same parent and teacher rating scales and did not meet criteria for any other behavioral disorder. Participants with a full-scale score below IQ = 70 on the Wechsler Intelligence Scale for Children (WISC-IV) were excluded from the current analyses (Table 2).

Table 2. Descriptive Statistics of IQ and Subscales (Group with ADHD)

	N	Min.	Max.	M	SD
Total IQ	169	77	132	100.21	10.542
Verbal comprehension	136	73	141	100.98	11.929
Perceptual reasoning	135	74	131	101.67	11.203
Working memory	127	67	127	98.09	11.660
Processing speed	127	67	131	98.87	12.498

To carry out the second objective of the study only those subjects in the sample whose EDAH results were registered item by item were considered. Thus, 227 children were recruited, 72 with ADHD diagnoses (37 inattentive and 35 combined-type) and 155 controls.

Measures

All children were administered the AULA based neuropsychological test, designed for a comprehensive evaluation of attention processes; parents received the ADHD Rating Scale-IV (DuPaul et al. 1998) and EDAH (Farré & Narbona, 2010) rating scale was filled out by teachers.

In those clinical subjects a standard assessment procedure consisting of an open clinical interview with parents and a more extensive and thorough cognitive evaluation was also conducted.

ADHD Rating Scale. ADHD Rating Scale-IV was administered with diagnostic purposes to all parents of children in the sample. This measure includes 18 symptom criteria for a diagnosis of ADHD according to the DSM-IV-TR criteria (DuPaul et al., 1998).

Parents were asked about the frequency for each child symptom on a 0 to 3 Likert's rating scale: 0 = *never*, 1 = *sometimes*, 2 = *often*, and 3 = *very often*. To fulfill the criteria for a symptomatic ADHD diagnosis, six of nine symptoms occurring *often* or *very often* in one or both the subscales must be present.

EDAH Rating Scale. EDAH is a revised Spanish version of the Conners' Teacher Rating Scale-Revised. It is typified in primary school children (6-12 years) and it has proven to be very helpful in the assessment of ADHD. This rating scale shows excellent indicators for reliability and validity. EDAH allows collecting information on the typical child behavior at school. It provides a structured observation method for the teacher, comprising 20 items (derived from DSM criteria) easy to understand and requires a minimum investment of time for its completion. Each item is scored from 0 to 3, with higher scores indicative of the presence of symptoms.

EDAH comprises three subscales: inattention, hyperactivity/impulsivity, and behavioral problems. However, since AULA only measures attention and hyperactivity and tests that can potentially quantify EDAH items related to behavior problems could not be found, EDAH scale items related to behavior problems were not considered in the present study. Items that were finally analyzed and subareas to which they belong are listed below:

Hyperactivity (H): item 1, "excessive motor activity", item 3, "frequently annoys other children", item 5, "demands immediate satisfaction", item 13, "moves constantly, uneasy", and item 17, "impulsive and irritable".

Inattention (I): item 2, "has academic learning problems", item 4, "distracts easily, show poor attention", item 7, "has head up in the clouds", item 8, "does not end what he/she starts", and item 19, "frustrates his efforts easily, inconstant".

Teachers needed to answer to each item according to a Likert scale depending on the degree to which the child shows the behavior that is described: 1 = *none*, 2 = *little,* 3 = *quite*, and 4 = *a lot*. In order to make appropriate statistical analyses, scores were grouped in the

Table 3. Pearson Correlations between the 5 Factors of AULA and Attention Deficit, Hyperactivity, and Behavioral Problems Variables with EDAH in the Study Sample (*N* = 407)

	EDAH Raw scores		
	Attention deficit	Hyperactivity	Behavioral disorder
AULA factor			
Sustained attention	$r = .26, p < .01$	$r = .15, p < .01$	$r = .10, p < .05$
Impulse control	$r = .29, p < .01$	$r = .27, p < .01$	$r = .18, p < .01$
Processing speed/reaction time	$r = -.026$	$r = -.13, p < .01$	$r = -.06$
Variability in reaction time	$r = .30, p < .01$	$r = .22, p < .01$	$r = .15, p < .01$
Motor activity control	$r = .37, p < .01$	$r = .40, p < .01$	$r = .30, p < .01$

following categories: 0 = *shows no symptoms* (1 and 2) and 1 = *shows symptoms* (3 and 4).

AULA Virtual Reality Test. AULA (classroom in Spanish) is a VR based neuropsychological test, designed to comprehensively assess attention processes (Climent & Banterla, 2011) and support an ADHD diagnosis in children between 6 and 16 years old (Díaz-Orueta et al., 2014; Zulueta, Iriarte, Díaz-Orueta, & Climent 2013), with high test-rest reliability (Fernandez-Fernandez, Morillo-Rojas, & Alonso-Romero, 2012; Iriarte et al., 2016), sensitivity, and specificity (Rufo-Campos, Cueto, Iriarte, & Rufo-Muñoz, 2012).

It is a CPT paradigm based test with different tasks and distracting conditions, presented in the virtual scenario of a school classroom (Climent & Banterla, 2011). AULA is composed of two main exercises: a non-X paradigm based exercise ("press the button when you DO NOT see or hear apple") and an X paradigm based exercise ("press the button whenever you DO see or hear seven"). Stimuli are presented both on a visual basis (on the classroom's blackboard) and on an auditory basis (the patient listens to them with the earphones). At the same time, visual, auditory, or combined distractors of ecological nature (i.e., equal to those that may appear in a real classroom environment, such as an ambulance going by, a child coughing, or the teacher walking through the classroom) are presented.

The main variables measured by AULA are as follows:

- Omission errors (inattention): patients do NOT press the button when they should.
- Commission errors (impulsivity): patients press the button when they should NOT.
- Reaction time (processing speed): measured for correct answers and commission errors.
- Variability (standard deviation) in reaction time (sustained attention): changes in reaction time patterns during the test.
- Motor activity (hyperactivity): head movement, tracked with a movement sensor placed in the 3D glasses.

The following scores include total scores, scores comparing non-X versus X task, scores comparing visual vs. auditory performance (divided attention), scores with versus without distractors (interference of external distractors). There is also a final score on the quality of attention focus that compares number of errors performed when the blackboard is in the viewing angle versus when it is not in the viewing angle.

Data Analyses

Statistical analyses were performed with SPSS version 21.0 and correspond to the two mayor objectives of this study: to test the external validity of AULA and to estimate whether AULA test contributes to the diagnosis of ADHD providing additional information.

In order to carry out the first study goal, the external validity in AULA, a factorial analysis of cognitive variables contained in AULA was performed and, subsequently, factors obtained were correlated with behavioral variables in EDAH (using Pearson correlation). Moreover, a study on AULA sensitivity and specificity was provided.

In order to achieve the second goal (i.e., to provide valuable and complementary information to clinical diagnosis by means of AULA), comparisons (Mann Whitney's *U*) were held between performance obtained in 5 AULA factors and items of EDAH Inattention and Hyperactivity scales. Finally, followed by a cluster or *k*-means conglomerate analysis, different subgroups in which the study sample may be classified were provided, according to their performance in AULA.

Results

External Validity

First, a factor analysis of AULA variables was conducted in order to reduce data to factors. The measure of sampling adequacy, Kaiser-Meyer-Olkin (KMO = .903), and the measure of adequacy of the sample for each variable (MSA) are greater than .50 indicating a good fit of the data to a factorial analysis model. Thus, 5 factors or components that account for 82.37% of the total variance were obtained from 407 subjects and are called: sustained attention, impulsivity control, processing speed, response variability, and control of motor activity.

Then, the correlation between variables of inattention and hyperactivity of EDAH with the five factors obtained with AULA was calculated by obtaining Pearson correlation coefficients and, as expected, a significant correlation between all analyzed variables was found (see Table 3).

A discriminant analysis was then performed on results obtained by participants from whom the five factors were obtained in AULA. In this way, it could be seen how AULA classifies correctly the study participants according to their membership to either the group with or without ADHD. AULA presents a moderate degree of specificity, identifying 75.3% of healthy participants, and a sensitivity of 68.1% of individuals with ADHD correctly identified (Table 4).

Table 4. Discriminant Analysis of AULA: Classification of all Participants (*N* = 407) according to Performance in the Factors in their Group Membership (Group with ADHD and Control Group)

	Performance in AULA VR factors	
	Controls *n* (%)	ADHD *n* (%)
Group membership according to DSM-5		
Controls	146 (75.3)[1]	48 (24.7)
ADHD	68 (31.9)	145 (68.1)[2]

Note. [1]Specificity; [2]sensitivity.

The current results provide support for the external validity of AULA, that classifies correctly study participants according to their membership to either the group with or without ADHD. That 24.7% of false positives detected, a percentage which AULA does not seem to properly discard as non ADHD, may be due to the possibility that AULA detects attention difficulties that are also linked to other disorders different from ADHD. In other words, AULA's ability to

Table 5a. Mean Differences using Mann-Whitney's *U* on Performance with AULA Based on the Presence/Absence of Each Item of EDAH Inattention Subscale (*N* = 227)

EDAH	AULA	Sustained attention	Impulse control	Processing speed/ Reaction time	Variability in reaction time	Motor activity control
Item #2	Mean ranks (No vs. Yes)	99.8 vs. 134.98	104.22 vs. 128.05	108.46 vs. 121.41	98.24 vs. 137.44	100.42 vs. 134.01
	Mann-Whitney's *U*	4,182	4,791.5	5,376	3,965.5	4,267
	z	-3.943	-2.672	-1.452	-4.395	-3.766
	p	.000	.008	.146	.000	.000
Item #4	Mean ranks (No vs. Yes)	97 vs. 129.71	104,86 vs. 121.99	106,54 vs. 120.34	98,01 vs. 128.71	96,61 vs. 130.09
	Mann-Whitney's *U*	4,535.5	5,416	5,604.5	4649.5	4,492.5
	z	-2.216	-1.651	-.088	-1.309	-3.651
	p	.000	.049	.113	.000	.000
Item #7	Mean ranks (No vs. Yes)	99.02 vs. 132.76	100.25 vs. 131.12	113.79 vs. 113.11	100.69 vs. 130.54	97 vs. 135.45
	Mann-Whitney's *U*	4,388.5	4,547.5	6,219	4,603.5	4,127.5
	z	-3.84	-3.513	-.077	-3.398	-4.376
	p	.000	.000	.939	.001	.000
Item #8	Mean ranks (No vs. Yes)	107.14 vs. 133.08	107.49 vs. 132.11	111.05 vs. 122.21	106.71 vs. 134.28	106.37 vs. 135.23
	Mann-Whitney's *U*	3,865	3,923.5	4,517.5	3,793	3,736.5
	z	-2.624	-2.490	-1.129	-2.789	-2.919
	p	.009	.013	.259	.005	.004
Item #19	Mean ranks (No vs. Yes)	107.18 vs. 130.55	104.65 vs. 138.19	110.98 vs. 119.10	103.29 vs. 142.29	103.62 vs. 141.30
	Mann-Whitney's *U*	3,749	3,321.5	4,390.5	3,091.5	3,147
	z	-2.328	-3.341	-.809	-3.886	-3.754
	p	.020	.001	.419	.000	.000

Table 5b. Mean Differences using Mann-Whitney's *U* on Performance with AULA Based on the Presence/Absence of Each Item of EDAH Hyperactivity Subscale (*N* = 227)

EDAH	AULA	Sustained attention	Impulse control	Processing speed/ Reaction time	Variability in reaction time	Motor activity control
Item #1	Mean ranks (No vs. Yes)	106.2 vs. 133.1	105.3 vs. 135.23	115.57 vs. 110.16	108.86 vs. 126.53	100.09 vs. 147.93
	Mann-Whitney's *U*	4,055	3,911.5	5,059.5	4,486	3,073.5
	z	-2.800	-3.119	-.564	-1.841	-4.984
	p	.005	.002	.573	.066	.000
Item #3	Mean ranks (No vs. Yes)	108.93 vs. 132.41	110.32 vs. 127.71	113.8 vs. 114.73	111.01 vs. 124.88	105.65 vs. 144.33
	Mann-Whitney's *U*	3,459	3,689	4,325	3,828	2,875
	z	-2.216	-1.651	-.088	-1.309	-3.651
	p	.027	.099	.930	.190	.000
Item #5	Mean ranks (No vs. Yes)	111.79 vs. 126.54	109.03 vs. 142.24	116.82 vs. 98	110.97 vs. 131.19	107,1 vs. 153.18
	Mann-Whitney's *U*	2,854.5	2,321	2,737	2,696.5	1,949
	z	-1.208	-2.719	-1.541	-1.655	-3.773
	p	.227	.007	.123	.098	.000
Item #13	Mean ranks (No vs. Yes)	108.62 vs. 126.22	105.12 vs. 136.79	114.81 vs. 107.54	107.6 vs. 129.29	103.36 vs. 142.08
	Mann-Whitney's *U*	3,991.5	3,400	4,426.5	3,820	3,103.5
	z	-1.754	-3.155	-.724	-2.16	-3.858
	p	.079	.002	.469	.031	.000
Item #17	Mean ranks (No vs. Yes)	112.25 vs. 121.07	107.91 vs. 138.62	115.51 vs. 107.9	109.28 vs. 133.08	1037.42 vs. 140.62
	Mann-Whitney's *U*	3,777	2,987	3,820.5	3,820	3,236.5
	z	-.806	-2.809	-.696	-2.176	-3.037
	p	.420	.005	.486	.030	.002

detect attention difficulties may not be only specific for ADHD but also for other learning or neuro-developmental disorders involving attention deficits. In terms of the 31.9% of false negatives (subjects who perform well in AULA despite their clinical diagnosis of ADHD), they may belong to a subgroup of ADHD patients who, in controlled or structured environments, may perform correctly.

Contribution of AULA to the Diagnosis of ADHD

The second objective of the study aims to investigate about what AULA provides for the diagnosis of ADHD in comparison to subjective reports provided by third parties. This analysis was performed with EDAH scores obtained item by item so only a part

of the sample that had these data (227 subjects, 72 with ADHD and 155 controls) was used.

First, pairs of groups were created based on the presence or absence of the symptom of each EDAH individual item taken from both attention deficit and hyperactivity subscales from the previously presented factor analysis (inattention, impulsivity, processing speed, variability in reaction time, and motor activity) (Tables 5a and 5b). Due to lack of normality of studied variables, a Mann Whitney's *U* test was applied to analyze these differences (*m* = mean ranks).

In relation to the presence or absence of the symptom of each EDAH individual item from attention deficit and hyperactivity subscales and performance on tasks measuring similar variables (AULA), the findings were the following: results for the items comprised

in EDAH inattention subscale showed significant differences in all variables except in processing speed (Table 5a). With regards to items comprised in EDAH hyperactivity subscale in Table 5b those items in which differences were observed are indicated. Thus, item 1, "excessive motor activity", showed significant differences in all variables except in processing speed; or item 13, "moves constantly, uneasy", showed significant differences in impulsivity, variability in reaction time, and motor activity.

To study whether AULA adds relevant information in the diagnosis of ADHD, a cluster analysis was performed. For this analysis the total sample of subjects (N = 407, 194 controls and 213 with ADHD) was again used. Thus, according to the performance obtained in AULA in the five variables, a cluster analysis with both groups was carried out. No statistically significant differences were found for performance in those 5 factors according to gender in any of the groups.

According to the performance in AULA for the five studied factors, we obtained 4 clusters in the analysis of conglomerates with the control group (Table 6a). As an example, healthy children with good performance form cluster 2 or subjects showing a performance below the mean form cluster 4. According to the performance in AULA, subjects with ADHD are distributed in 6 groups of clusters (Table 6b) (for example, cluster 3 comprises children with normal performance; cluster 6 is an extreme subgroup that presents deficits throughout all variables; cluster 1 is formed by children with impulsivity and hyperactivity; cluster 2 and 5 show a low performance in impulsivity). Tables 6c and 6d specify results obtained for ADHD subgroups of combined and inattentive subtypes, respectively.

Table 6a. Results from the Analysis of Conglomerates. Grouping of Control Group Subjects (n = 194) in 4 Clusters according to Performance in AULA Factors (T scores, M = 50, SD = 10)

	Clusters according to performance in AULA Control group			
	1	2	3	4
Sustained attention	**57.59**	47.39	52.31	**63.82**
Impulse control	45.11	44.39	**59.51**	**58.22**
Processing speed/Reaction time	**60.07**	48.64	47.89	**57.87**
Variability in reaction time	52.45	39.74	49.47	**63.51**
Motor activity control	48.52	36.82	48.12	**65.61**
n = 194	66	65	48	15

Note. Cluster 1 = mild inattention and quick responses.
Cluster 2 = average performance.
Cluster 3 = mild impulsiveness.
Cluster 4 = mild inattention and impulsiveness; quick responses; moderate hyperactivity.

Table 6c. Results from the Analysis of Conglomerates. Grouping of ADHD Subjects, Combined Subtype (n = 108) in 4 Clusters according to Performance in AULA Factors (T scores, M = 50, SD = 10)

	Clusters according to performance in AULA ADHD-C group			
	1	2	3	4
Sustained attention	**65.52**	54.18	**58.59**	47.43
Impulse control	**56.00**	51.47	**61.52**	**57.07**
Processing speed/Reaction time	**61.95**	55.31	45.14	**38.96**
Variability in reaction time	**66.02**	53.59	**57.51**	41.26
Motor activity control	**69.25**	45.52	**64.10**	51.00
n = 108	22	31	39	16

Note. Cluster 1 = moderate inattention and hyperactivity; mild impulsiveness; quick responses.
Cluster 2 = average performance.
Cluster 3 = mild inattention and impulsiveness; moderate hyperactivity.
Cluster 4 = mild impulsiveness and slowness.

Table 6d. Results from the Analysis of Conglomerates. Grouping of ADHD Subjects, Inattentive Subtype (n = 105) in 4 Clusters according to Performance in AULA Factors (T scores, M = 50, SD = 10)

	Clusters according to performance in AULA ADHD-I group			
	1	2	3	4
Sustained attention	**62.15**	48.95	**59.99**	**56.80**
Impulse control	54.78	51.84	**65.99**	50.29
Processing speed/Reaction time	**63.37**	44.49	**42.10**	55.76
Variability in reaction time	**62.16**	45.37	**59.07**	51.92
Motor activity control	**62.72**	48.64	**62.75**	40.63
n = 105	27	25	22	31

Note. Cluster 1 = mild inattention and impulsiveness; moderate hyperactivity; quick responses.
Cluster 2 = average performance.
Cluster 3 = mild inattention and impulsiveness; slowness; moderate hyperactivity.
Cluster 4 = mild inattention.

Discussion

Accurate ADHD diagnosis is crucial to appropriate treatment. In this section, we first discuss the implications of the current results for the external validity and the complementary diagnostic information that AULA provides. Then we examine the broader clinical and theoretical implications of these findings, we consider possible limitations, and conclude by highlighting several key directions for future research on ADHD.

Table 6b. Results from the Analysis of Conglomerates. Grouping of ADHD Subjects (n = 213) in 6 Clusters according to Performance in AULA Factors (T scores, M = 50, SD = 10)

	Clusters according to performance in AULA ADHD group					
	1	2	3	4	5	6
Sustained attention	**68.02**	49.45	50.26	**58.02**	**58.56**	**61.08**
Impulse control	54.11	**59.06**	49.92	48.93	**58.99**	**69.44**
Processing speed/Reaction time	**65.63**	**39.66**	48.45	**60.18**	52.57	**35.64**
Variability in reaction time	**67.61**	45.80	47.19	54.91	**57.03**	**63.33**
Motor activity control	**70.90**	53.66	37.69	48.23	**61.04**	**67.96**
n = 213	25	31	37	38	63	19

Note. Cluster 1 = inattention and moderate hyperactivity; quick responses.
Cluster 2 = mild impulsiveness and slowness.
Cluster 3 = average performance.
Cluster 4 = mild inattention and quick responses.
Cluster 5 = inattention, impulsiveness and mild hyperactivity.
Cluster 6 = mild inattention; slowness; impulsiveness and moderate hyperactivity.

First, in relation of external validity of AULA, it could be seen that AULA generally does well in classifying subjects participating in the study according to whether they belonged to the group of ADHD or to the non-ADHD group, showing a moderate degree of specificity and sensitivity.

Despite the false negatives identified (those clinically diagnosed of ADHD who perform well in AULA), this could suggest that they may belong to a subgroup of ADHD patients who, in controlled or structured environments, may perform correctly. This fact does not invalidate the diagnosis, as they constitute a group of children that potentially self-regulate better despite their ADHD. Hence, AULA can contribute to the diagnosis of ADHD but can also identify the presence of other type of attention difficulties that may be linked to different developmental or learning conditions. As these results contradict those obtained by Rufo et al. (2012), in which AULA features of sensitivity and specificity showed to be significantly higher, other factors or the effect of potential confounding variables, such as the influence IQ (especially for those children above 130), and comorbidities (other coexisting disorders which were absent in the study of Rufo et al., 2012, but which were not conclusively ruled out for this study) should be considered as reasons for this discrepancy, provided that, as stated above, behavioral subscales and items of teachers' EDAH and Conners' parent scales do not track cognition as AULA, and AULA only tracks hyperactivity in terms of movement (not in terms of disruption of home or class environment).

Second, it can be established that there is a correlation between the presentation of behavioral symptoms (in a moderate intensity) with low or medium performance on cognitive tasks that measure similar variables. Thus, the presence or absence of a symptom of inattention or impulsivity/hyperactivity significantly correlates with a low to moderate performance in cognitive tasks. This seems to hold even for other executive variables different from those affected in ADHD (as it is the case with all the inattention subscale items, or, for example, item 5 (i.e., "demands immediate satisfaction") correlates with impulsivity control and control of motor activity, thus, they are measuring the same thing. Therefore, in agreement with previous researches (Areces, Rodríguez, García, Cueli, & González-Castro, in press) it can be established that AULA confirms the diagnosis of ADHD and even anticipates difficulties of different etiology that is not specific of individuals with ADHD. These findings are consistent with other studies that highlight the importance of relying on rating scales for the assessment of ADHD (Charach, Chen, Hogg-Johnson, & Schachar, 2009; Parker & Corkum, 2016).

In relation to the complementary information provided by AULA for the diagnosis of ADHD and oriented to a cognitive behavioral intervention, AULA enhances the relevance of dimensions related about how to behave, to cope with tasks, or even how to face and solve problems in a way that is not provided by the behavioral diagnosis obtained with DSM scales and EDAH. However, as it has been observed in obtained clusters, groups are differentiated on the basis of the presence of hyperactivity and/or impulsivity. Thus, AULA will differentiate subgroups in an efficient way, that is, it will specify executive profiles not previously detected by mean of traditional diagnostic tools or procedures. In summary, AULA provides information about performance in different domains and discriminates between impulsivity and hyperactivity.

In this study, children below the threshold of symptoms for a clinical diagnosis of ADHD but showing inattention and hyperactivity problems demanding an intervention have been detected. In the control group, 15 children without an ADHD diagnosis showed a low performance in all variables, hence, they would require some kind of treatment. With regards to the group with a diagnosis of ADHD, data obtained with the application of AULA allows dividing subjects in one group that present impulsivity and another group showing hyperactivity, for which the target intervention will be different. In this sense, these findings will allow the improvement and the specification of intervention proposals or designs. Moreover,

the finding of a subgroup of ADHD diagnosed children who were able to show a good performance despite their diagnosis has to be reported. For this subgroup, formed by children who may be able to control themselves in highly structured environments and tasks, the prognosis seems positive.

In summary, AULA can reasonably confirm the diagnosis of ADHD and complement the information obtained by means of observation scales with regards to different ADHD subtypes or clinical presentations. Finally, the current study, in agreement with numerous previous researches (Areces et al., in press; Bioulac et al., 2012; Epstein et al., 2011; Etchepareborda et al., 2009; Negut et al., 2017; Nolin et al., 2016; Parsey & Schmitter-Edgecombe, 2013), provides also support for the value and ecological validity of technological-based measures in neuropsychological assessment.

The interpretation of these results must be seen in the context of the current study's limitations. There are limitations related to the sample group. First, the ADHD group included only two of the three ADHD subtypes. The ADHD predominantly hyperactive-impulsive subtype was not included separately but merged with the group of mixed or combined subtype. Nevertheless, excluding this subtype allows drawing conclusions across all three subtypes.

Other limitations refer to the control group. In this study, we only considered typically-developing children. We did not include a group with psychopathology. As a result, we are not able to assert if the intergroup differences are only due to the ADHD condition or to other learning or potential neurodevelopment conditions not previously identified, at least partially, for the individuals comprised in the control group.

To carry out the second objective of the study, only those subjects in the sample whose EDAH results were registered item by item were recruited. This resulted in a significant decrease in the number of subjects in this analysis and consequently, a decrease in the general representativeness.

Limitations in the extent to which virtual reality simulations reflect real experiences of individuals as well as limitations related to psychological concerns (influence of VR on cognitive performance, internal and ecological validity of tools, utility for all populations, lack of normative data properties for different clinical populations) must be also taken into account.

Finally, although the AULA may be a useful clinical tool for ADHD diagnosis and the identification of subtypes of attentional disorders in children, its usage does not seem to be vastly superior to other instruments. The sensitivity and specificity of this instrument in relation to other measures of attention, such as seen in the Continuous Performance Test, are similar (Negut et al., 2017). In this sense, the AULA VR becomes another effective possibility for evaluating attentional processes.

Technology development, including VR, is continuing to grow and new applications are definitely emerging for the field of neuropsychology. VR technology offers a great opportunity to work in real-world simulation scenarios by means of safe, attractive, and controlled environments (Nolin et. al. 2016; Tarantino et al. 2013). Neuropsychologists should continue to move beyond in research for new strategies to assess cognitive abilities that allow them to offer patients effective and high quality diagnosis services.

Future research may try to obtain different neuropsychological profiles for ADHD and other developmental disorders in order to help increase accuracy in differential diagnosis, not only between different ADHD subtypes, or between ADHD and healthy controls, but also between ADHD and other learning, behavioral or neurodevelopmental disorders that sometimes may coexist with ADHD and mask actual results by increasing or decreasing performance. More specifically, the role and impact of high and very high IQs, other learning disorders (such as dyslexia), and conditions related to processing speed that may be inappropriately diagnosed as ADHD will have to be included in future studies.

Conflict of Interest

The authors of this article declare no conflict of interest.

References

Achenbach, T. M., & McConaughy, S. H. (1996). Relations between DSM-IV and empirically based assessment. *School Psychology Review, 25,* 329-341.

American Psychiatric Association - APA. (1996). *Diagnostic and statistical manual of mental disorders (*4th ed.). Washington, DC: Author.

American Psychiatric Association - APA. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association - APA. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Areces, D., Rodríguez, C., García, T., Cueli, M., & González-Castro, P. (2016). Efficacy of a continuous performance test based on virtual reality in the diagnosis of ADHD and its clinical presentations. *Journal of Attention Disorders.* https://doi.org/10.1177/1087054716629711.

Bioulac, S., Lallemand, S., Rizzo, A., Philip, P., Fabrigoule, C., & Bouvard, M. P. (2012). Impact of time on task on ADHD patient's performances in a virtual classroom. *European Journal of Paediatric Neurology, 16,* 514-521. https://doi.org/10.1016/j.ejpn.2012.01.006

Blázquez-Almería, G., Joseph-Munné, D., Burón-Masó, E., Carrillo-González, C., Joseph-Munné, M., Cuyàs-Reguera, M., & Freile-Sánchez, R. (2005). Resultados del cribado de la sintomatología del trastorno por déficit de atención con o sin hiperactividad en el ámbito escolar mediante la escala EDAH. *Revista de Neurología, 41,* 586-590.

Bolea-Alamañac, B., Nutt, D. J., Adamou, M., Asherson, P., Bazire, S., Coghill, D., ... British Association of Psychopharmacology (2014). Evidence-based guidelines for the pharmacological management of attention deficit hyperactivity disorder: Update on recommendations from the British Association for Psychopharmacology. *Journal of Psychopharmacology, 28,* 179-203. https://doi.org/10.1177/0269881113519509

Bruchmüller, K., Margraf, J., & Schneider, S. (2012). Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. *Journal of Consulting and Clinical Psychology, 80* 128-138. https://doi.org/10.1037/a0026582

Cáceres, J., & Herrero, D. (2011). Cuantificación y análisis de la concordancia entre padres y tutores en el diagnóstico del trastorno por déficit de atención/hiperactividad. *Revista de Neurología, 52,* 527-535.

Capdevila-Brophy, C., Navarro-Pastor, J. B., Artigas-Pallarés, J., & Obiols-Llandrich, J. (2007). Complicaciones obstétricas y médicas en el trastorno de déficit atencional/hiperactividad (TDAH): ¿hay diferencias entre los subtipos? *International Journal of Clinical and Health Psychology, 7,* 679-695.

Cardó, E., Severa-Barceló, M., & Llobera, J. (2007). Estimación de la prevalencia del trastorno por déficit de atención e hiperactividad en población normal de la isla de Mallorca. *Revista de Neurología, 44,* 10-14.

Charach, A., Chen, S., Hogg-Johnson, S., & Schachar, R. J. (2009). Using the Conners' Teacher Rating Scale-Revised in school children referred for assessment. *Canadian Journal of Psychiatry, 54,* 232-241. https://doi.org/10.1177/070674370905400404

Chi, T., & Hinshaw, S. (2002). Mother-child relationships of children with ADHD. *Journal of Abnormal Child Psychology, 30*(4), 387-400. https://doi.org/10.1023/A:1015770025043

Climent, G., & Banterla, F. (2011). *AULA: evaluación ecológica de los procesos atencionales.* San Sebastian, Spain: Nesplora.

Conners, C. K., Epstein, J. N., Angold, A., & & Klaric, J. (2003). Continuous performance test performance in a normative epidemiological sample. *Journal of Abnormal Child Psychology, 31,* 555-562. https://doi.org/10.1023/A:1025457300409

Díaz-Orueta, U. (2017). Advances in neuropsychological assessment of attention. From initial computerized continous performance tests to AULA. In: R. L. Kane & T. D. Parsons (Eds.), *The Role of Technology in Clinical Neuropsychology* (pp. 103-142). Oxford, UK: Oxford University Press.

Díaz-Orueta, U., García-López, C., Crespo-Eguílaz, N., Sánchez-Carpintero, R., Climent, G., & Narbona, J. (2014). AULA virtual reality test as an attention measure: convergent validity with Conners' Continuous Performance Test. *Child Neuropsychology, 20,* 328-342. https://doi.org/10.1080/09297049.2013.792332

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD Rating Scale-IV: Checklists, norms, and clinical interpretation.* New York, NY: The Guilford Press. https://doi.org/10.1177/0734282905285792

Eisenberg, D., & Schneider, H. (2007). Perceptions of academic skills of children diagnosed with ADHD. *Journal of Attention Disorders, 10,* 390-397. https://doi.org/10.1177/1087054706292105

Epstein, J. N., Erkanli, A., Conners, C. K., Klaric, J., Costello, J. E., & Angold, A. (2003). Relations between continuous performance test performance measures and ADHD behaviours. *Journal of Abnormal Child Psychology, 31,* 543-554.

Epstein, J. N., Langberg, J. M., Rosen, P. J., Graham, A., Narad, M. E, Antonini, T. N., ... Altaye, M. (2011). Evidence for higher reaction time variability for children with ADHD on a range of cognitive tasks including reward and event rate manipulations, *Neuropsychology, 25,* 427-441. https://doi.org/10.1037/a0022155

Etchepareborda, M. C., Paiva-Barón, H., & Abad, L. (2009). Ventajas de las baterías de exploración neuropsicológica en el trastorno por déficit de atención/hiperactividad. *Revista de Neurología, 48*(Supl.2), 89-93.

Farré, A., & Narbona, J. (2001). *Escala para la evaluación del trastorno por déficit de atención con hiperactividad.* Madrid, Spain: TEA.

Fernández-Fernández, M., Morillo-Rojas, M., & Alonso-Romero, L. (2012). *Utilidad del estudio Aula Nesplora en la valoración del TDAH.* Communication presented at the XXXVI Annual Meeting of the Spanish Society of Pediatric Neurology. Santander, May 31-June 2.

Gualtieri, C. T., & Johnson, L. G. (2005). ADHD: Is objective diagnosis possible? *Psychiatry (Edgmont), 2*(11), 44-53.

Holmes, J., Gathercole, S. E., Place, M., Alloway, T. P., Elliott, J. G., & Hilton, K. A. (2010). The diagnostic utility of executive function assessments in the identification of ADHD in children. *Children and Adolescent Mental Health, 15,* 37-43. https://doi.org/10.1111/j.1475-3588.2009.00536.x

Iriarte, Y., Díaz-Orueta, U., Cueto, E., Irazustabarrena, P., Banterla, F., & Climent, G. (2016). AULA - advanced virtual reality tool for the assessment of attention: Normative study in Spain. *Journal of Attention Disorders, 20,* 542-568. https://doi.org/10.1177/1087054712465335

Lange, K. W., Hauser, J., Lange, K. M., Makulska-Gertruda, E., Takano, T., Takeuchi, & Tucha, O. (2014). Utility of cognitive neuropsychological assessment in attention-deficit/hyperactivity disorder. *ADHD Attention Deficit and Hyperactivity Disorders, 6,* 241-248. https://doi.org/10.1007/s12402-014-0132-3

Madon, S., Guyll, M., & Spoth, R. L. (2004). The self-fulfilling prophecy as an intrafamily dynamic. *Journal of Family Psychology, 18,* 459-469. https://doi.org/10.1037/0893-3200.18.3.459

Narad, M. E., Garner, A. A., Peugh, J.L., Tamm, L., Antonini, T. N., Kingery, K. M., ... Epstein, J. N. (2015). Parent–teacher agreement on ADHD symptoms across development. *Psychological Assessment, 27,* 239-248. https://doi.org/10.1037/a0037864

Negut, A., Jurma, A. M., & David, D. (2017). Virtual-reality-based attention assessment of ADHD: ClinicaRV: Classroom-CPT versus a traditional continuous performance test, *Child Neuropsychology, 23,* 692-712.

Negut, A., Matu, S. A., Sava, F. A., & David, D. (2016). Virtual reality measures in neuropsychological assessment: A meta-analytic review. *The Clinical Neuropsychologist, 30,* 165-184. https://doi.org/10.1080/13854046.2016.1144793

Nolin, P., Stipanicic, A., Henry, M., Lachapelle, Y., Lussier-Desrochers, D., & Allain, P. (2016). ClinicaVR: Classroom-CPT: A virtual reality tool for assessing attention and inhibition in children and adolescents. *Computers in Human Behavior, 59,* 327-333. https://doi.org/10.1016/j.chb.2016.02.023

Parker, A., & Corkum, P. (2016). ADHD Diagnosis: As simple as administering a questionnaire or a complex diagnostic process? *Journal of Attention Disorders, 20,* 478-486. https://doi.org/10.1177/1087054713495736

Parsey, C. M., & Schmitter-Edgecombe, M. (2013). Applications of technology in neuropsychological assessment. *The Clinical Neuropsychologist, 27,* 1328-1361. https://doi.org/10.1080/13854046.2013.834971

Parsons, T. D., Carlew, A. R., Magtoto, J., & Stonecipher, K. (2017). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychological Rehabilitation, 37,* 777-807. https://doi.org/10.1080/09602011.2015.1109524

Parsons, T. D. (2016). *Clinical neuropsychology and technology: What's new and how we can use it.* New York, NY: Springer Press.

Polanczyk, G., De Lima, M.S., Horta, B.L., Biederman, J., & Rohde, L.A. (2007). The worldwide prevalence of ADHD: A systematic review and metaregression analysis. *American Journal of Psychiatry, 164,* 942-948. https://doi.org/10.1176/ajp.2007.164.6.942

Roe, D. (1998). Young children with attention difficulties: How can we help? *AECR Research and Practice Series, 5,* 1-23.

Rufo-Campos, M., Cueto, E., Iriarte, Y., & Rufo-Muñoz, M. (2012). *Estudio de sensibilidad de un nuevo método diagnóstico para el TDAH: Aula Nesplora.* Communication presented at the XXXVI Annual Meeting of the Spanish Society of Pediatric Neurology. Santander, May 31-June 2.

Sandberg, S. (2002). *Hyperactivity and attention disorders of childhood.* Cambridge, UK: Cambridge University Press.

Tarantino, V., Cutini, S., Mogentale, C., & Bisiacchi, P. S. (2013). Time-on-task in children with ADHD: An ex-Gaussian analysis. *Journal of the International Neuropsychological Society, 19,* 1-9. https://doi.org/10.1017/S1355617713000623

Wheeler, J., & Carlson, C. L. (1994). The social functioning of children with ADD with hyperactivity and ADD without hyperactivity: A comparison of their peer relations and social deficits. *Journal of Emotional and Behavioral Disorders, 2,* 2-13. htpps://doi.org/10.1177/106342669400200101

Willcutt, E.G. (2012). The prevalence of DSM-IV attention-deficit/hyperactivity disorder: A meta-analytic review. *Neurotherapeutics, 9,* 490-499. https://doi.org/10.1007/s13311-012-0135-8

Wolraich, M. L., Lambert, E. W., Bickman, L., Simmons, T., Doffing, M. A., & Worley, K. A. (2004). Assessing the impact of parent and teacher

agreement on diagnosing attention-deficit hyperactivity disorder. *Journal of Developmental and Behavioral Pediatrics, 25*, 41-47.

World Health Organization. (1992). *The international classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines* (10th ed.). Geneva, Switzerland: World Health Organization.

Zulueta, A., Iriarte, Y., Díaz-Orueta, U., & Climent, G. (2013). AULA Nesplora: avance en la evaluación de los procesos atencionales. Estudio de la validación convergente con el test de percepción de diferencias "caras" (versión ampliada). *ISEP Science, 4*, 4-10.