Papeles del Psicólogo, 2016. Vol. 37(1), pp. 3-13
http://www.papelesdelpsicologo.es

ADVANCES IN PSYCHOLOGICAL ASSESSMENT

Introduction

Eduardo Fonseca-Pedrero[1,3] and José Muñiz[2,3]
[1]*Universidad de La Rioja.* [2]*Universidad de Oviedo.* [3]*Centro de Investigación Biomédica en Red de Salud Mental* (CIBERSAM)

Spanish Psychology is in good health, both from an academic and professional point of view. This assertion is backed up by numerous indicators, such as the scientific production of our researchers, the fifty-five centres currently teaching psychology studies, or the more than fifty thousand members who make up the National Association of Psychologists. Not forgetting the recent milestone that has been the creation of the Psychology Academy of Spain. Of course, it has not been completely problem-free, but the path taken by Spanish psychology in recent decades has been astonishing. This special issue is our small contribution, which aims to show the progress of psychological assessment in different areas, from the point of view of the young, but highly qualified, researchers, upon whom the near future of Spanish psychology will depend. The contributors are notably scholarship holders of the *Formación del Personal Investigador (FPI)* [Research Staff Training] Program, assistant lecturers and contract doctors, postdoctoral students, psychologist intern residents (PIR), and university professors of the latest generation. They all belong to Spanish and foreign research groups, and in most cases have managed to establish a close and fruitful collaboration between the applied field and academic research. It will fall to them to build the Spanish psychology of the future. Of course, they are not all here, this monographic issue has no more space, but the sample is reasonably significant and clearly testifies that the future of Spanish psychology is in excellent health.

The studies included in this special issue were selected by Eduardo Fonseca Pedrero, Associate Lecturer at the University of La Rioja. The main idea was to choose rigorous and current studies, in different substantive areas, which were led by young researchers.

The first study, conducted by Davinia Fernández Espejo, analyses and discusses the main progress in assessing state of consciousness, from standardised clinical assessment scales to advanced neuroimaging techniques. Next, Javier Ortuño Sierra, Eduardo Fonseca Pedrero, Félix Inchausti and Silvia Sastre i Riba present a review of the Strengths and Difficulties Questionnaire (SDQ) in the child-youth population. This is followed by Beatriz Lucas Molina, Alicia Pérez Albéniz and Marta Giménez Dasí, who analyse the current situation and discuss future challenges in the assessment of cyberbullying. Meanwhile, in fourth position, the team of Nuria Ordóñez Camblor, Eduardo Fonseca Pedrero, Mercedes Paino, Leticia García Álvarez, Juan Pablo Pizarro Ruiz and Serafín Lemos Giráldez contribute a review of the evaluation of early traumatic experiences. In fifth place, José Luis Carballo, Ainhoa Coloma Carmona, Dana Mrozowicz Gaudyn, Verónica Vidal Arenas, Carlos van-der Hofstadt and Jesús Rodríguez-Marín highlight the increase in the prescription of opioid analgesic drugs and the negative consequences associated with their inappropriate use, presenting a proposal for the psychological assessment of the abuse of such substances. In the sixth study, Sergio Fernández Artamendi and Sara Weidberg López discuss some of the main advances in the assessment of addictions. Seventh, Ignacio Pedrosa and Javier Suárez Álvarez address the difficult problem of assessing the entrepreneurial personality, analysing the current state of the question and posing some future directions for the research in this field. Eighth, Isabel Benítez takes us through the present and future methodological challenges in the evaluation of quality of life. Finally, Ana María Ruiz Ruano and Jorge López Puga present an excellent introduction to the statistical program R, a free environment for statistical analysis which may be useful for psychological assessment in clinical and/or research settings.

We hope that the studies included in this special issue are of interest to both psychology practitioners and the readers that are more focused on research, since these two aspects, research and practice, are two sides of the same coin. Psychological research is carried out with the idea that one day the results will be applied to help people, and responsible professional practice must be based on empirical evidence from rigorous and replicable research. In summary, research and practice must walk hand in hand, enriching each other mutually; this is the only guaranteed formula for building rigorous and respectable psychology as a science and as a profession.

Correspondence: Eduardo Fonseca-Pedrero. *C/ Luis de Ulloa, s/n, Edificio VIVES. 26002 Logroño. La Rioja. España. E-mail: eduardo.fonseca@unioja.es*

Papeles del Psicólogo, 2016. Vol. 37(1), pp. 3-13
http://www.papelesdelpsicologo.es

ADVANCES IN THE ASSESSMENT OF DISORDERS OF CONSCIOUSNESS: THE ROLE OF BEDSIDE ASSESSMENT AND NEUROIMAGING TECHNIQUES IN THE DIAGNOSTIC PROCESS

Davinia Fernández-Espejo
University of Birmingham

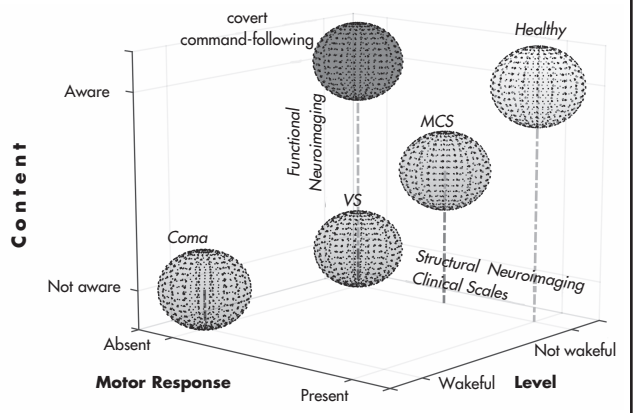
El estado vegetativo se define clínicamente por la falta de conciencia de uno mismo y el entorno, junto con el mantenimiento de funciones básicas como la respiratoria, cardiaca, o los ciclos de sueño y vigilia. Se trata de pacientes incapaces de reaccionar de un modo intencional a la estimulación externa y que no manifiestan ninguna capacidad comunicativa. Estudios recientes han demostrado que en torno al 40% de estos pacientes han sido incorrectamente diagnosticados y se encuentran, en realidad, conscientes. Sin embargo, en los últimos años se ha producido una revolución en las herramientas disponibles para evaluar a estos pacientes. El presente artículo tiene como objetivo discutir el papel de las escalas de evaluación clínica estandarizadas, así como técnicas avanzadas de neuroimagen, en la reducción del alarmante error diagnóstico. Se revisarán el alcance y las limitaciones de cada aproximación para identificar signos de conciencia externos o encubiertos, y se presentará evidencia a favor de una evaluación multimodal, combinando la información clínica, estructural y funcional para garantizar el diagnóstico correcto en cada caso individual.
Palabras clave: *Alteraciones de conciencia, Evaluación clínica, Resonancia magnética.*

Patients in a vegetative state are considered to lack awareness of themselves or the environment, but preserve respiratory and cardiac functions, as well as sleep/wake cycles. These patients are incapable of producing intentional responses to external stimulation and do not demonstrate any communication skills. Recent studies have shown that around 40% of vegetative state patients have been misdiagnosed. However, in recent years there has been a revolution in the tools that are available for the assessment of these patients. The objective of this article is to discuss the role of behavioural scales, as well as advanced neuroimaging techniques, in reducing the misdiagnosis rate. We review the scope and limitations of these approaches for the identification of overt and covert signs of awareness, and we present evidence to support a multimodal assessment that combines information from behavioural, structural, and functional imaging tools to ensure an accurate diagnosis for each individual patient.
Key words*: Disorders of consciousness, Clinical assessment, Magnetic resonance imaging.*

Defining consciousness and understanding its nature is one of the great challenges of contemporary neuroscience. Although the scientific community does not have a universally accepted definition of consciousness, the field of clinical neuroscience uses an operational definition that can be utilised in the bedside examination of the patient (Laureys, Perrin, & Bredart, 2007). In this context, consciousness is considered to be a complex system with two key dimensions: the level of consciousness or alertness (known as 'wakefulness' in the literature), and the content of consciousness or consciousness *per se* (known as 'awareness' in the literature) (Plum & Posner, 1982). The former refers to a state in which the eyes are open and there is a motor response. The latter encompasses both self-awareness and awareness of the environment, and refers to the ability to have subjective experiences of any kind. These two dimensions are related to very different brain mechanisms and networks. The level of consciousness depends on the ascending reticular activating system and its cortical connections (Parvizi & Damasio, 2001), while the content of consciousness depends on complex cortical-cortical and subcortical-cortical networks, which are not yet fully understood (Schiff, 2008). Because of the hierarchical relationship between these two systems, it is considered that a patient must have some preservation of the structures of wakefulness in order to have subjective experiences (awareness). Conversely, a patient may be alert but it should not be assumed that they are therefore aware of

Correspondence: Davinia Fernández-Espejo. *Lecturer. School of Psychology. 3.43 Hills Building. University of Birmingham. B15 2TT. E-mail: D.Fernandez-Espejo@bham.ac.uk*

FIGURE 1
DIMENSIONS IN ASSESSING DISORDERS OF CONSCIOUSNESS, AND TECHNIQUES OF CHOICE FOR THE DIFFERENTIAL DIAGNOSIS ON EACH AXIS



* Figure based on Monti, Coleman, & Owen, 2009 (with important modifications and including additional information).

themselves and their environment. This dissociation is crucial in understanding altered states of consciousness in patients who have suffered a severe brain injury (see Figure 1).

The most common causes of such injuries are traumatic brain injury (TBI) and hypoxic-ischemic encephalopathy (HIE) (The Multi-Society Task Force on PVS, 1994). The TBI that lead to severe consciousness disorders are often related to traffic accidents, although also, to a lesser extent, to falls or assaults. HIE, however, occurs after a prolonged lack of oxygen, in most cases after cardiorespiratory arrest, but also after drowning, carbon monoxide poisoning, etc. (The Multi-Society Task Force on PVS, 1994). Thanks to advances in emergency medical care, the widespread use of assisted ventilation and the presence of defibrillators in public places, a large number of patients survive both types of accidents today (Fernandez-Espejo & Owen, 2013). However, because of the extreme gravity of the accidents, many patients sustain severe brain injuries and enter a phase of coma.

A coma is an acute state (usually lasting a few days or weeks), in which the patient shows no signs of wakefulness or awareness: there is no spontaneous eye opening and the patient cannot be awakened with the application of vigorous sensory stimulation (Plum & Posner, 1982). Once the phase of the coma has passed, some of the patients regain consciousness and evolve favourably (albeit with cognitive sequelae of varying severity). A significant percentage of them, however, come out of the coma (open their eyes) but do not regain consciousness and fall into what is known as a **vegetative state** (VS).

Unlike the coma, the VS is defined by the preservation of wakefulness, manifested by the presence of sleep-wake cycles in the absence of consciousness (Jennett & Plum, 1972). These patients regain their autonomic function, the ability to regulate their breathing and heart rate without the aid of mechanical ventilation, but they do not react intentionally to stimulation, they do not respond to simple commands, and they do not have any communication skills (Royal College of Physicians, 2003). It is therefore considered that they are not aware of themselves or their surroundings. The VS is considered persistent when the patient remains unchanged one month after the accident, and permanent when no improvement has been recorded after 12 months after the TBI, or 3/6 months (according to American and British standards, respectively) in cases of HIE (Royal College of Physicians, 2003; The Multi-Society Task Force on PVS, 1994). Once the diagnosis of a permanent VS has been reached, it is considered that this state is irreversible and there is no possibility of recovery.

Before reaching the criterion of permanent, some patients begin to show fluctuating but clear signs of awareness and progress to what is known as a **minimally conscious state** (MCS) (Giacino et al., 2002). This category includes a heterogeneous group of patients, who at the bottom of the spectrum are able to follow a moving object with their eyes, and at the top are able to follow simple commands. In cases where there are no concomitant pathologies that reduce life expectancy, the patient may survive in a VS or MCS for decades. Some patients begin to be able to use everyday objects such as a cup, or a comb, or they manifest functional communication skills (i.e., they are able to respond to basic situational questions correctly). It is considered in this case that the patient has emerged from the MCS (Giacino et al., 2002), and would go on to receive a complete neuropsychological examination to determine the profile of the cognitive sequelae and, in cases where it is deemed appropriate, to design the rehabilitation program (Rosenbaum & Giacino, 2015; Royal College of Physicians, 2013).

At present there are no official statistics on the incidence or prevalence of disorders of consciousness partly because, with the exception of the coma, these clinical conditions are not listed in the International Classification of Diseases (ICD-10-ES, 2016 version[1]). In 2005, it was estimated that there were 46 new cases of VS patients per million inhabitants in the United States; and 14 per million inhabitants in the UK (Jennett, 2005). However, these data are based solely on trauma cases, since non-trauma cases have a more varied etiology and are referred to different specialists, making them difficult to identify. Although in Spain we do not have official figures for VS or MCS, in 2003 the *Institut de Neurorehabilitació Guttmann* estimated an annual incidence of TBI resulting in severe disability of 20 cases per 100,000 inhabitants/year (Alberdi Odriozola, Iriarte Ibarrán, Mendía Gorostidi, Murgialdai, & Marco Garde, 2009). With regards to the prevalence, several studies in Austria, the Netherlands and France have recently described the existence of between 0.2 and 6.1 VS patients, and 1.5 MCS patients per 100,000 inhabitants (Lavrijsen, van den Bosch, Koopmans, & van Weel, 2005; Pisa, Biasutti, Drigo, & Barbone, 2014; Saout et al., 2010). These figures are useful as a reference, but they cannot be easily extrapolated to other countries, among other reasons because there are important differences in the decisions concerning the termination of life in the acute phase, in cases such as these where the patient has catastrophic injuries. Although the incidence and prevalence are relatively low, the social, family and economic impact associated with the care of these patients is extremely high (Moretta et al., 2014), so it is necessary to carry out a proper assessment to identify the cognitive functions that the patient has preserved or lost, in order to ensure a good allocation of the resources. In the sections below, a critical review will be presented of the options available for diagnosing patients with disorders of consciousness in the areas of clinical assessment, as well as structural and functional neuroimaging. The most important contributions of each area will be discussed as well as the scope and limitations for identifying both overt and covert signs of awareness. Finally, this paper will argue the need for a multimodal assessment of patients with disorders of consciousness in order to ensure a correct diagnosis in each individual case.

CLINICAL ASSESSMENT OF DISORDERS OF CONSCIOUSNESS

Currently there are no objective biomarkers or laboratory analysis to determine whether a patient is or is not aware of themselves or their environment. The differential diagnosis of VS and MCS is based solely on the clinical examination of the patient, and observing the behavioural repertoire that they are capable of displaying; both spontaneously

[1] http://eciemaps.mspsi.es/ecieMaps/browser/index_10_mc.html

and in response to external stimulation provided by the examiner (Royal College of Physicians, 2003). The main function of the clinician is to determine whether the behaviours the patient displays are reflexes or whether they in fact indicate that the patient is able to interact intentionally with the environment. It is extremely complex to determine whether a behaviour is a reflex or voluntary, and in many cases the diagnostic process is further complicated by the presence of concomitant motor or language deficits that hinder the assessment (Majerus, Bruno, Schnakers, Giacino, & Laureys, 2009; Majerus, Gill-Thwaites, Andrews, & Laureys, 2005; Schnakers et al., 2015). In two studies in specialised neurorehabilitation centres in the US and the UK in the 90s, it was found that, due to these difficulties, 37% and 43% (respectively) of patients admitted with a diagnosis of VS had been incorrectly diagnosed. When these patients were re-evaluated by qualified personnel with experience in the diagnosis of disorders of consciousness, signs of partial consciousness were identified, or in some cases even complete consciousness (Andrews, Murphy, Munday, & Littlewood, 1996; Childs, Mercer, & Childs, 1993). The authors pointed to a lack of familiarity with the diagnostic criteria, and a lack of standardised assessments as being primarily responsible for these errors.

After these studies, the scientific and clinical communities agreed in recommending that the classic bedside examination is not sufficient to diagnose these patients, and it is necessary to use standardised batteries and to incorporate family and caregivers into the process (Bernat, 2006; Gill-Thwaites, 2006). In 2010, The Brain-Injury Interdisciplinary Special Interest Group, Disorders of Consciousness Task Force was established during the American Congress of Rehabilitation Medicine, with the aim of reviewing the scientific literature and formalising a recommendation of diagnostic scales, based on their content validity, diagnostic validity, reliability and prognostic value (American Congress of Rehabilitation Medicine, Brain Injury-Interdisciplinary Special Interest Group, Disorders of Consciousness Task Force et al., 2010). The main conclusions were that the only scales that were adequate for diagnosing patients with disorders of consciousness are the following: the Coma Recovery Scale-Revised (CRS- R) (Giacino, Kalmar, & Whyte, 2004), the Sensory Stimulation Assessment Measure (SSAM) (Rader & Ellis, 1994), the Wessex Head Injury Matrix (WHIM) (Shiel et al., 2000), the Western Neuro Sensory Stimulation Profile (WNSSP) (Ansell & Keenan, 1989), the Sensory Modality Assessment Technique (SMART) (Gill-Thwaites, 1997) and the Coma/Near-Coma Scale (CNC) (Rappaport, 2005) although the recommendation of this latter scale was with reservations. At the same time, they spoke out against the use of other scales that are widely used in neurological practice, such as the Full Outline of UnResponsiveness Score (FOUR) (Wijdicks, Bamlet, Maramattom, Manno, & McClelland, 2005), the Comprehensive Levels of Consciousness Scale (CLOCS) (Stanczak et al., 1984), the Innsbruck Coma Scale (INNS) (Benzer et al., 1991), the Glasgow-Liege Coma Scale (Born, 1988), the Swedish Reaction Level Scale-1985 (Johnstone et al., 1993), and the Loewenstein Communication Scale (Borer-Alafi, Gil, Sazbon, & Korn, 2002) due to their lack of content validity, standardisation or reliability (see Table 1).

Among the recommended scales, the two most complete ones, which contain specific elements for the differential diagnosis between VS and MCS, and which have received the most support in the scientific literature, are the SMART and CRS-R scales. The SMART scale was developed by occupational therapists at the Royal Hospital for

TABLE 1
RECOMMENDATIONS OF THE DISORDERS OF CONSCIOUSNESS TASK FORCE ON SCALES FOR THE DIAGNOSIS OF DISORDERS OF CONSCIOUSNESS. *(American Congress of Rehabilitation Medicine, Brain Injury-Interdisciplinary Special Interest Group, Disorders of Consciousness Task Force,* 2010)

Scale	Complete Name	Reference	Recommendation
CRS-R*	*Coma Recovery Scale-Revised*	Giacino, Kalmar & Whyte, 2004 (Spanish version: Noé et al., 2012)	Yes
CNC	*Coma/Near-Coma Scale*	Rappaport, 2005	With reservations
CLOCS	*Comprehensive Levels of Consciousness Scale*	Stanczak et al., 1984	No
INNS	*Glasgow-Liege Coma Scale* *Innsbruck Coma Scale (INNS)* *Loewenstein Communication Scale*	Born, 1988 Benzer et al., 1991 Borer-Alafi, Gil, Sazbon, & Korn, 2002	No No No
MATADOC	*Music Therapy Assessment Tool for Awareness in Disorders of Consciousness*	Magee, Siegert, Daveson, Lenton-Smith, & Taylor, 2013	Not studied
SMART*	*Sensory Modality Assessment Technique*	Gill-Thwaites, 1997	Yes
SSAM	*Sensory Stimulation Assessment Measure*	Rader & Ellis, 1994	Yes
FOUR	*Swedish Reaction Level Scale-1985* *The Full Outline of UnResponsiveness Score*	Johnstone et al., 1993 (Wijdicks, Bamlet, Maramattom, Manno, & McClelland, 2005	No No
WHIM	*Wessex Head Injury Matrix*	Shiel et al., 2000	Yes
WNSSP	*Western Neuro Sensory Stimulation Profile*	Ansell & Keenan, 1989	Yes

* Most widely accepted scales in the scientific community

Neuro-disability in London, as a tool for assessing and planning a rehabilitation program and monitoring its effects on the patient (Gill-Thwaites & Munday, 2004). It contains 29 sub-scales that allow a full exploration of the five sensory modalities, the motor function, functional communication and level of alertness. The patient's responses are classified hierarchically according to the functional level they represent (no response, reflex response, withdrawal response, localising response, or differentiating response). In order to access this scale, specific training by the team that developed it must be received, which, together with its high cost, in practice makes it inaccessible to most clinicians and researchers who are not resident in the UK (American Congress of Rehabilitation Medicine, Brain Injury-Interdisciplinary Special Interest Group, Disorders of Consciousness Task Force et al., 2010).

With a similar emphasis on rehabilitation, in this case based on music therapy, the Royal Hospital for Neuro-disability recently published the Music Therapy Assessment Tool for Awareness in Disorders of Consciousness scale (MATADOC) (Magee, Siegert, Daveson, Lenton-Smith, & Taylor, 2013). In the first standardisation study, the scale showed good internal validity and diagnosis consistent with that obtained using the SMART and CRS-R scales (Magee et al., 2013). As it does not rely on linguistic stimulation, this scale is especially useful in patients with aphasia (Schnakers et al., 2015), or paediatric patients (Magee, Ghetti, & Moyer, 2015). However, it suffers from access problems similar to those of the SMART scale, so its use is not yet widespread.

The CRS-R scale specifically evaluates all behaviours described by the Aspen Workgroup for the differential diagnosis of VS and MCS (Giacino et al., 2002), and has excellent content validity (American Congress of Rehabilitation Medicine, Brain Injury-Interdisciplinary Special Interest Group, Disorders of Consciousness Task Force et al., 2010). In addition, unlike the previous two scales, it is free to access and no formal training is required to administer it (Giacino et al., 2004). This has become the most widely used scale in the specialised scientific literature. It consists of 25 items hierarchically ordered and distributed in 6 subscales that assess the following different functions: auditory, visual, motor, oromotor/verbal, communication and arousal. The score for each subscale is based on the presence or absence of specific behaviours in response to sensory stimulation that the assessor presents in a standardised way. Low scores reflect reflex behaviours, while higher scores represent cognitively mediated behaviours (Giacino et al., 2004). This scale has recently been adapted to Spanish by the team of Enrique Noé at the Neuro-rehabilitation and Brain Damage Service of the NISA Hospital in Valencia (Noé et al., 2012).

Despite the high availability of these assessment scales (particularly the easy access to the CRS-R), the publication of differential diagnostic criteria (Giacino et al., 2002), and the recommendation to carry out standardised tests that appears in the clinical practice guidelines for dealing with patients with disorders of consciousness (*Royal College of Physicians*, 2003), a recent study, in which 103 patients were evaluated in Belgium, found a diagnostic error rate similar to the rate described in the 90s (Schnakers et al., 2009). The authors compared the diagnosis reached by clinical consensus in the medical team with that obtained after repeated assessments by qualified personnel using the CRS-R scale. The study found that 41% of patients with a clinical diagnosis of VS were actually in MCS, 10% of patients diagnosed with MCS had emerged from this state, and 89% of patients about whom the medical staff had not reached a consensus diagnosis were in MCS.

THE ROLE OF NEUROIMAGING

Diagnostic errors like those described in the previous section can have serious consequences. Firstly, the MCS has a more favourable prognosis than the VS (Giacino & Kalmar, 1997; Luauté et al., 2010), so misdiagnosis could influence the resources made available to the patient to facilitate recovery. Likewise, patients in MCS retain a higher cognitive processing capacity, which reaches more complex brain areas, than patients in VS (Boly et al., 2004; Laureys et al., 2000; 2002; Silva et al., 2010). For example, several studies have suggested that patients in MCS are capable of experiencing pain (Boly, Faymonville, et al., 2008a; Laureys et al., 2002), which must be taken into account when administering invasive clinical procedures. Finally, at present and in most jurisdictions in Western countries, legal proceedings relating to the withdrawal of life support (in this case artificial nutrition and hydration) are only initiated in cases where the patient has a diagnosis of VS (Andrews, 2004; Fernandez-Espejo & Owen, 2013).

In order to reduce this alarming misdiagnosis rate, several research groups have begun to use advanced structural neuroimaging techniques to identify objective biomarkers that provide complementary information to the clinical assessment. The foundations for this line of work were established in neuropathological studies carried out in the 90s, before the explosion of modern neuroimaging. After analysing 178 cases published in the scientific literature to date, Kinney and Samuel (1994) identified three general patterns of brain damage: diffuse axonal injury in trauma cases, destruction of the cortical rim in cases with hypoxic-ischemic etiology and thalamic lesions in both etiologies. These findings were confirmed in a series of successive studies (Adams, Graham, & Jennett, 2000; Adams, Jennett, McLellan, Murray, & Graham, 1999; Jennett, Adams, Murray, & Graham, 2001), which also reported a greater severity of diffuse axonal injury and traumatic lesions in VS patients than in those in MCS (Jennett et al., 2001).

The first morphometric studies based on magnetic resonance imaging (MRI) confirmed the previous findings (Ammermann et al., 2007; Juengling, Kassubek, Huppertz, Krause, & Els, 2005; Kampfl, Franz, et al., 1998a; Kampfl, Schmutzhard, et al., 1998b). However, it was not until 2011 that the first study was published in which it was possible to identify diagnostic biomarkers *in vivo*, by diffusion tensor imaging analysis (Fernandez-Espejo et al., 2011). This type of imaging allows us to characterise the microstructure of the brain tissue by observing the movement of the water molecules, and is particularly sensitive in detecting subtle changes that are not observable with other conventional forms of MRI (Bruno et al., 2011; Le Bihan et al., 2001). Firstly, this study confirmed differences in the severity of damage to the white matter and the thalamus between patients in VS and MCS. However, the real importance of this study is that, using only objective indices of the damage in these areas, it was possible to correctly diagnose 95% of the patients analysed (Fernandez-Espejo et al., 2011).

The thalamus is a structure of tremendous structural and functional complexity, with connections distributed across multiple cortical areas (Morel, Magnin, & Jeanmonod, 1997). Several studies have attempted to determine whether there is regional specificity in the thalamic damage in patients with disorders of consciousness (Fernandez-Espejo, Junque, Bernabeu, et al., 2010a; Lutkenhoff et al., 2015; 2013; Maxwell, MacKinnon, Smith, McIntosh, & Graham, 2006; Maxwell et al., 2004; Schiff, 2008). It has been shown that this atrophy particularly affects the central body (the dorsomedial nucleus and the internal medullary lamina), and is more pronounced in VS patients than in those in MCS (Fernandez-Espejo, Junque, Bernabeu, et al., 2010a; Maxwell et al., 2004; 2006), and in patients with TBI than those with HIE (Lutkenhoff et al., 2015). In trauma cases, the degree of acute atrophy present in the dorsomedial and anterior-medial nuclei has also been linked with prognosis at 6 months (Lutkenhoff et al., 2013).

Similar to the case of the thalamus, the regional distribution of white matter damage was profiled in a recent study of 52 patients of varying severity (Fernandez-Espejo et al., 2012). It was found to affect specifically the tracts connecting the cortical regions that make up part of the default mode network (medial prefrontal cortex, posterior cingulate/precuneus and inferior parietal lobes), as well as those connecting the posterior cingulate/precuneus with the thalamus. Numerous studies have found activation of this network in periods when we are resting, daydreaming or letting the mind wander (Buckner, Andrews-Hanna, & Schacter, 2008; Mason et al., 2007), and its functional integrity has been suggested as a prerequisite for the existence of conscious experience (Boly, Phillips, et al., 2008b; Laureys et al., 2007; Vanhaudenhuyse et al., 2010). In the previous study (Fernandez-Espejo et al., 2012), the severity of damage in the connections between the posterior and lateral nodes of this network as well as with the thalamus, correlated with the severity of the disorder of consciousness and the patient diagnosis.

Together, these findings open the way for the potential identification of more specific diagnostic biomarkers within the white matter and the thalamus. Although to date there have been no formal attempts in this direction, it is to be expected that this increased specificity will improve the diagnostic accuracy obtained previously (95%) (Fernandez-Espejo et al., 2011). The potential for clinical application of these techniques is clear, given that, as highlighted by the Royal College of Physicians in its latest guide (Royal College of Physicians, 2013), they do not require the participation of the patient and they can easily be performed in centres that are not specialised and that do not have research experience. Thus, if adopted as part of routine clinical assessment, they can help facilitate the identification of patients in MCS in cases where the diagnosis is not clear, or when the patient cannot be evaluated by teams of specialists (Schnakers et al., 2009).

In parallel, thanks to advances in *functional* neuroimaging techniques, a new group of conscious patients has been discovered whose detection is not possible even with assessments by teams of experts (Owen, 2013), or structural techniques. These patients retain complex cognitive skills but are unable to show them with external behaviour and, therefore, are incorrectly diagnosed as VS (see Figure 1). It is only possible to identify these cases through the use of techniques such as functional magnetic resonance imaging (fMRI) and electroencephalography (EEG), which enable us to relate changes in brain activation after presenting sensory stimuli with specific cognitive processes, without needing the patient to produce external verbal or motor responses (Owen, Epstein, & Johnsrude, n.d.). The early studies of brain activation in patients in VS or MCS were based on the presentation of passive stimulation, and showed that some of these patients retain emotional processing capabilities and are able to react, for example, to their own name (Di et al., 2007; Fischer, Luaute, & Morlet, 2010; Qin et al., 2010; 2008; Staffen, Kronbichler, Aichhorn, Mair, & Ladurner, 2006), familiar voices (Bekinschtein et al., 2004; de Jong, Willemsen, & Paans, 1997; Machado et al., 2007), familiar faces (Menon et al., 1998), or music with personal emotional content (O'Kelly et al., 2013; Okumura et al., 2014; Varotto et al., 2012). Successive studies have also found evidence of sensorimotor (Moritz et al., 2001; Schiff et al., 2005), visual (Monti, Pickard, & Owen, 2013; Moritz et al., 2001; Zhu et al., 2009), and linguistic processing (Bekinschtein et al., 2005; Fernandez-Espejo, Junque, Cruse, et al., 2010b; Fernandez-Espejo et al., 2008; Moritz et al., 2001; Owen et al., 2005; Schiff et al., 2005).

Several authors have stressed the need to carry out passive stimulation tasks hierarchically, starting with studying the simplest cognitive processes and progressively increasing their complexity (Laureys, Owen, & Schiff, 2004; Owen & Coleman, 2008a). Following this reasoning, Rodd and collaborators developed an auditory paradigm which proceeds from the basic acoustic processing of non-linguistic stimuli to semantic processing and linguistic comprehension (Rodd, Davis, & Johnsrude, 2005). In 2009, Coleman and colleagues used this paradigm in a group of 41 patients (22 VS, 19 MCS) (Coleman et al., 2009), revealing that 19 of them (7 VS 12 MCS) showed evidence of recognising linguistic stimuli (compared to other sounds) and 4 patients (2 VS, 2 MCS) showed evidence of linguistic comprehension, despite what might be inferred from their diagnosis. It is worth noting that the 7 patients in VS that showed linguistic responses in this paradigm progressed to MCS at 6 months, which suggests that the information obtained in fMRI tasks may have prognostic value. In fact, in a review of 15 studies published up to 2008 on fMRI and positron emission tomography in VS patients, Di and colleagues found that the presence of activation in association areas predicts a favourable outcome with 93% specificity and 69% sensitivity (Di, Boly, Weng, Ledoux, & Laureys, 2008). In a linguistic study similar to the previous one but carried out in Spain, it was also found that the only VS patient that showed linguistic responses in the fMRI (Fernandez-Espejo et al., 2008) had a favourable progression and regained consciousness one year after the initial injury (Fernandez-Espejo, Junque, Cruse, et al., 2010b).

The main advantage of passive fMRI paradigms is that they do not require the voluntary participation of the patient, so they can provide information about specific cognitive processes, regardless of the patient's ability or intention to collaborate. However, despite the fact that they can find cognitive functions contrary to diagnosis (e.g., linguistic comprehension), these types of paradigm do not allow us to make inferences about the state of consciousness of the patient. The only exception is the paradigm recently published by Naci and colleagues

(Naci, Cusack, Anello, & Owen, 2014), in which it was shown that when several people watch a movie (in this case a fragment of a short film by Hitchcock) their brain activity synchronises with that of the other spectators, and correlates with the film's executive demands. The same fragment was presented to a VS patient and it was shown that the patient's brain activity was highly correlated with those of the healthy volunteers, which was interpreted as evidence that the patient shared the conscious experience of the film with the healthy individuals.

Other attempts to overcome this limitation have been based on the use made in clinical practice of following simple commands (e.g., 'open your mouth', 'look at the ceiling', etc.) as definitive proof of consciousness (Giacino et al., 2004). With this idea in mind, active fMRI paradigms, in which the patient is asked, rather than to respond to these orders externally, to do so by voluntarily modulating their neuronal activity (Fernandez-Espejo & Owen, 2013). This approach is based on the fact that certain mental imagery tasks are associated with specific patterns of brain activation. Thus, the presence of these patterns can be used to determine that the patient followed the instructions and did the visualisation when asked to do so (Owen & Coleman, 2008b). Specifically, the paradigm that has proven most successful in identifying the following of orders in VS patients is based on motor imagery and spatial navigation. The patient is instructed to imagine moving their hand to hit a tennis ball repeatedly every time they hear the word 'tennis', or to imagine that they are going around the different rooms of their house and to try to visualise the objects that they would find every time they hear the word 'house' (Boly et al., 2007). In healthy volunteers, the two tasks elicit a very similar brain activation to that which would be obtained if the participant were actually moving his hand (supplementary motor area), or performing a spatial navigation task (parahippocampal cortex, posterior parietal lobe, and lateral premotor cortex) (Boly et al., 2007).

In 2006, Owen and colleagues used this task with a VS patient and found that the patient's brain activity was indistinguishable from that obtained with healthy volunteers, which showed that the patient was able to understand and follow instructions and therefore was not actually in a VS (Owen et al., 2006). Several subsequent studies have successfully used this paradigm to identify the following of orders in unresponsive patients (Fernandez-Espejo & Owen, 2013; Gibson et al., 2014; Monti et al., 2010). For example, Monti and colleagues studied a group of 23 VS patients and found evidence of following orders in 17% of them (Monti et al., 2010). What is even more important, one of these patients successfully managed to use activation in these two tasks (motor and spatial imagery) to communicate with the researchers; i.e., the patient used one type of visualisation to answer 'yes' and the other to answer 'no', and answered 5 autobiographical questions correctly (e.g., "Is your father's name Alexander?") (Monti et al., 2010). Recently, this technique allowed another patient, who had been in a VS for 12 years, to answer questions with important implications for his quality of life (e.g., whether he was suffering any pain) (Fernandez-Espejo & Owen, 2013). This patient also showed he knew the name of the person who had been his primary caregiver since the accident, whom he did not know before, showing that he was able to create memories of events that had occurred while he was diagnosed as being in a VS. Although so far this paradigm is only available in specialised research centres (Royal College of Physicians, 2013), it has been proven that it can be successfully performed on a clinical MRI scanner (Fernandez-Espejo, Norton, & Owen, 2014).

Other active tasks that have been applied to evaluate the following of orders in VS and MCS patients using fMRI include visualising motor activities such as swimming (Bardin et al., 2011; Forgacs et al., 2014), motor preparation (Bekinschtein, Manes, Villarreal, Owen, & Della-Maggiore, 2011), or attention directed to specific stimuli presented aurally (Monti et al., 2015; Naci & Owen, 2013; Naci, Cusack, Jia, & Owen, 2013), or visually (Hampshire et al., 2013; Monti et al., 2013). To date, the only one of these tasks that has been successfully used to communicate with patients in VS or MCS, is based on selective attention to the words 'yes' or 'no' (depending on the response) presented aurally (Naci & Owen, 2013).

Despite the great success of fMRI in this field, it is an expensive technique, it is not available in many hospitals, and it cannot be performed on patients with, for example, excessive agitation, certain metallic implants, or those who are unable to lie supine on a flat surface. Therefore, several research groups have developed active paradigms similar to the above, but based on the EEG (Coyle, Stow, McCreadie, McElligott, & Carroll, 2015; Cruse et al., 2011; Cruse, Chennu, Chatelle, et al., 2012a; Cruse, Chennu, Fernandez-Espejo, et al., 2012b; Gibson et al., 2014; Horki et al., 2014; Lulé et al., 2013; Pan et al., 2014; Schnakers et al., 2008). For example, Cruse and colleagues succeeded in identifying responses in two tasks of motor imagery (imagining that you close your hand and open it again, and imagining that you are moving your toes) in 19% of 16 VS patients (Cruse et al., 2011). This technique is portable, so the patient does not need to be transferred. The technique can be performed with the patient lying down or sitting up, and it has a much lower cost. However, to date, no patients in VS or MCS have managed to use EEG to communicate.

CONCLUSIONS

Three main conclusions can be drawn from the studies discussed here:

Firstly, the clinical assessment for the diagnosis of patients with disorders of consciousness should include the repeated administration of standardised scales by qualified personnel, to ensure the identification of subtle signs of consciousness that the patient is capable of displaying.

Secondly, structural MRI techniques have shown great potential for assisting in the diagnostic process through the objective identification of markers that enable us to differentiate between patients in VS and MCS. Their contribution is vital in cases where the clinical assessment does not provide a clear diagnosis, or when there is no team of experts available to evaluate the patient.

Finally, functional neuroimaging techniques (fMRI and EEG) are necessary in identifying covert cognitive functions, which some patients are not able to show externally. These functions can range from the basic processing of sensory stimuli to language comprehension, executive functions, or even the ability to follow simple orders in some cases. In fact, it is estimated that at least 17-19% of patients in VS are able to

follow orders in fMRI or EEG tests, and therefore have been diagnosed incorrectly.

Achieving a correct diagnosis and appropriately identifying the cognitive abilities of the patient has profound clinical implications, but also ethical and moral ones (Weijer et al., 2014). Because of this, and given that to date these tests are only available as part of research studies, there is a need to share the findings from these studies with the medical staff responsible for the patient and the family. To this end, and in collaboration with professionals of bioethics, we researchers in this area have recently developed an ethical framework for the disclosure of information obtained in our studies (Graham et al., 2014).

On the other hand, the evidence gathered to date, and discussed in this article, advocates the urgent need to re-evaluate the existing diagnostic categories to include this new group of patients, who are still conscious, but completely non-responsive externally (Fernandez-Espejo & Owen, 2013). It also points to the need to incorporate into the routine assessment of patients with disorders of consciousness functional and structural neuroimaging tests such as the ones reviewed here. To make this possible, first progress must be made in adapting the tasks and acquisition protocols so they are compatible with the equipment normally available in nonspecialised clinical centres (e.g., less powerful MRI machines, limited equipment for presenting stimulation to the patient, etc.). This will facilitate access to a greater number of patients and the realisation of future studies of validation and standardisation of neuroimaging tests in large samples, so that they may be included in the clinical practice guidelines for the management of patients with disorders of consciousness.

REFERENCES

Adams, J. H., Graham, D. I., & Jennett, B. (2000). The neuropathology of the vegetative state after an acute brain insult. *Brain*, *123 (Pt 7)*, 1327–1338.

Adams, J. H., Jennett, B., McLellan, D. R., Murray, L. S., & Graham, D. I. (1999). The neuropathology of the vegetative state after head injury. *Journal of Clinical Pathology*, *52*(December 2006), 804–806.

Alberdi Odriozola, F., Iriarte Ibarrán, M., Mendía Gorostidi, A., Murgialdai, A., & Marco Garde, P. (2009). Pronóstico de las secuelas tras la lesión cerebral [Prognosis of the sequelae after brain injury]. *Medicina Intensiva*, *4*.

American Congress of Rehabilitation Medicine, Brain Injury-Interdisciplinary Special Interest Group, Disorders of Consciousness Task Force, Seel, R. T., Sherer, M., Whyte, J., Katz, D. I., Giacino, J. T., et al. (2010). Assessment scales for disorders of consciousness: evidence-based recommendations for clinical practice and research. *Archives of Physical Medicine and Rehabilitation*, *91*(12), 1795–1813. doi:10.1016/j.apmr.2010.07.218

Ammermann, H., Kassubek, J., Lotze, M., Gut, E., Kaps, M., Schmidt, J., et al. (2007). MRI brain lesion patterns in patients in anoxia-induced vegetative state. *Journal of the Neurological Sciences*, *260*(1-2), 65–70.

Andrews, K. (2004). Medical decision making in the vegetative state: withdrawal of nutrition and hydration. *NeuroRehabilitation*, *19*(4), 299–304.

Andrews, K., Murphy, L., Munday, R., & Littlewood, C. (1996). Misdiagnosis of the vegetative state: retrospective study in a rehabilitation unit. *BMJ (Clinical Research Ed.)*, *313*(7048), 13–16.

Ansell, B. J., & Keenan, J. E. (1989). The Western Neuro Sensory Stimulation Profile: a tool for assessing slow-to-recover head-injured patients. *Archives of Physical Medicine and Rehabilitation*, *70*(2), 104–108.

Bardin, J. C., Fins, J. J., Katz, D. I., Hersh, J., Heier, L. A., Tabelow, K., et al. (2011). Dissociations between behavioural and functional magnetic resonance imaging-based evaluations of cognitive function after brain injury. *Brain*, *134*(Pt 3), 769–782. doi:10.1093/brain/awr005

Bekinschtein, T. A., Manes, F. F., Villarreal, M., Owen, A. M., & Della-Maggiore, V. (2011). Functional imaging reveals movement preparatory activity in the vegetative state. *Frontiers in Human Neuroscience*, *5*, 5. doi:10.3389/fnhum.2011.00005

Bekinschtein, T., Leiguarda, R., Armony, J., Owen, A., Carpintiero, S., Niklison, J., et al. (2004). Emotion processing in the minimally conscious state. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*(5), 788.

Bekinschtein, T., Tiberti, C., Niklison, J., Tamashiro, M., Ron, M., Carpintiero, S., et al. (2005). Assessing level of consciousness and cognitive changes from vegetative state to full recovery. *Neuropsychological Rehabilitation*, *15*(3-4), 307–322.

Benzer, A., Mitterschiffthaler, G., Marosi, M., Luef, G., Pühringer, F., La Renotiere, De, K., et al. (1991). Prediction of non-survival after trauma: Innsbruck Coma Scale. *The Lancet*, *338*(8773), 977–978.

Bernat, J. L. (2006). Chronic disorders of consciousness. *Lancet*, *367*(9517), 1181–1192.

Boly, M., Coleman, M. R., Davis, M. H., Hampshire, A., Bor, D., Moonen, G., et al. (2007). When thoughts become action: an fMRI paradigm to study volitional brain activity in non-communicative brain injured patients. *NeuroImage*, *36*(3), 979–992. doi:10.1016/j.neuroimage.2007.02.047

Boly, M., Faymonville, M.-E., Peigneux, P., Lambermont, B., Damas, P., Del Fiore, G., et al. (2004). Auditory processing in severely brain injured patients: differences between the minimally conscious state and the persistent vegetative state. *Archives of Neurology*, *61*(2), 233–238.

Boly, M., Faymonville, M.-E., Schnakers, C., Peigneux, P., Lambermont, B., Phillips, C., et al. (2008a). Perception of pain in the minimally conscious state with PET activation: an observational study. *The Lancet Neurology*, *7*(11), 1013–1020.

Boly, M., Phillips, C., Tshibanda, L., Vanhaudenhuyse, A., Schabus, M., Dang-Vu, T. T., et al. (2008b). Intrinsic brain activity in altered states of consciousness: how conscious is the default mode of brain function? *Annals of the New York Academy of Sciences*, *1129*, 119–129. doi:10.1196/annals.1417.015

Borer-Alafi, N., Gil, M., Sazbon, L., & Korn, C. (2002). Loewenstein communication scale for the minimally responsive patient. *Brain Injury : [BI]*, *16*(7), 593–609. doi:10.1080/02699050110119484

Born, J. D. (1988). The Glasgow-Liège Scale. Prognostic value and evolution of motor response and brain stem reflexes after severe head injury. *Acta Neurochirurgica*, *91*(1-2), 1–11.

Bruno, M. A., Fernández-Espejo, D., Lehembre, R., Tshibanda, L., Vanhaudenhuyse, A., Gosseries, O., et al. (2011). Multimodal neuroimaging in patients with disorders of consciousness showing "functional hemispherectomy". *Progress in Brain Research*, *193*, 323–333. doi:10.1016/B978-0-444-53839-0.00021-1

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. doi:10.1196/annals.1440.011

Childs, N. L., Mercer, W. N., & Childs, H. W. (1993). Accuracy of diagnosis of persistent vegetative state. *Neurology*, *43*(8), 1465–1467.

Coleman, M. R., Davis, M. H., Rodd, J. M., Robson, T., Ali, A., Owen, A. M., & Pickard, J. D. (2009). Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. *Brain*, *132*(Pt 9), 2541–2552. doi:10.1093/brain/awp183

Coyle, D., Stow, J., McCreadie, K., McElligott, J., & Carroll, Á. (2015). Sensorimotor modulation assessment and brain-computer interface training in disorders of consciousness. *Archives of Physical Medicine and Rehabilitation*, *96*(3 Suppl), S62–70. doi:10.1016/j.apmr.2014.08.024

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. A., Fernandez-Espejo, D., Pickard, J. D., et al. (2011). Bedside detection of awareness in the vegetative state: a cohort study. *Lancet*, *378*(9809), 2088–2094. doi:10.1016/S0140-6736(11)61224-5

Cruse, D., Chennu, S., Chatelle, C., Fernández-Espejo, D., Bekinschtein, T. A., Pickard, J. D., et al. (2012a). Relationship between etiology and covert cognition in the minimally conscious state. *Neurology*, *78*(11), 816–822. doi:10.1212/WNL.0b013e318249f764

Cruse, D., Chennu, S., Fernandez-Espejo, D., Payne, W. L., Young, G. B., & Owen, A. M. (2012b). Detecting awareness in the vegetative state: electroencephalographic evidence for attempted movements to command. *PloS One*, *7*(11), e49933. doi:10.1371/journal.pone.0049933

de Jong, B. M., Willemsen, A. T., & Paans, A. M. (1997). Regional cerebral blood flow changes related to affective speech presentation in persistent vegetative state. *Clinical Neurology and Neurosurgery*, *99*(3), 213–216.

Di, H. B., Yu, S. M., Weng, X. C., Laureys, S., Yu, D., Li, J. Q., et al. (2007). Cerebral response to patient's own name in the vegetative and minimally conscious states. *Neurology*, *68*(12), 895–899.

Di, H., Boly, M., Weng, X., Ledoux, D., & Laureys, S. (2008). Neuroimaging activation studies in the vegetative state: predictors of recovery? *Clinical Medicine (London, England*, *8*(5), 502–507.

Fernandez-Espejo, D., & Owen, A. M. (2013). Detecting awareness after severe brain injury. *Nature Reviews Neuroscience*. doi:10.1038/nrn3608

Fernandez-Espejo, D., Bekinschtein, T., Monti, M. M., Pickard, J. D., Junque, C., Coleman, M. R., & Owen, A. M. (2011). Diffusion weighted imaging distinguishes the vegetative state from the minimally conscious state. *NeuroImage*, *54*(1), 103–112. doi:10.1016/j.neuroimage.2010.08.035

Fernandez-Espejo, D., Junque, C., Bernabeu, M., Roig-Rovira, T., Vendrell, P., & Mercader, J. M. (2010a). Reductions of thalamic volume and regional shape changes in the vegetative and the minimally conscious states. *Journal of Neurotrauma*, *27*(7), 1187–1193. doi:10.1089/neu.2010.1297

Fernandez-Espejo, D., Junque, C., Cruse, D., Bernabeu, M., Roig-Rovira, T., Fábregas, N., et al. (2010b). Combination of diffusion tensor and functional magnetic resonance imaging during recovery from the vegetative state. *BMC Neurology*, *10*, 77. doi:10.1186/1471-2377-10-77

Fernandez-Espejo, D., Junque, C., Vendrell, P., Bernabeu, M., Roig, T., Bargallo, N., & Mercader, J. M. (2008). Cerebral response to speech in vegetative and minimally conscious states after traumatic brain injury. *Brain Injury : [BI]*, *22*(11), 882–890. doi:10.1080/02699050802403573

Fernandez-Espejo, D., Norton, L., & Owen, A. M. (2014). The clinical utility of fMRI for identifying covert awareness in the vegetative state: a comparison of sensitivity between 3T and 1.5T. *PloS One*, *9*(4), e95082. doi:10.1371/journal.pone.0095082

Fernandez-Espejo, D., Soddu, A., Cruse, D., Palacios, E. M., Junque, C., Vanhaudenhuyse, A., et al. (2012). A role for the default mode network in the bases of disorders of consciousness. *Annals of Neurology*, *72*(3), 335–343. doi:10.1002/ana.23635

Fischer, C., Luaute, J., & Morlet, D. (2010). Event-related potentials (MMN and novelty P3) in permanent vegetative or minimally conscious states. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *121*(7), 1032–1042.

Forgacs, P. B., Conte, M. M., Fridman, E. A., Voss, H. U., Victor, J. D., & Schiff, N. D. (2014). Preservation of electroencephalographic organization in patients with impaired consciousness and imaging-based evidence of command-following. *Annals of Neurology*, *76*(6), 869–879. doi:10.1002/ana.24283

Giacino, J. T., & Kalmar, K. (1997). The vegetative and minimally conscious states: a comparison of clinical features and functional outcome. *Journal of Head Trauma Rehabilitation*, *12*(4), 36–51.

Giacino, J. T., Ashwal, S., Childs, N., Cranford, R., Jennett, B., Katz, D. I., et al. (2002). The minimally conscious state: definition and diagnostic criteria. *Neurology*, *58*(3), 349–353.

Giacino, J. T., Kalmar, K., & Whyte, J. (2004). The JFK Coma Recovery Scale-Revised: Measurement characteristics and diagnostic utility. *Archives of Physical Medicine and Rehabilitation*, *85*(12), 2020–2029.

Gibson, R. M., Fernandez-Espejo, D., Gonzalez-Lara, L. E., Kwan, B. Y., Lee, D. H., Owen, A. M., & Cruse, D. (2014). Multiple tasks and neuroimaging modalities increase the likelihood of detecting covert awareness in patients with disorders of consciousness. *Frontiers in Human Neuroscience*, *8*, 950. doi:10.3389/fnhum.2014.00950

Gill-Thwaites, H. (1997). The Sensory Modality Assessment Rehabilitation Technique - A tool for assessment and treatment of patients with

severe brain injury in a vegetative state. *Brain Injury*, *11*(10), 723–734. Retrieved from http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=9354248&retmode=ref&cmd=prlinks

Gill-Thwaites, H. (2006). Lotteries, loopholes and luck: Misdiagnosis in the vegetative state patient. *Brain Injury : [BI]*, *20*(13-14), 1321–1328. Retrieved from http://informahealthcare.com/doi/abs/10.1080/02699050601081802

Gill-Thwaites, H., & Munday, R. (2004). The Sensory Modality Assessment and Rehabilitation Technique (SMART): a valid and reliable assessment for vegetative state and minimally conscious state patients. *Brain Injury : [BI]*, *18*(12), 1255–1269.

Graham, M., Weijer, C., Peterson, A., Naci, L., Cruse, D., Fernandez-Espejo, D., et al. (2014). Acknowledging awareness: informing families of individual research results for patients in the vegetative state. *Journal of Medical Ethics*. doi:10.1136/medethics-2014-102078

Hampshire, A., Parkin, B., Cusack, R., Fernández-Espejo, D., Allanson, J., Kamauc, E., et al. (2013). Assessing residual reasoning ability in overtly non-communicative patients using fMRI. *NeuroImage: Clinical*, *2*, 174–183.

Horki, P., Bauernfeind, G., Klobassa, D. S., Pokorny, C., Pichler, G., Schippinger, W., & Müller-Putz, G. R. (2014). Detection of mental imagery and attempted movements in patients with disorders of consciousness using EEG. *Frontiers in Human Neuroscience*, *8*, 1009. doi:10.3389/fnhum.2014.01009

Jennett, B. (2005). Thirty years of the vegetative state: clinical, ethical and legal problems. *Progress in Brain Research*, *150*, 537–543.

Jennett, B., & Plum, F. (1972). Persistent vegetative state after brain damage. A syndrome in search of a name. *The Lancet*, *1*(7753), 734–737.

Jennett, B., Adams, J. H., Murray, L. S., & Graham, D. I. (2001). Neuropathology in vegetative and severely disabled patients after head injury. *Neurology*, *56*(4), 486–490.

Johnstone, A. J., Lohlun, J. C., Miller, J. D., McIntosh, C. A., Gregori, A., Brown, R., et al. (1993). A comparison of the Glasgow Coma Scale and the Swedish Reaction Level Scale. *Brain Injury : [BI]*, *7*(6), 501–506.

Juengling, F. D., Kassubek, J., Huppertz, H.-J., Krause, T., & Els, T. (2005). Separating functional and structural damage in persistent vegetative state using combined voxel-based analysis of 3-D MRI and FDG-PET. *Journal of the Neurological Sciences*, *228*(2), 179–184.

Kampfl, A., Franz, G., Aichner, F., Pfausler, B., Haring, H. P., Felber, S., et al. (1998a). The persistent vegetative state after closed head injury: clinical and magnetic resonance imaging findings in 42 patients. *Journal of Neurosurgery*, *88*(5), 809–816.

Kampfl, A., Schmutzhard, E., Franz, G., Pfausler, B., Haring, H. P., Ulmer, H., et al. (1998b). Prediction of recovery from post-traumatic vegetative state with cerebral magnetic-resonance imaging. *Lancet*, *351*(9118), 1763–1767.

Kinney, H. C., & Samuels, M. A. (1994). Neuropathology of the persistent vegetative state. A review. *Journal of Neuropathology and Experimental Neurology*, *53*(6), 548–558. Retrieved from http://journals.lww.com/jneuropath/Abstract/1994/11000/Neuropathology_of_the_Persistent_Vegetative_State_.2.aspx

Laureys, S., Faymonville, M. E., Degueldre, C., Fiore, G. D., Damas, P., Lambermont, B., et al. (2000). Auditory processing in the vegetative state. *Brain : a Journal of Neurology*, *123 (Pt 8*, 1589–1601.

Laureys, S., Faymonville, M. E., Peigneux, P., Damas, P., Lambermont, B., Del Fiore, G., et al. (2002). Cortical processing of noxious somatosensory stimuli in the persistent vegetative state. *NeuroImage*, *17*(2), 732–741.

Laureys, S., Owen, A. M., & Schiff, N. D. (2004). Brain function in coma, vegetative state, and related disorders. *The Lancet Neurology*, *3*(9), 537–546.

Laureys, S., Perrin, F., & Bredart, S. (2007). Self-consciousness in non-communicative patients. *Consciousness and Cognition*, *16*(3), 722–41– discussion 742–5. Retrieved from https://www2.bc.edu/~slotnics/articles/ps571_C_Laureys2007.pdf

Lavrijsen, J. C. M., van den Bosch, J. S. G., Koopmans, R. T. C. M., & van Weel, C. (2005). Prevalence and characteristics of patients in a vegetative state in Dutch nursing homes. *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(10), 1420–1424.

Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., & Chabriat, H. (2001). Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging : JMRI*, *13*(4), 534–546.

Luauté, J., Maucort-Boulch, D., Tell, L., Quelard, F., Sarraf, T., Iwaz, J., et al. (2010). Long-term outcomes of chronic minimally conscious and vegetative states. *Neurology*, *75*(3), 246–252.

Lulé, D., Noirhomme, Q., Kleih, S. C., Chatelle, C., Halder, S., Demertzi, A., et al. (2013). Probing command following in patients with disorders of consciousness using a brain-computer interface. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *124*(1), 101–106. doi:10.1016/j.clinph.2012.04.030

Lutkenhoff, E. S., Chiang, J., Tshibanda, L., Kamau, E., Kirsch, M., Pickard, J. D., et al. (2015). Thalamic and extrathalamic mechanisms of (un)consciousness after severe brain injury. *Annals of Neurology*. doi:10.1002/ana.24423

Lutkenhoff, E. S., McArthur, D. L., Hua, X., Thompson, P. M., Vespa, P. M., & Monti, M. M. (2013). Thalamic atrophy in antero-medial and dorsal nuclei correlates with six-month outcome after severe brain injury. *NeuroImage: Clinical*, *3*, 396–404. doi:10.1016/j.nicl.2013.09.010

Machado, C., Korein, J., Aubert, E., Bosch, J., Alvarez, M. A., Rodriguez, R., et al. (2007). Recognizing a mother's voice in the persistent vegetative state. *Clinical Neurophysiology*, *38*(3), 124–126.

Magee, W. L., Ghetti, C. M., & Moyer, A. (2015). Feasibility of the music therapy assessment tool for awareness in disorders of consciousness (MATADOC) for use with pediatric populations. *Frontiers in Psychology*, *6*, 698. doi:10.3389/fpsyg.2015.00698

Magee, W. L., Siegert, R. J., Daveson, B. A., Lenton-Smith, G., & Taylor, S. M. (2013). Music Therapy Assessment Tool for Awareness in

Disorders of Consciousness (MATADOC): Standardisation of the principal subscale to assess awareness in patients with disorders of consciousness. *Neuropsychological Rehabilitation*. doi:10.1080/09602011.2013.844174

Majerus, S., Bruno, M.-A. E. L., Schnakers, C., Giacino, J. T., & Laureys, S. (2009). The problem of aphasia in the assessment of consciousness in brain-damaged patients. *Progress in Brain Research*, *177*, 49–61.

Majerus, S., Gill-Thwaites, H., Andrews, K., & Laureys, S. (2005). Behavioral evaluation of consciousness in severe brain damage. *Progress in Brain Research*, *150*, 397–413.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science*, *315*(5810), 393–395. doi:10.1126/science.1131295

Maxwell, W. L., MacKinnon, M. A., Smith, D. H., McIntosh, T. K., & Graham, D. I. (2006). Thalamic nuclei after human blunt head injury. *Journal of Neuropathology and Experimental Neurology*, *65*(5), 478–488.

Maxwell, W. L., Pennington, K., MacKinnon, M. A., Smith, D. H., McIntosh, T. K., Wilson, J. T. L., & Graham, D. I. (2004). Differential responses in three thalamic nuclei in moderately disabled, severely disabled and vegetative patients after blunt head injury. *Brain*, *127*(Pt 11), 2470–2478. doi:10.1093/brain/awh294

Menon, D. K., Owen, A. M., Williams, E. J., Minhas, P. S., Allen, C. M., Boniface, S. J., & Pickard, J. D. (1998). Cortical processing in persistent vegetative state. Wolfson Brain Imaging Centre Team. *Lancet*, *352*(9123), 200.

Monti, M. M., Coleman, M. R., & Owen, A. M. (2009). Neuroimaging and the vegetative state: resolving the behavioral assessment dilemma? *Annals of the New York Academy of Sciences*, *1157*, 81–89.

Monti, M. M., Pickard, J. D., & Owen, A. M. (2013). Visual cognition in disorders of consciousness: from V1 to top-down attention. *Human Brain Mapping*, *34*(6), 1245–1253. doi:10.1002/hbm.21507

Monti, M. M., Rosenberg, M., Finoia, P., Kamau, E., Pickard, J. D., & Owen, A. M. (2015). Thalamo-frontal connectivity mediates top-down cognitive functions in disorders of consciousness. *Neurology*, *84*(2), 167–173. doi:10.1212/WNL.0000000000001123

Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., et al. (2010). Willful modulation of brain activity in disorders of consciousness. *New England Journal of Medicine*, *362*(7), 579–589.

Morel, A., Magnin, M., & Jeanmonod, D. (1997). Multiarchitectonic and stereotactic atlas of the human thalamus. *The Journal of Comparative Neurology*, *387*(4), 588–630.

Moretta, P., Estraneo, A., De Lucia, L., Cardinale, V., Loreto, V., & Trojano, L. (2014). A study of the psychological distress in family caregivers of patients with prolonged disorders of consciousness during in-hospital rehabilitation. *Clinical Rehabilitation*, *28*(7), 717–725. doi:10.1177/0269215514521826

Moritz, C. H., Rowley, H. A., Haughton, V. M., Swartz, K. R., Jones, J., & Badie, B. (2001). Functional MR imaging assessment of a non-responsive brain injured patient. *Magnetic Resonance Imaging*, *19*(8), 1129–1132.

Naci, L., & Owen, A. M. (2013). Making every word count for nonresponsive patients. *JAMA Neurology*, *70*(10), 1235–1241. doi:10.1001/jamaneurol.2013.3686

Naci, L., Cusack, R., Anello, M., & Owen, A. M. (2014). A common neural code for similar conscious experiences in different individuals. *Proceedings of the National Academy of Sciences*, *111*(39), 14277–14282. doi:10.1073/pnas.1407007111

Naci, L., Cusack, R., Jia, V. Z., & Owen, A. M. (2013). The brain's silent messenger: using selective attention to decode human thought for brain-based communication. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *33*(22), 9385–9393. doi:10.1523/JNEUROSCI.5577-12.2013

Noé, E., Olaya, J., Navarro, M. D., Noguera, P., Colomer, C., García-Panach, J., et al. (2012). Behavioral recovery in disorders of consciousness: a prospective study with the Spanish version of the Coma Recovery Scale-Revised. *Archives of Physical Medicine and Rehabilitation*, *93*(3), 428–33.e12. doi:10.1016/j.apmr.2011.08.048

O'Kelly, J., James, L., Palaniappan, R., Taborin, J., Fachner, J., & Magee, W. L. (2013). Neurophysiological and behavioral responses to music therapy in vegetative and minimally conscious States. *Frontiers in Human Neuroscience*, *7*, 884. doi:10.3389/fnhum.2013.00884

Okumura, Y., Asano, Y., Takenaka, S., Fukuyama, S., Yonezawa, S., Kasuya, Y., & Shinoda, J. (2014). Brain activation by music in patients in a vegetative or minimally conscious state following diffuse brain injury. *Brain Injury : [BI]*, *28*(7), 944–950. doi:10.3109/02699052.2014.888477

Owen, A. M. (2013). Detecting consciousness: a unique role for neuroimaging. *Annual Review of Psychology*, *64*, 109–133. doi:10.1146/annurev-psych-113011-143729

Owen, A. M., & Coleman, M. R. (2008a). Detecting awareness in the vegetative state. *Annals of the New York Academy of Sciences*, *1129*, 130–138. doi:10.1196/annals.1417.018

Owen, A. M., & Coleman, M. R. (2008b). Functional neuroimaging of the vegetative state. *Nature Reviews Neuroscience*, *9*(3), 235–243. doi:10.1038/nrn2330

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science*, *313*(5792), 1402–1402.

Owen, A. M., Epstein, R., & Johnsrude, I. S. (n.d.). *Functional Magnetic Resonance Imaging. An Introduction to Methods*. (P. Jezzard, P. M. Mathews, & S. M. Smith) (2001st ed., pp. 311–328). Oxford: Oxford University Press Inc.

Owen, A., Coleman, M., Menon, D., Johnsrude, I., Rodd, J., Davis, M., et al. (2005). Residual auditory function in persistent vegetative state: a combined pet and fmri study. *Neuropsychological Rehabilitation*, *15*(3-4), 290–306.

Pan, J., Xie, Q., He, Y., Wang, F., Di, H., Laureys, S., et al. (2014). Detecting awareness in patients with disorders of consciousness using a hybrid brain-computer interface. *Journal of Neural Engineering*, *11*(5), 056007. doi:10.1088/1741-2560/11/5/056007

Parvizi, J., & Damasio, A. R. (2001). Consciousness and the brainstem. *Cognition*, *79*(1-2), 135–160.

Pisa, F. E., Biasutti, E., Drigo, D., & Barbone, F. (2014). The prevalence of vegetative and minimally conscious states: a systematic review and methodological appraisal. *The Journal of Head Trauma Rehabilitation*, *29*(4), E23–30. doi:10.1097/HTR.0b013e3182a4469f

Plum, F., & Posner, J. (1982). *The diagnosis of stupor and coma.* (3rd ed.). Philadelphia: Oxford University Press.

Qin, P., Di, H., Liu, Y., Yu, S., Gong, Q., Duncan, N., et al. (2010). Anterior cingulate activity and the self in disorders of consciousness. *Human Brain Mapping*, *31*(12), 1993–2002. doi:10.1002/hbm.20989

Qin, P., Di, H., Yan, X., Yu, S., Yu, D., Laureys, S., & Weng, X. (2008). Mismatch negativity to the patient's own name in chronic disorders of consciousness. *Neuroscience Letters*, *448*(1), 24–28.

Rader, M. A., & Ellis, D. W. (1994). The Sensory Stimulation Assessment Measure (SSAM): a tool for early evaluation of severely brain-injured patients. *Brain Injury : [BI]*, *8*(4), 309–321.

Rappaport, M. (2005). The Disability Rating and Coma/Near-Coma scales in evaluating severe head injury. *Neuropsychological Rehabilitation*, *15*(3-4), 442–453. doi:10.1080/09602010443000335

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex (New York, N.Y.: 1991)*, *15*(8), 1261–1269. doi:10.1093/cercor/bhi009

Rosenbaum, A. M., & Giacino, J. T. (2015). Clinical management of the minimally conscious state. *Handbook of Clinical Neurology*, *127*, 395–410. doi:10.1016/B978-0-444-52892-6.00025-8

Royal College of Physicians. (2013). *Prolonged disorders of consciousness: National clinical guidelines*. (RCP). London. Retrieved from https://www.rcplondon.ac.uk/sites/default/files/pdoc_web_final_navigable_2014.pdf

Royal College of Physicians, W. P. (2003). The vegetative state: guidance on diagnosis and management. *Clinical Medicine (London, England)*, *3*(3), 249–254.

Sao u t, V., Ombredane, M. P., Mouillie, J. M., Marteau, C., Math e, J.-F., & Richard, I. (2010). Patients in a permanent vegetative state or minimally conscious state in the Maine-et-Loire county of France: A cross-sectional, descriptive study. *Annals of Physical and Rehabilitation Medicine*, *53*(2), 96–104.

Schiff, N. D. (2008). Central thalamic contributions to arousal regulation and neurological disorders of consciousness. *Annals of the New York Academy of Sciences*, *1129*(1), 105–118.

Schiff, N. D., Rodriguez-Moreno, D., Kamal, A., Kim, K., Giacino, J. T., Plum, F., & Hirsch, J. (2005). fMRI reveals large-scale network activation in minimally conscious patients. *Neurology*, *64*, 515–523.

Schnakers, C., Bessou, H., Rubi-Fessen, I., Hartmann, A., Fink, G. R., Meister, I., et al. (2015). Impact of aphasia on consciousness assessment: a cross-sectional study. *Neurorehabilitation and Neural Repair*, *29*(1), 41–47. doi:10.1177/1545968314528067

Schnakers, C., Perrin, F., Schabus, M., Majerus, S., Ledoux, D., Damas, P., et al. (2008). Voluntary brain processing in disorders of consciousness. *Neurology*, *71*(20), 1614–1620. doi:10.1212/01.wnl.0000334754.15330.69

Schnakers, C., Vanhaudenhuyse, A., Giacino, J., Ventura, M., Boly, M., Majerus, S., et al. (2009). Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. *BMC Neurology*, *9*(1), 35. doi:10.1186/1471-2377-9-35

Shiel, A., Horn, S. A., Wilson, B. A., Watson, M. J., Campbell, M. J., & McLellan, D. L. (2000). The Wessex Head Injury Matrix (WHIM) main scale: a preliminary report on a scale to assess and monitor patient recovery after severe head injury. *Clinical Rehabilitation*, *14*(4), 408–416.

Silva, S., Alacoque, X., Fourcade, O., Samii, K., Marque, P., Woods, R., et al. (2010). Wakefulness and loss of awareness. Brain and brainstem interaction in the vegetative state. *Neurology*, *74*, 313–320.

Staffen, W., Kronbichler, M., Aichhorn, M., Mair, A., & Ladurner, G. (2006). Selective brain activity in response to one's own name in the persistent vegetative state. *Journal of Neurology, Neurosurgery & Psychiatry*, *77*(12), 1383–1384.

Stanczak, D. E., White, J. G., Gouview, W. D., Moehle, K. A., Daniel, M., Novack, T., & Long, C. J. (1984). Assessment of level of consciousness following severe neurological insult. A comparison of the psychometric qualities of the Glasgow Coma Scale and the Comprehensive Level of Consciousness Scale. *Journal of Neurosurgery*, *60*(5), 955–960. doi:10.3171/jns.1984.60.5.0955

The Multi-Society Task Force on PVS. (1994). Medical aspects of the persistent vegetative state (I). *New England Journal of Medicine*, *330*, 1499–1508.

Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L. J.-F., Bruno, M.-A., Boveroux, P., Schnakers, C., et al. (2010). Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. *Brain*, *133*(Pt 1), 161–171. doi:10.1093/brain/awp313

Varotto, G., Fazio, P., Rossi Sebastiano, D., Avanzini, G., Franceschetti, S., Panzica, F., & CRC. (2012). Music and emotion: an EEG connectivity study in patients with disorders of consciousness. *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *2012*, 5206–5209. doi:10.1109/EMBC.2012.6347167

Weijer, C., Peterson, A., Webster, F., Graham, M., Cruse, D., Fernandez-Espejo, D., et al. (2014). Ethics of neuroimaging after serious brain injury. *BMC Medical Ethics*, *15*, 41. doi:10.1186/1472-6939-15-41

Wijdicks, E. F. M., Bamlet, W. R., Maramattom, B. V., Manno, E. M., & McClelland, R. L. (2005). Validation of a new coma scale: The FOUR score. *Annals of Neurology*, *58*(4), 585–593.

Zhu, J., Wu, X., Gao, L., Mao, Y., Zhong, P., Tang, W., & Zhou, L. (2009). Cortical activity after emotional visual stimulation in minimally conscious state patients. *Journal of Neurotrauma*, *26*(5), 677–688.

Papeles del Psicólogo, 2016. Vol. 37(1), pp. 14-26
http://www.papelesdelpsicologo.es

ASSESSING BEHAVIOURAL AND EMOTIONAL DIFFICULTIES IN THE CHILD-ADOLESCENT POPULATION: THE STRENGTHS AND DIFFICULTIES QUESTIONNAIRE (SDQ)

Javier Ortuño-Sierra[1], Eduardo Fonseca-Pedrero[2,3,4], Félix Inchausti[5] y Sylvia Sastre i Riba[2]
[1]Universidad Loyola. [2]Universidad de La Rioja. [3]Programa de Prevención en Psicosis (P3). [4]Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM). [5]Fundación Proyecto Hombre de Navarra

El Cuestionario de Capacidades y Dificultades (SDQ) es una herramienta de screening que permite la evaluación de dificultades emocionales y comportamentales así como del comportamiento prosocial en la infancia y adolescencia, desde una perspectiva multi-informante. El objetivo de este trabajo es llevar a cabo una revisión selectiva de las características epidemiológicas así como de las principales evidencias a nivel psicométrico del SDQ. Las propiedades psicométricas referidas a la fiabilidad de las puntuaciones son adecuadas y el modelo dimensional de cinco factores (Problemas Emocionales, Problemas Conductuales, Problemas con los Compañeros, Hiperactividad y Prosocial) es el más ampliamente replicado. Asimismo, se han obtenido evidencias de validez que apoyan la utilidad de este instrumento de medida para su uso en el contexto escolar y clínico. Los resultados también indican que el género y la edad influyen en la expresión fenotípica de las dificultades emocionales y comportamentales. En conclusión, el SDQ es un instrumento de medida breve, sencillo de administrar y útil para la valoración de este tipo de problemática en la infancia y adolescencia y puede ser de sumo interés para su uso en población infanto-juvenil española.
Palabras clave: *SDQ, Evaluación, Adolescencia, Problemas emocionales, Revisión, Propiedades psicométricas.*

The Strengths and Difficulties Questionnaire (SDQ) is a screening tool that enables the assessment of emotional and behavioural difficulties, and prosocial behaviour in children and adolescents from a multi-informant perspective. The main goal of this article is to carry out a selective review on the main evidence concerning the psychometric and epidemiologic characteristics of the SDQ. The psychometric properties are adequate with regard to the reliability of the scores and the five-factor structure is the most accepted (emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behaviour). In addition, different studies support the appropriateness of the SDQ for use as an evaluation tool in clinical and school contexts. The results show that gender and age have an influence on the phenotypic expression of emotional and behavioural difficulties. In conclusion, the SDQ is a short, easy to use, and useful measurement tool for evaluating problems, difficulties, and capacities related to childhood and adolescence and it may be used with Spanish children and adolescents.
Key words: *SDQ, Assessment, Adolescence, Emotional problems, Review, Psychometric properties.*

MENTAL HEALTH IN CHILDHOOD AND ADOLESCENCE

The *Encuesta Nacional de Salud España 2006* [Spain National Health Survey 2006], conducted with the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997), indicated that between 19.2 and 26.6% of Spanish children and adolescents between the ages of 4 and 15 years old were at risk of mental health problems (Fajardo, León, Felipe, & Ribeiro, 2012). Previous studies carried out in Spain show similar prevalence rates of behavioural and emotional symptoms and disorders, both in the general population (Blanco et al., 2015; Bones, Pérez, Rodríguez-Sanz, Borrell, & Obiols, 2010; Haro et al., 2006) and in the child population (Cuesta et al., 2015; Diaz de Neira et al., 2015; Fonseca-Pedrero, Paino, Lemos-Giraldez, & Muñiz, 2012; Ortuño, Fonseca-Pedrero, Paino, & Aritio-Solana, 2014). Considering the possible methodological differences, these rates are similar to those found in epidemiological studies worldwide (Olfson, Blanco, Wang, Laje, & Correll, 2014; Polanczyk, Salum, Sugaya, Caye, & Rohde, 2015; Wichstrøm et al., 2012). For example, in an excellent review conducted by Polanczyk et al. (2015), which included 41 studies conducted in 27 countries in all regions of the world, it was found that the prevalence of mental disorders in children and adolescents was 13.4% (95% confidential interval: 11.3 to 15.9).

The previous literature indicates that a significant percentage of children and adolescents present difficulties in psychological adjustment throughout their life, which has a clear impact not only on the personal, academic, family and social areas, but also at the health and economic levels (Blanchard, Gurka, & Blackman, 2006; Domino et al., 2009; Simpson, Bloom, Cohen, Blumberg, & Bourdon, 2005). Such symptoms tend to start in about 50% of cases before the age of 15 and they usually remain stable until adulthood (Copeland, Shanahan, y Costello, 2011; Costello, Copeland, & Angold, 2011; Davies et al., 2015; Widiger, De Clercq, & De Fruyt, 2009). Furthermore, the presence of emotional and behavioural subclinical symptoms at these ages increases the subsequent risk of developing a severe form of mental disorder (e.g., depression, psychosis) and general health problems of various kinds (Cullins & Mian, 2015; Klein, Shankman, Lewinsohn, & Seeley, 2009; Najman et al., 2008; Welham et al., 2009).

There is no doubt that the patterns of health and disease in childhood and adolescence have changed in recent decades, in what is known as the "new morbidity" (Cullins & Mian, 2015; Palfrey, Tonniges, Green, & Richmond, 2005). While epidemics have been declining gradually, mental health problems, such as emotional or behavioural disorders, have become more significant (Drabick & Kendall, 2010; Polanczyk et al., 2015), generating a progressive social process of becoming aware

Correspondence: Javier Ortuño Sierra. *Departamento de Psicología. Universidad Loyola. C/ Energía Solar, 1. 41014 Sevilla. España. E-mail: jortuno@uloyola.es*

of the needs in matters of child and adolescent mental health (Mulloy, Evangelista, Betkowski, & Weist, 2011). To this could be added the significant increase in the prevalence rates of certain mental disorders, which begin in childhood and adolescence and are clearly linked to our current lifestyle (Mulloy et al., 2015). In this sense, it is necessary to have rigorous tools for screening and assessment as well as effective psychological interventions for this sector of the population (Fonagy et al., 2015), to reduce or mitigate the global burden and the associated disability and morbidity, and ultimately, to help improve the quality of life of individuals and society.

Within this context, the main objective of this work is to carry out a selective review of the Strengths and Difficulties Questionnaire (Goodman, 1997), as a tool for screening and evaluating emotional and behavioural difficulties as well as prosocial type skills during childhood and adolescence. Specifically, once the importance of carrying out early identification and detection in this sector of the population has briefly been analysed, this study will include the following steps: 1) expose some of the measuring instruments for assessing psychopathology as well as emotional and behavioural problems, focusing on the multi-informant SDQ system; 2) analyse the psychometric properties of the SDQ regarding the reliability of scores and obtaining different evidence of validity in studies carried out both nationally and internationally; 3) analyse the influence of gender and age on the phenotypic expression of the difficulties and capabilities through the SDQ; and 4) finally, to recapitulate, the main conclusions are discussed as well as the possible directions of future research.

THE IMPORTANCE OF EARLY DETECTION IN MENTAL HEALTH

There is now greater awareness among mental health professionals regarding the consequences of a lack of early detection of such difficulties and the benefits associated with early prophylactic intervention in childhood and adolescence (Moscoso, Jovanovic, & Rojnic, 2015; Mulloy et al., 2011; Steinberg & Morris, 2001). The delay in identifying the clinical or subclinical conditions (e.g., affective symptoms) can be associated, among other things, with increased symptoms in adulthood, as well as a worse outcome or prognosis in the medium to long term (e.g., Drancourt et al., 2013). Helping mental health practitioners in the early detection of this set of experiences and symptoms, both clinical and subclinical, it is a goal of great interest with clear practical implications.

The scientific evidence suggests the need to increase and improve the early detection of the indicators of psychological maladjustment in child population (Moscoso et al, 2015). All of this has the aim of preventing the possible consequences and managing the existing resources (e.g., healthcare, school) more effectively. The detection, prevention and treatment of these types of emotional and behavioural problems is a cardinal issue, not only in order to solve specific problems, but also to improve adult functioning and prevent the consolidation of difficulties and problems in future generations (Brimblecombe et al., 2015; Ford, Goodman, & Meltzer, 2003; Stockings et al., 2015). Similarly, the early detection of existing psychological difficulties enables us to identify subclinical symptoms that may go unnoticed and become the potential cause of other, bigger, personal, social and economic problems, given the possibility of escalation and worse prognosis (Aebi, Giger, Plattner, Winkler Metzke, & Steinhausen, 2014; Levitt, Saka, Romanelli, & Hoagwood, 2007).

Despite the efforts dedicated to early identification and detection, various investigations suggest that only a minority of the child and youth population in need of intervention in the area of mental health comes to specialised services (Angold et al., 1998; Ford, Hamilton, Meltzer, & Goodman, 2008). In other words, strategies for primary prevention and secondary prevention are not yet well established in this sector of the population (Du, Kou, & Coghill, 2008). This causes the increase of tertiary prevention in the treatment of psychological problems, once the clinical condition has manifested, which results in the intervention being more difficult, with poorer results and additional costs (Ford et al., 2008).

At present, within the field of education, school psychologists are focusing their functions and tasks beyond mere intervention, paying greater attention to prevention, prioritising universal screening over selective and indicated screening (Cummings et al., 2004; Hoagwood & Johnson, 2003). This has generated the need for short, simple instruments with adequate psychometric characteristics that enable the rigorous evaluation and measurement of the emotional, behavioural and prosocial adjustment of children and adolescents (Hill & Hughes, 2007). The school context also is of great significance in the analysis of different types of mental health problems and difficulties, since it is in this context that many of these problems occur, so it is an ideal and crucial framework for the detection of different mental health problems (Mulloy et al., 2011).

PSYCHOPATHOLOGICAL ASSESSMENT IN CHILDREN AND ADOLESCENTS: AN INTRODUCTION TO THE SDQ

Over recent years there have been great advances in the measurement and evaluation of the psychological adjustment of children and adolescents. Two of the main measuring instruments, now classics, for the assessment of psychopathology and behavioural and emotional problems in childhood and adolescence are the Rutter questionnaires (Rutter & Graham, 1966) and those belonging to the ASEBA system (Achenbach System of Empirically Basic Assessment) (Achenbach, 1991a, 1991b, 1991c; Achenbach & Rescorla, 2001, 2007). Also, and more recently, the SDQ has also become particularly significant (Goodman, 1997), since, as will be discussed below, it has a number of features such as its brevity or the inclusion of a subscale of prosocial behaviour, which could make it more recommendable compared to the previous two, always bearing in mind the objective of the assessment and intervention, of course.

In the sixties, the Rutter questionnaires (Rutter & Graham, 1966) were developed for detecting emotional and behavioural problems, and the adequate reliability of their scores and their validity evidence were confirmed (Goodman, 1994; Rutter & Graham, 1966). However, these measuring instruments do not include a number of current areas of interest for psychology and child and adolescent psychiatry, such as prosocial type capabilities or hyperactivity (Koskelainen, Sourander, & Kaljonen, 2000) (see Table 1). The ASEBA system, originally built by Achenbach (Achenbach, 1991a, 1991b, 1991c), is one of the best-known multiaxial assessment systems that has been extensively validated and has proven useful in the detection of mental health problems in child-youth population. In its 2001 version, it was enriched by the inclusion of updated versions aimed at both young people (*Youth Self Report*, YSR/11-18 years), as well as teachers (*Teachers' Report Form*, TRF/6-18 years) and parents (*Child Behavior Checklist*, CBCL/6-18 years) (Achenbach & Rescorla, 2001). This version also offered the

possibility of generating scores that are equivalent to the criteria of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

However, the different versions of the ASEBA system have a number of disadvantages. Specifically, while it is true that the CBCL and YSR (Achenbach & Rescorla, 2001) have the advantage of being more current than the Rutter questionnaires and they cover a larger number of facets in the evaluation, they are also more time-consuming in their administration since they contain more than 100 items (Bourdon, Goodman, Rae, Simpson, & Koretz, 2005; Koskelainen et al., 2000). Thus, the measuring instruments belonging to the ASEBA system, in any of their forms, as well as others used for similar purposes, such as the *Behavior Assessment System for Children* (BASC) (Reynolds & Kamphaus, 1992), have the disadvantage of being slow to administer, resulting arduous and repetitive, in some cases, for the children and adolescents (Ruchkin, Koposov, & Schwab-Stone, 2007). The brevity of the measuring instrument is a characteristic that enables people to engage more with it and value it more positively. It can be used in assessment situations where there is a lack of time or economic resources, or where it is necessary to carry out a more holistic evaluation, not only considering variables related to the mere psychopathological exploration of emotional and behavioural problems.

In this regard, as shown in Table 1, using the SDQ allows us to obtain reliable scores as it is a short questionnaire and it is easy to administer, correct and interpret (Ruchkin, Jones, Vermeiren, & Schwab-Stone, 2008; Vostanis, 2006). At the same time, it is a screening tool that could be of significant value for school psychologists with regard to practices aimed at prevention and public healthcare (Hoagwood & Johnson, 2003). Finally, the SDQ multi-informant system is a screening tool available for free use on the internet (http://www.sdqinfo.com/). From the website you can download the SDQ in different formats and languages, together with the systems of scoring and correcting and various supplementary materials (e.g., syntax for SPSS).

TABLE 1
COMPARISON OF THE MAIN FEATURES OF THE RUTTER QUESTIONNAIRES, ACHENBACH SYSTEM OF EMPIRICALLY BASED ASSESSMENT (ASEBA) AND THE STRENGTHS AND DIFFICULTIES QUESTIONNAIRE (SDQ)

	SDQ	Rutter	ASEBA
Number of items	25-34	More than 100	More than 100
Positive items	+		
Versions:			
Parents	+	+	+
Teachers	+	+	+
Self-report	+	-	+
Follow-up version available	+	-	-
Coverage of:			
Behavioural problems	+	+	+
Emotional symptoms	+	+	+
Hyperactivity/attention deficit	+	+	+
Peer relationships	+	+	+
Prosocial behaviour	+	-	-
Impact of symptoms	+	-	-

The Difficulties And Capabilities Questionnaire comprises a total of five dimensions or sub-scales (Goodman, 1997), namely: 1) Emotional Problems, 2) Behavioural Problems, 3) Peer Problems, 4) Hyperactivity, and 5) Prosocial Behaviour. Each dimension is evaluated using five items. The first four subscales constitute a Total Difficulties score. The simple version of the SDQ, with its 25 items, is complemented with an extended version called the *impact supplement*, aimed at parents, as well as teachers and the children/adolescents themselves. In both versions there is a Likert response format with three options: No, not at all; Sometimes; Yes, always (scoring 0, 1 and 2, respectively). It is true however, that other response formats have also been used in the literature (e.g., Likert format with five options, according to the degree of adhesion) (Ortuño-Sierra et al., 2015), with the aim of improving the reliability of the scores of the different facets that make up the SDQ in its self-report version.

Goodman (1997) established a set of criteria for the construction of the SDQ which corresponds to its final form. It must: a) not be longer than one page; b) meet at least an age range between 4 and 17 years old; c) have identical versions for parents and teachers, and a very similar self-report version (11-16 years); d) address both the challenges and the strengths of the person; and e) have the same number of items in each dimension of the measuring instrument (Goodman, 1997).

As mentioned there are three versions of the SDQ: one for parents, one for teachers and another self-report version. The versions for parents and teachers are intended for children and adolescents aged 4 to17 years old, while the self-report version is recommended to be administered from the age of 11, because at this age a level of introspection is assumed that is necessary to complete the assessment. There is also an extended version (SDQ Extended Version) (Goodman, 1997), as well as versions for parents and teachers which cover only the ages of 3-4 years in which the items corresponding to antisocial behaviour have been replaced with items that measure opposition to rules, due to their greater adjustment to the characteristics of this stage of development.

THE PSYCHOMETRIC PROPERTIES OF THE SDQ

Psychometric studies on the SDQ internationally

The psychometric properties of the SDQ, in its different versions have been analysed extensively (see Brown, 2006; Kersten et al., 2015; Niclasen et al., 2012). Estimating the reliability of the scores has found adequate levels of internal consistency in most studies. However, the Behavioural Problems subscale and, especially, the subscale of Peer Problems show in some cases levels lower than 0.70 (Essau et al., 2012; Goodman, 2001; Mieloo et al., 2014; Niclasen, Skovgaard, Andersen, Somhovd, & Obel, 2013; Ortuño-Sierra, Fonseca-Pedrero, Paino, Sastre i Riba, & Muñiz, 2015b; Ortuño-Sierra et al., 2015c; Ruchkin et al., 2008; Ruchkin et al., 2007; Stevanovic et al., 2014; Sveen, Berg-Nielsen, Lydersen, & Wichstrøm, 2013; Theunissen, Vogels, De Wolff, & Reijneveld, 2013; Williamson et al., 2014; Yao et al., 2009). For example, in the study by Rothenberg et al. (2008) a Cronbach's alpha of 0.82 was found for the Total Difficulties score, while values for the subscales of Behavioural Problems and Peer Problems were the lowest, with values of 0.58 and 0.62, respectively. Other studies have examined the test-retest reliability of the SDQ (Borg, Pälvi, Raili, Matti, & Tuula, 2012; Downs, Strand, Heinrichs, & Cerna, 2012; Mellor, 2004; Svedin & Priebe, 2008), obtaining adequate values ranging between 0.47 and 0.76.

With regards to the analysis of the internal structure of the SDQ by exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), the different investigations carried out on the three versions of the SDQ reveal a five-factor structure as the most suitable (Downs et al., 2012; Niclasen et al., 2012; Ortuño-Sierra et al., 2015b; Richter, Sagatun, Heyerdahl, Oppedal, & Røysamb, 2011; Ruchkin et al., 2008; Stevanovic et al., 2014; Van Roy, Veenstra, & Clench-Aas, 2008; Williamson et al., 2014; Yao et al., 2009). This five-factor model corresponds to the scales of the SDQ: Emotional Problems, Behavioural Problems, Peer Problems, Hyperactivity and Prosocial. However, other studies have shown that the five-dimensional structure does not fit the data well, suggesting a four-factor solution as the most appropriate (Muris, Meesters, Eijkelenboom, & Vincken, 2004) and, in some cases, a three-factor one (Goodman, Lamping, & Ploubidis, 2010). On the other hand, a recent study proposes the inclusion of a bifactor model as the most relevant for explaining the factorial structure underlying the SDQ scores (Caci, Morin, & Tran, 2015). The bifactor model postulates that in addition to the five factors above (or the specific factors), a general factor can be added that explains the variability of scores on the SDQ items. Table 2 presents a selective review of the scientific literature published on the analysis of the factorial structure of the SDQ, both the versions for parents and teachers, and the self-report version.

Also, different validity evidence has been obtained in the previous research. For example, in its version for parents and teachers the SDQ has shown evidence of concurrent validity with different measuring instruments and diagnostic interviews (Downs et al., 2012; Mieloo et al., 2014; Theunissen et al., 2013). Likewise, evidence of the discrimination capacity of the SDQ has been proven in several studies (De Giacomo et al., 2012; Petermann, Petermann, & Schreyer, 2010). For example, a recent study shows the usefulness of the SDQ as a screening tool in the child-youth population, noting adequate levels of diagnostic sensitivity for internalising behavioural problems (Silva, Osorio, & Loureiro, 2015). Similarly, a longitudinal study conducted in the UK with children aged between 3 and 7 years old (Croft, Stride, Maughan, & Rowe, 2015), showed the predictive validity of the SDQ in detecting problems such as autism spectrum disorders or attention deficit hyperactivity disorder (ADHD). However, it is equally true that other research shows inadequate levels of sensitivity and specificity (Bekker, Bruck, & Sciberras, 2013; Mathai, Anderson, & Bourne, 2004). In general terms, there is sufficient empirical evidence supporting the validity of the SDQ as a tool for detection and screening in the child-youth population.

Psychometric studies of the SDQ nationwide

As we have seen, a large number of works have studied the psychometric properties of the SDQ both in Europe, and in America and Asia, though for the moment, there have been few studies in Spain and in Spanish-speaking countries (Ortuño-Sierra et al., 2015b). Some studies focus on the analysis of the psychometric properties of the Spanish version of the SDQ (García et al., 2000) in the child population and the versions for parents and teachers, in both cases revealing a structure of five factors as the most appropriate (Ezpeleta, Granero, de la Osa, Penelo, & Doménech, 2012; Fajardo et al., 2012; Gómez-Beneyto et al., 2013; Rodríguez-Hernández et al., 2012). For example, in the study by Rodríguez-Hernández et al. (2012), conducted with 595 children aged 7-10 years and administered to parents and teachers, a five-factor structure was found to be the most appropriate through principal component analysis and CFA. Another study conducted in Spain by Ezpeleta et al. (2012), with three year olds, revealed a factorial structure of five factors with two second order factors, which include the Internalising factor (Emotional Problems and Peer Problems) and the Externalising factor (Behavioural Problems and Hyperactivity), as the most appropriate for explaining the underlying dimensionality of the scores, both in the version for parents and the one for teachers.

Moreover, a recent study has highlighted the validity of the instrument as a tool for detection of ADHD in the version for parents (Carballo, Rodriguez-Blanco, Garcia-Nieto, & Baca-Garcia, 2014). The SDQ has shown evidence of discriminant validity in the Spanish version, obtaining the optimal diagnosis point 20 in the Total Difficulties score, which is the one that reveals the best values of sensitivity (0.96) and specificity (0.95) (Fajardo et al., 2012). The normative values in the parent version of the SDQ have been calculated and are available for use in Spain (Barriuso-Lapresa, Hernando-Arizaleta, & Rajmil, 2014).

With regards to the psychometric properties of the SDQ in its self-report version, various studies have found evidence of its validity and adequate levels of internal consistency for use in adolescents (Fajardo et al., 2012; Ortuño-Sierra, Chocarro, Fonseca-Pedrero, Sastre i Riba, & Muñiz, 2015a; Ortuño-Sierra et al., 2015b). As is the case with the versions for parents and teachers, the self-report version reveals a five-factor structure as the most appropriate. For example, in the study by Ortuño-Sierra et al. (2015a) the five-factor structure revealed goodness of fit indices superior to the three-factor model, however, various modifications to the original model were necessary to achieve optimal goodness of fit indices. Similarly a bifactor model (Caci et al., 2015) has been proposed as an alternative, although it is also true that its suitability has not yet been confirmed (Ortuño-Sierra et al., 2015a).

In conclusion, the SDQ is an interesting and useful tool for the measurement and detection of emotional and behavioural problems in this sector of the population. Most of the psychometric properties have been proven in the different versions within Spain for use by professionals as a screening tool in educational and/or care centres as well as in research. Future studies should continue to analyse and seek new evidence of validity which will allow us to make informed decisions and make inferences from the scores obtained with the SDQ.

INFLUENCE OF GENDER AND AGE ON SDQ SCORES

The studies analysed in this section refer to the impact of gender and age on the phenotypic expression of emotional and behavioural difficulties as well as prosocial behaviour, assessed with the SDQ.

As seen in Table 3, in terms of gender, the majority of the studies reviewed internationally find that females earn higher mean scores than males in Emotional and Prosocial Behaviour; however, males tend to earn higher mean scores than females in Behaviour Problems, Hyperactivity and/or Relationship Problems (Di Riso et al., 2010; Giannakopoulos et al., 2009; Koskelainen, Sourander, & Vauras, 2001; Ortuño et al., 2014; Svedin & Priebe, 2008; Van Roy, Grøholt, Heyerdahl, & Clench-Aas, 2006; van Widenfelt, Goedhart, Treffers, & Goodman, 2003; Yao et al., 2009). For example, in the study conducted by Giannakopoulos et al. (2009) on Greek adolescents, the mean scores were found to be higher in Prosocial Behaviour and Emotional Problems in girls but not in Behavioural Problems, Hyperactivity or Relationships among boys.

Other studies reveal results that contradict the previous ones, as is the case of the study in Finland conducted by Koskelainen et al. (2001) with

TABLE 2
THE MAIN STUDIES THAT ANALYSE THE INTERNAL STRUCTURE OF THE SCORES OF THE STRENGTHS AND DIFFICULTIES QUESTIONNAIRE (SDQ)

Study	Sample				
	Nationality	*N* / Age Range	SDQ Version	Type of factor analysis	Number of factors found
Koskelainen, Sourander & Kaljonen (2000)	Finland	735 / 7-15	SDQ (P,T,S)	PCA	5
Thabet, Stretch & Vostanis (2000)	Gaza	322 / 3-16	SDQ (P,T,S)	CFA	5
Goodman (2001)	Great Britain	10438 / 5-15	SDQ (P,T,S)	PCA	5
Koskelainen, Sourander & Vauras (2001)	Finland	1458 / 13-17	SDQ (S)	PCA	53
Muris, Meesters & van den Berg (2003)	Holland	562 / 9-15	SDQ (P,T,S)	PCA	5
Becker et al., (2004)	Germany	214 / 11-17	SDQ (S)	PCA	5
Dickey & Blumberg (2004)	USA	9574 / 4-17	SDQ (P)	EFA, PCACFA	3
Muris, Meesters, EijKelenboom & Vincken (2004)	Holland	1111 / 8-13	SDQ (S)	PCA	45
Rønning, Helge Handegaard, Sourander & Mørch (2004)	Norway	4167 / 11-16	SDQ (S)	CFA	5 but with poor fit
Kashala, Elgen, Sommerfelt & Tylleskar (2005)	Congo	1187 / 7-9	SDQ (T)	PCA	52 of Hyperactivity
Mojtabai (2006)	USA	8034	SDQ (P)	CFA	3
	U.K.	7970 / 5-16	SDQ (P)	CFA	3
Van Leeuwen, Meerschaert, Bosmans, De medts & Braet (2006)	Germany	3179 / 4-8	SDQ (P, T)	CFA EFA	35
Mellor & Stokes (2007)	Australia	914 / 7-17	SDQ (P,T,S)	CFA	5 with poor fit
Palmieri & Smith (2007)	USA	733 / *M*= 56.1	SDQ (P)	PCA	34 better
Ruchkin, Koposov & Schwab-Stone (2007)	Russia	2892 / 13-18	SDQ (S)	CFA	5
Mazur, Tabak & Kololo (2007)	Poland	774 / 14	SDQ (S)	EFA	5
d'Acremont (2008)	Switzerland	557 / Adolescents	SDQ (T)	CFA	5
Matsuishi et al. (2008)	Japan	2899 / 4-12	SDQ (P)	EFA	5
Percy, McCrystal & Higgins (2008)	Ireland	3753 / 12	SDQ (S)	EFA CFA	EFA: 3 CFA: 5 questionable
Rothenberg et al. (2008)	Germany	2406 / 7-16	SDQ (P,S)	EFA CFA	5
Ruchkin, Jones, Vermeiren & Schwab-Stone (2008)	USA	>5000 / 13-14	SDQ (S)	EFA and PCACFA	53 better
Svedin & Priebe (2008)	Sweden	1015 / 17-19	SDQ (S)	CFA	75
Van Roy, Veenstra & Clench-Aas (2008)	Norway	26269 / 10-19	SDQ (S)	CFA	5
Giannakopoulos et al. (2009)	Greece	1194 / 11-17	SDQ (P,S)	CFA	5
Sanne, Torsheim, Heiervang & Stormark (2009)	Norway	6430 Parents / 8999 Teachers	SDQ (P,T)	CFA EFA	5 better3
Yao et al., (2009)	China	1135 / 11-18	SDQ (S)	CFA	5
Di Riso (2010)	Italy	1394 / *M*= 9.04	SDQ (P)	CFA	3
Goodman, Lamping & Ploubidis (2010)	Great Britain	18222 / 5-16	SDQ (P,T,S)	PCA CFA	35 and 2 of second order
Petermann, Petermann & Schreyer (2010)	Germany	1738 / 3-5	SDQ (P)	CFA	5
Stone, Otten, Engels, Vermults & Janssens (2010)	Review of 48 studies	131223 / 4-12	SDQ	CFA	8 studies= 45 studies= 5
Richter, Sagatun, Heyerdahl, Oppedal & Røysamb (2011)	Norway	>6000 / 15-16	SDQ (S)	CFA	5
Van de Looij-Jansen, Goedhart, de Wilde & Treffers (2011)	Holland	11881 / 11-16	SDQ (S)	CFA	5 better4
Ezpeleta, Granero, de la Osa, Penelo & Domènech (2012)	Spain	1341 / 3-4	SDQ (P, T)	CFA	55 and 2 second order
Gómez (2012)	Australia	2021 / 2-17	SDQ (P,T,S)	CFA	5
Mieloo et al., (2012)	Germany	5514 / 5-6	SDQ (P,T)	CFA	5
Nidasen, Teasdale, Andersen, Skovgaard, Elberling & Obel (2012)	Denmark	71840 / 5-12	SDQ (P,T)	CFA	55 and 2 second order
Rodríguez-Hernández et al. (2012)	Spain	595 / 7-10	SDQ (P, T)	PCA CFA	5
Ruchkin, Koposov, Vermeiren & Schwab-Stone (2012)	Russia	528 / 13-18	SDQ	CFA	5
Essau et al. (2012)	5 European countries	2418 / 12-17	SDQ (S)	CFA	35
Shevlin, Murphy & McElearney (2012)	Ireland	202	SDQ (P, S)	CFA	5
Liu, Chien, Shang, Lin, Liu & Gau (2013)	China	3534 / 6-15	SDQ (P,T,S)	PCA	4 (P,T)5 (S)
He, Burstein, Schmitz & Merikangas (2013)	USA	6483 / 13-18	SDQ (P)	CFA	5
Theunissen et al. (2013)	Holland	839 / 3-4	SDQ (P)	CFA	5

TABLE 2
THE MAIN STUDIES THAT ANALYSE THE INTERNAL STRUCTURE OF THE SCORES OF THE STRENGTHS AND DIFFICULTIES QUESTIONNAIRE (SDQ) (Continued)

Study	Sample				
	Nationality	*N* / Age Range	SDQ Version	Type of factor analysis	Number of factors found
Sveen et al. (2013)	Norway	845 / 4	SDQ (P, T)	CFA	5
Williamson et al. (2014)	Australia	717 / 4-17	SDQ (P)	CFA	5
Stevanovic et al. (2014)	7 European, African and Asian countries	2367 / 13-18	SDQ (S)	CFA	Bifactor and 5
Ortuño-Sierra et al. (2015a)	Spain	1547 / 11-19	SDQ (S)	CFA	5
Ortuño-Sierra et al. (2015b)	5 European countries	3012 / 12-17	SDQ (S)	CFA	5
Caci, Morin & Tran (2015)	France	889 / 4-17	SDQ (P)	CFA	Bifactor

Note. SDQ (P,T,S): *Strengths and Difficulties Questionnaire* (*Parent, Teacher, Self-Report*); PCA: Principal Component Analysis; CFA: Confirmatory Factor Analysis; EFA: Exploratory Factor Analysis; CBCL: *Child Behaviour Checklist*; YSR: *Youth Self Report.*

a sample of 1458 adolescents aged 13-17 years, which showed mean scores that were significantly higher in girls in the total difficulties and problems associated with hyperactivity compared with boys. More recently, Reinholdt-Dunne et al. (2011), in a sample of 834 Danish adolescents aged 12-14 years, found higher mean scores on the subscale Emotional Problems in favour of girls, with no statistically significant differences according to gender in the rest of the subscales.

Regarding age, the results are more inconsistent than in the case of gender. Some studies show an increase in problems with increasing age (Giannakopoulos et al., 2009; Koskelainen et al., 2001; Rønning, Helge Handegaard, Sourander, & Mørch, 2004; Yao et al., 2009), while others find a reverse trend (Muris, Meesters, & van den Berg, 2003; Van Roy et al., 2006), or they do not find any association (Prior, Virasinghe, & Smart, 2005). Theoretically, it is speculated that adolescents are more exposed to the presentation of behavioural or relational problems in the early years, giving way at the beginning of middle and late adolescence to a greater capacity for problem management, behavioural regulation and control, management of social behaviours and increased capacity for prosocial behaviours (see Table 3).

For example, the study by Van Widenfelt et al. (2003), with 970 German teenagers aged between 11 and 16 years old, showed higher mean scores on Emotional Problems, Behavioural Problems and Hyperactivity among the younger participants. Similarly, Armand et al. (2012), in their study with 2,000 Iranian children and adolescents aged 6-18 years old, found that the problems of Hyperactivity and Total Difficulties were higher among those of a younger age. In Norway, Lien, Green, Welander-Vatn and Bjertness (2009), with a sample of 3,790 schoolchildren aged 15-19, found higher scores for internalising problems in older participants, whereas externalising problems were more frequent among the younger participants. However, as noted, other research contradicts the above in part. For example, the results achieved in Italy by di Riso et al. (2010) showed a greater number of Relationship Problems in the older participants.

Moreover, the literature reviewed includes the interrelationship between gender and age, revealing, for example, that the levels of total difficulties increase with age for females, while males show the opposite trend (Van Roy, Grøholt, Heyerdahl, & Clench-Aas, 2010) or higher levels of prosocial behaviour at higher ages in males (Rønning et al., 2004). Other studies show more emotional problems with increasing age in females (Armand et al., 2012; d'Acremont & Van der Linden, 2008).

RECAPITULATION

A significant percentage of children and adolescents present mental health difficulties throughout their life, potentially having a clear impact not only on the personal areas but also at the health and economic levels (Blanchard, Gurka, & Blackman, 2006; Domino et al., 2009; Drabick & Kendall, 2010; Polanczyk et al., 2015; Simpson, Bloom, Cohen, Blumberg, & Bourdon, 2005). Among the different measuring instruments available for the assessment and detection of psychological difficulties in children is the Strengths and Difficulties Questionnaire (SDQ). The SDQ has a number of features that make it interesting for use by mental health practitioners. These include, for example, the fact that it is an instrument that is easy to access, available for free on the Internet, its brevity, its ease of administration and correction, the fact that it provides a multi-informant system, the inclusion of prosocial type behaviours, and its adequate psychometric properties.

The reliability studies reviewed found adequate levels of reliability in the SDQ scores, although the subscales of Behavioural Problems and Peer Problems show, in some cases, discrete or moderate levels. Data has also been obtained regarding the stability of the scores. We have collected various sources of validity evidence for the SDQ. The dimensional structure of the SDQ scores seems to be able to be explained through a five-factor model, although it is also true that other models (e.g., a two-factor model of second order or a bifactor model) are also factorial solutions for which some empirical support has been found. Similarly, other studies have analysed different sources of validity in relation to external variables, and adequate levels of sensitivity and specificity were obtained as well as the prediction of various mental health problems in child and adolescent population.

Moreover, in view of the different studies conducted with the SDQ, despite the existence of research that shows inconsistent results, there

TABLE 3
THE MAIN STUDIES CONCERNING GENDER AND AGE WITH THE STRENGTHS AND DIFFICULTIES QUESTIONNAIRE

Study	Sample		SDQ Version	Gender		Age	
	Nationality	N Age Range		Males Higher Score	Females Higher Score	Higher age	Lower age
Koskelainen, Sourander & Vauras (2001)	Finland	1458 13-17	SDQ (S)	Behavioural Peers	Prosocial Hyperactivity Emotional Total	Emotional	
Knyazev et al. (2003)	Russia	146 7-17	SDQ (P,T, S)	Hyperactivity	Prosocial		
Muris, Meesters & van den Berg (2003)	Holland	562 9-15	SDQ (P,T,S)	Behavioural	Emotional Prosocial		Total Peers
Van Widenfelt, Goedhart, Treffers & Goodman (2003)	Germany	1476 11-16	SDQ (P,T,S)	Behavioural Hyperactivity	Emotional Prosocial		Emotional Hyperactivity Peers
Becker et al. (2004)	Germany	214 11-17	SDQ (S)	Behavioural	Prosocial Emotional	Emotional Prosocial	
Muris, Meesters, EijKelenboom & Vincken (2004)	Holland	1111 8-13	SDQ (S)	Behavioural	Emotional Prosocial		
Rønning, Helge Handegaard, Sourander & Mørch (2004)	Norway	4167 11-16	SDQ (S)	Behavioural Hyperactivity	Emotional Prosocial		
Bourdon, Goodman, Rae, Simpson & Koretz (2005)	USA	10367 4-17	SDQ (P)	Total			
Kashala, Elgen, Sommerfelt & Tylleskar (2005)	Congo	1187 7-9	SDQ (T)	Total Behavioural	Prosocial		
Prior, Virasinghe & Smart (2005)	Sri Lanka	1809 11-13	SDQ (P,T,S)	Total Behavioural	Prosocial	No differences	No differences
Simpson, Bloom, Cohen, Blumberg & Bourdon (2005)	USA	> 25000 4-17	SDQ (P,S)	Total		Total	
Becker et al. (2006)	European countries	1573 ADHD *M*=8,8	SDQ (P)		Emotional Prosocial		Total Hyperactivity Peers
Mojtabai (2006)	USA UK	8034 7970 5-16	SDQ (P)	Behavioural	Emotional		
Thabet, Karim & Vostanis (2006)	Gaza	309	SDQ (P)	Hyperactivity			
Van Leeuwen, Meerschaert, Bosmans, De medts & Braet (2006)	Germany	3179 4-8	SDQ (P,T)	Total Behavioural Hyperactivity	Prosocial	Hyperactivity Emotional Total	 Behavioural
Van Roy, Grøholt, Heyerdahl & Clench-Aas (2006)	Norway	29631 10-19	SDQ (S)	Behavioural Peers	Emotional		Behavioural Peers
Capron, Therond, & Duyme (2007)	France	1400 *M*=12,8	SDQ (P, S)	Behavioural Hyperactivity Peers	Prosocial Emotional		
d'Acremont & Van der Linden (2008)	Switzerland	557 13-18	SDQ (T)	Behavioural Peers Hyperactivity	Prosocial		
Du, Kou, & Coghill (2008)	China	2655 3-17	SDQ (P,T)	Hyperactivity Prosocial Behavioural			
Matsuishi et al. (2008)	Japan	2899 4-12	SDQ (P)	Total Peers Hyperactivity Behavioural	Emotional Prosocial		
Ravens-Sieberer et al. (2008)	Germany	2863 7-17	SDQ (P)				
Rothenberg et al. (2008)	Germany	2406 7-16	SDQ (P,S)	Total Behavioural Hyperactivity Peers	Emotional	Prosocial	Hyperactivity Total

TABLE 3
THE MAIN STUDIES CONCERNING GENDER AND AGE WITH THE STRENGTHS AND DIFFICULTIES QUESTIONNAIRE (Continued)

Study	Sample		SDQ Version	Gender		Age	
	Nationality	N Age Range		Males Higher Score	Females Higher Score	Higher age	Lower age
Shojai, Wazana, Pitrou & Kovess (2008)	France	1348 6-11	SDQ (P)	Hyperactivity Behavioural	Prosocial		
Svedin & Priebe (2008)	Sweden	1015 17-19	SDQ (S)	Behavioural Peers	Emotional Prosocial		
Giannakopoulos et al. (2009)	Greece	1194 11-17	SDQ (P,S)		Prosocial Emotional	Hyperactivity Behavioural	
Lien, Green, Welander-Vatn & Bjertness (2009)	Norway	3790 15-19	SDQ			Internalising	Externalising
Ullah-Syed, Abdul-Hussein & Haidry (2009)	Pakistan	675 5-11	SDQ (P,T)	Behavioural Hyperactivity Total	Emotional		
Yao et al., (2009)	China	1135 11-18	SDQ (S)	Behavioural Peers	Emotional	Hyperactivity Prosocial	Peers
Di Riso et al. (2010)	Italy	1394 M= 9.04	SDQ (P)	Behavioural Hyperactivity	Prosocial Emotional	Peers	
Keskin & Çam (2010)	Turkey	38411-16	SDQ	Peers	Emotional Prosocial	Hyperactivity	Prosocial
Lai et al. (2010)	Hong Kong	> 4000 6-12	SDQ (P)	Behavioural Hyperactivity Peers	Emotional Prosocial		
Van Roy, Groholt, Heyerdahl & Clench-Aas (2010)	Norway	8154 10-13	SDQ (P,S)	Behavioural Hyperactivity Peers	Emotional Prosocial		
Fonseca-Pedrero, Paíno, Lemos-Giráldez & Muñiz (2011)	Spain	1319 13-17	SDQ (S)	Behavioural Hyperactivity Peers	Emotional Prosocial	Hyperactivity Total	
Reinholdt-Dunne et. (2011)	Denmark	834 12-14	SDQ (S)		Emotional		
Wichstrøm et al. (2012)	Norway	2475 4	SDQ (P)	Hyperactivity			
Arman, Keypour, Maracy & Attari (2012)	Iran	2000 6-18	SDQ (P)	Behavioural Hyperactivity	Emotional		Hyperactivity Total
Mieloo et al. (2012)	Germany	5514 5-6	SDQ (P,T)	Total Behavioural Hyperactivity			
Ruchkin, Koposov, Vermeiren & Schwab-Stone (2012)	Russia	528 13-18	SDQ (T)	Behavioural Hyperactivity			
Shoval et al. (2012)	Israel	1402 14-17	SDQ (P)	Externalising			
Liu et al. (2013)	China	3534 6-15	SDQ (P,T,S)	Emotional (S) Behavioural Peers Hyperactivity	Prosocial		Behavioural Peers
Armand, Amel & Maracy (2013)	Iran	1934 11-18	SDQ (S,P)	Behavioural Hyperactivity	Emotional	Emotional Behavioural Total (S)	Prosocial Total (P)
Sveen et al. (2013)	Norway	845 4	SDQ (P, T)	Behavioural	Emotional		
Barriuso-Lapresa, Hernando-Arizaleta & Rajmil (2014)	Spain	6266 4-15	SDQ (P)	Hyperactivity	Emotional Prosocial		Total Behavioural
Ortuño-Sierra et al. (2014)	Spain	508 11-18	SDQ (S)	Behavioural	Emotional Prosocial	Emotional Behavioural Hyperactivity Total	

Note. SDQ (P,T,S): *Strengths and Difficulties Questionnaire* (*Parent, Teacher, Self-Report*); Emotional: Emotional Problems; Behavioural: Behavioural Problems; Peers: Peer Problems.

appears to be some consensus that externalising problems, such as behavioural problems and hyperactivity are more common among boys during adolescence. However, emotional problems are more common among girls, who also show higher values of prosocial behaviour. In terms of age, the results are more inconsistent, and there is research that reflects an increase in difficulties with increasing age as well as other studies that reveal the opposite; consequently, the heterogeneity of the results prevents us from being able to specify what kind of difficulties are more typical during early or late adolescence. There are also no conclusive results regarding the degree of presentation of prosocial type behaviours in relation to age.

In conclusion, the study of psychological adjustment and difficulties during these developmental stages is a subject of great importance given the impact and repercussions that these problem have on multiple levels (e.g., social, family, healthcare, etc.). The assessment and accurate detection of such difficulties is of great importance with regards to a possible early prevention and to avoid their potential consolidation in adulthood. Vulnerable or "high risk" groups should be identified as early as possible during childhood and adolescence in order to develop effective preventive interventions to prevent, mitigate or reduce the overall burden and the associated morbidity, and ultimately to help to improve one of the main causes of disability in our society.

ACKNOWLEDGEMENTS

This research was funded by the Ministry of Science and Innovation of Spain (MICINN) (reference PSI2014-56114-P), by the Carlos III Institute, the Biomedical Research Networking Centre in Mental Health (CIBERSAM) and the 2015 Call for BBVA Foundation grants for researchers and cultural creators.

REFERENCES

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist 4-18 and 1991 profile*. Burlington: University of Vermont Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Teacher's Report Form and 1991 profile*. Burlington: University of Vermont Department of Psychiatry.

Achenbach, T. M. (1991c). *Manual for the Youth Self -Report and 1991 profile*. Burlington: University of Vermont Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms y Profiles*. University of Vermont, Research Center for Children, Youth, and Families: Burlington, VT.

Achenbach, T. M., & Rescorla, L. A. (2007). *Multicultural supplement to the Manual for the ASEBA School-Age Forms and Profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families.

Aebi, M., Giger, J., Plattner, B., Winkler Metzke, C., & Steinhausen, H.-C. (2014). Problem coping skills, psychosocial adversities and mental health problems in children and adolescents as predictors of criminal outcomes in young adulthood. *European Child* and *Adolescent Psychiatry, 23*(5), 283-293.

Angold, A., Messer, S. C., Stangl, D., Farmer, E. M. Z., Costello, E. J., & Burns, B. J. (1998). Perceived parental burden and service use for child and adolescent psychiatric disorders. *American Journal of Public Health, 88*, 75-80.

Armand, S., Amel, A. K., & Maracy, M. R. (2013). Comparison of parent adolescent scores on Strengths and Difficulties Questionnaire. *Journal of Research in Medical Science, 18*, 501-505.

Armand, S., Keypour, M., Maracy, M. R., & Attari, A. (2012). Epidemiological study of youth mental health using Strengths and Difficulties Questionnaire (SDQ). *Iranian Red Crescent Medical Journal, 14*, 371-375.

Barriuso-Lapresa, L. M., Hernando-Arizaleta, L., & Rajmil, L. (2014). Reference values of the Strengths and Difficulties Questionnaire (SDQ) version for parents in the Spanish population, 2006. *Actas Españolas de Psiquiatría, 42*, 43-48.

Becker, A., Hagenberg, N., Roessner, N., Woerner, W., & Rothenberg, A. (2004). Evaluation of the self-reported SDQ in a clinical setting: Do self-reports tell us more than ratings by adult informants? *European Child* and *Adolescent Psychiatry, 13*(2), 17-24.

Becker, A., Steinhausen, H.-C., Baldursson, G., Dalsgaard, S., Lorenzo, M. J., Ralston, S. J., . . . Rothenberger, A. (2006). Psychopathological screening of children with ADHD: Strengths and Difficulties Questionnaire in a pan-European study. *European Child and Adolescent Psychiatry, 15(suppl. 1)*, 56-62.

Bekker, J., Bruck, D., & Sciberras, E. (2013). Congruent validity of the Strength and Difficulties Questionnaire to screen for comorbidities in children with ADHD. *Journal of Attention Disorders*. doi:10.1177/1087054713496462

Blanco, C., Wall, M. M., He, J. P., Krueger, R. F., Olfson, M., Jin, C. J., . . . Merikangas, K. R. (2015). The space of common psychiatric disorders in adolescents: comorbidity structure and individual latent liabilities. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*, 45-52.

Blanchard, L. T., Gurka, M. J., & Blackman, J. A. (2006). Emotional, developmental, and behavioral health of American children and their families: a report from the 2003 National Survey of Children's Health. *Pediatrics, 117*, 1202-1212.

Bones, K., Pérez, K., Rodríguez-Sanz, M., Borrell, C., & Obiols, J. E. (2010). Prevalencia de problemas de salud mental y su asociación con variables socioeconómicas, de trabajo y salud: resultados de la Encuesta nacional de Salud de España [The prevalence of mental health problems and their association with socioeconomic, work and health variables: results from the National Health Survey of Spain]. *Psicothema, 22*, 389-395.

Borg, A.-M., Pälvi, K., Raili, S., Matti, J., & Tuula, T. (2012). Reliability of the Strengths and Difficulties Questionnaire among Finnish 4-9-year-old children. *Nordic Journal of Psychiatry, 66*, 403-413.

Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G., & Koretz, D. S. (2005). The Strengths and Difficulties Questionnaire: U.S. Normative Data and Psychometric Properties. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 557-564.

Brimblecombe, N., Knapp, M., Murguia, S., Mbeah-Bankas, H., Crane, S., Harris, A., . . . King, D. (2015). The role of youth mental health services in the treatment of young people with serious mental illness: 2-year outcomes and economic implications. *Early Intervention in Psychiatry*. doi:10.1111/eip.12261

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press.

Caci, H., Morin, A. J., & Tran, A. (2015). Investigation of a bifactor model of the Strengths and Difficulties Questionnaire. *European Child Adolescent Psychiatry, 24*,1291-301.

Capron, C., Therond, C., & Duyme, M. (2007). Psychometric Properties of the French Version of the Self-Report and Teacher Strengths and

Difficulties Questionnaire (SDQ). *European Journal of Psychological Assessment, 23*, 79-88.

Carballo, J. J., Rodriguez-Blanco, L., Garcia-Nieto, R., & Baca-Garcia, E. (2014). Screening for the ADHD phenotype using the Strengths and Difficulties Questionnaire in a clinical sample of newly referred children and adolescents. *Journal of Attentional Disorders*. doi:1087054714561858

Copeland, W., Shanahan, L., & Costello, J. (2011). Cumulative prevalence of psychiatric disorders by young adult: A prospective cohort analysis from great smoky mountains study. *Journal American Academy Child Adolescent Psychiatry, 50*, 252-261.

Costello, E. J., Copeland, W., & Angold, A. (2011). Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults? *Journal Of Child Psychology and Psychiatry, 52*, 1015-1025.

Croft, S., Stride, C., Maughan, B., & Rowe, R. (2015). Validity of the strengths and difficulties questionnaire in preschool-aged children. *Pediatrics, 135*, e1210-1219.

Cuesta, M. J., Sanchez-Torres, A. M., Cabrera, B., Bioque, M., Merchan-Naranjo, J., Corripio, I., . . . Bernardo, M. (2015). Premorbid adjustment and clinical correlates of cognitive impairment in first-episode psychosis. The PEPsCog Study. *Schizophrenia Research, 164*, 65-73. doi:10.1016/j.schres.2015.02.022. Epub

Cullins, L. M., & Mian, A. I. (2015). Global child and adolescent mental health: A Culturally informed focus. *Journal of the Amercian Child and Adolescent Psychiatric Clinics of North America, 24*, 823-830.

Cummings, J. A., Harrison, P. L., Dawson, M. N., Short, R. J., Gorin, S., & Palomares, R. S. (2004). The 2002 conference on the future of school psychology: Implications for consultation, intervention, and prevention services *Journal of Educational and Psychological Consultation, 15*, 239-256.

d'Acremont, M., & Van der Linden, M. (2008). Questionnaire in a Community Sample of French-Speaking Adolescents. *European Journal of Psychological Assessment, 24*, 1-8.

Davies, S. J., Pearson, R. M., Stapinski, L., Bould, H., Christmas, D. M., Button, K. S., . . . Evans, J. (2015). Symptoms of generalized anxiety disorder but not panic disorder at age 15 years increase the risk of depression at 18 years in the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort study. *Psychological Medicine*, 1-13.

De Giacomo, A., Lamanna, A. L., Craig, F., Santoro, N., Goffredo, S., & Cecinati, V. (2012). The SDQ in Italian clinical practice: Evaluation between three outpatient groups compared. *Rivista di Psichiatria, 47*, 400-406.

Di Riso, D., Salcuni, S., Chessa, D., Raudino, A., Lis, A., & Altoè, G. (2010). The Strengths and Difficulties Questionnaire (SDQ). Early evidence of its reliability and validity in a community sample of Italian children. *Personality and Individual Differences, 49*, 570-575.

Diaz de Neira, M., Garcia-Nieto, R., de Leon-Martinez, V., Perez Fominaya, M., Baca-Garcia, E., & Carballo, J. J. (2015). Prevalence and functions of self-injurious thoughts and behaviors in a sample of Spanish adolescents assessed in mental health outpatient departments. *Revista de Psiquiatría y Salud Mental, 8*, 137-145.

Dickey, W. C., & Blumberg, S. J. (2004). Revisiting the Factor Structure of the Strengths and Difficulties Questionnaire: United States, 2001. *Journal American Academy Child Adolescent Psychiatry, 43*, 1159-1167.

Domino, M. E., Burns, B. J., Mario, J., Reinecke, M. A., Vitiello, B., Weller, E. B., . . . March, J. S. (2009). Service use and costs of care for depressed adolescents: who uses and who pays? *Journal of Clinical Child and Adolescent Psychology, 38*, 826-836.

Downs, A., Strand, P. S., Heinrichs, N., & Cerna, S. (2012). Use of the teacher version of the Strengths and Difficulties Questionnaire with German and American preschoolers. *Early Education and Development, 23*, 493-516.

Drabick, D. A., & Kendall, P. C. (2010). Developmental psychopathology and the diagnosis of mental health problems among youth. *Clinical Psychology: Science and Practice, 17*, 272-280.

Drancourt, N., Etain, B., Lajnef, M., Henry, C., Raust, A., Cochet, B., . . . Bellivier, F. (2013). Duration of untreated bipolar disorder: missed opportunities on the long road to optimal treatment. *Acta Psychiatrica Scandinavica, 127*, 136-144.

Du, Y., Kou, J., & Coghill, D. (2008). The validity, reliability and normative scores for the parent, teacher and self report versions of the strengths and Difficulties Questionnaire in China. *Child and Adolescent Psychiatry and Mental Health, 2*, 8. doi:10.1186/1753-2000-2-8

Essau, C. A., Olaya, B., Anastassiou-Hadjicharalambous, X., Pauli, G., Gilvarry, C., Bray, D., . . . Ollendick, T. H. (2012). Psychometric properties of the Strength and Difficulties Questionnaire from five European countries. *International Journal of Methods in Psychiatry Research, 21*, 232-245.

Ezpeleta, L., Granero, R., de la Osa, N., Penelo, E., & Doménech, J. M. (2012). Psychometric properties of the Strengths and Difficulties Questionnaire in 3-year-old preschoolers. *Comprehensive Psychiatry, 54*, 282-291.

Fajardo, F., León, B., Felipe, E., & Ribeiro, E. J. (2012). Salud mental en el grupo de edad 4-15 años a partir de los resultados de la Encuesta Nacional de Salud 2006 [Mental health in the age group of 4-15 years based on the results of the National Health Survey 2006]. *Revista Española de Salud Pública, 86*.

Fonseca-Pedrero, E., Paino, M., Lemos-Giraldez, S., & Muñiz, J. (2012). Patrones desadaptativos de la personalidad del cluster B en adolescentes españoles [Maladaptive patterns of cluster B personality in Spanish adolescents]. *Revista de Psiquiatría y Salud Mental, 6*, 129–138.

Fonseca-Pedrero, E., Paino, M., Lemos Giraldez, S., & Muñiz, J. (2011). Prevalencia de la sintomatología emocional y comportamental en adolescentes españoles a través del Strength and Difficulties Questionnaire (SDQ) [The prevalence of emotional and behavioural symptoms in Spanish adolescents through the Strength and Difficulties Questionnaire (SDQ)]. *Revista de Psicopatología y Psicología Clínica, 16*, 15-25.

Ford, T., Goodman, R., & Meltzer, H. (2003). The British Child and Adolescent Mental Health Survey 1999: The Prevalence of DSM-IV Disorders. *Journal of the American Academic of Child and Adolescence Psychiatry, 42*, 1203-1211. doi:10.1097/01.chi.0000081820.25107.ae

Ford, T., Hamilton, H., Meltzer, H., & Goodman, R. (2008). Predictors of service use for mental health problems among British school children. *Child and Adolescent Mental Health, 13*, 32-40.

García, P., Goodman, R., Mazaira, J., Torres, A., Rodríguez-Sacristán, J., Hervas, A., & Fuentes, J. (2000). El cuestionario de Capacidades y Dificultades [The Strengths and Difficulties Questionnaire]. *Revista de Psiquiatría Infanto-Juvenil, 1*, 12-17.

Giannakopoulos, G., Tzavara, C., Dimitrakaki, C., Kolaitis, G., Rotsika, V., & Tountas, Y. (2009). The factor structure of the Strengths and Difficulties Questionnaire (SDQ) in Greek adolescents. *Annals of General Psychiatry, 8*, 20.

Gómez-Beneyto, M., Nolasco, A., Moncho, J., Pereyra-Zamora, P., Tamayo-Fonseca, N., Munarriz, M., . . . Girón, M. (2013). Psychometric behaviour of the strengths and difficulties questionnaire (SDQ) in the Spanish national health survey 2006. *BMC Psychiatry, 13*(95). doi:10.1186/1471-244X-13-95

Gómez, R. (2012). Correlated trait–Correlated method minus one analysis of the convergent and discriminant validities of the Strengths and Difficulties Questionnaire. *Assessment, 21*, 372-382. d

Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to Use Broader Internalising and Externalising Subscales Instead of the Hypothesised Five Subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British Parents, Teachers and Children. *Journal of Abnormal Child Psychology, 38*, 1179-1191.

Goodman, R. (1994). A modified version of the Rutter Parent Questionnaire including extra items on children's strengths: a research note. *Journal of Child Psychology and Psychiatry, 35*, 1483-1494.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry, 38*, 581-586.

Goodman, R. (2001). Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academic of Child and Adolescence Psychiatry, 40*, 1337-1345.

Haro, J. M., Palacín, C., Vilagut, G., Martínez, M., Luque, I., Codony, M., . . . Alonso, J. (2006). Prevalencia de los trastornos mentales y factores asociados: resultados del estudio ESEMeD-España [The prevalence of mental disorders and associated factors: The results of the ESEMeD-Spain study]. *Medicina Clínica, 126*, 63-77.

He, J. P., Burstein, M., Schmitz, A., & Merikangas, K. R. (2013). The Strengths and Difficulties Questionnaire (SDQ): the factor structure and scale validation in U.S. adolescents. *Journal of Abnormal Child Psychology, 41*, 583-595.

Hill, C. R., & Hughes, J. N. (2007). An examination of the convergent and discriminant validity of the Strengths and Difficulties Questionnaire. *School Psychology Quarterly, 22*, 380-406.

Hoagwood, K., & Johnson, J. (2003). A public health framework: I. From evidence-based practices to evidence-based policies. *Journal of School Psychology, 41*, 3-22.

Kashala, E., Elgen, I., Sommerfelt, K., & Tylleskar, T. (2005). Teacher ratings of mental health among school children in Kinshasa, Democratic Republic of Congo. *European Child and Adolescent Psychiatry, 14*, 208-215.

Kersten, P., Czuba, K., McPherson, K., Dudley, M., Elder, H., Tauroa, R., & Vandal, A. (2015). A systematic review of evidence for the psychometric properties of the Strengths and Difficulties Questionnaire. *International Journal of Behavioral Development*, 1-12.

Keskin, G., & Çam, O. (2010). Adolescents' strengths and difficulties: approach to attachment styles. *Journal of Psychiatric and Mental Health Nursing, 17*, 433-441.

Klein, D. N., Shankman, S. A., Lewinsohn, P. M., & Seeley, J. R. (2009). Subthreshold depressive disorder in adolescents: predictors of escalation to full-syndrome depressive disorders. *Journal of American Academy of Child and Adolescent Psychiatry, 48*, 703-710.

Knyazeva, G. G., Slobodskayaa, H. R., Safronovaa, M. V., Sorokinb, O. V., Goodmanc, R., & Wilsonc, G. D. (2003). Personality, psychopathology and brain oscillations. *Personality and Individual Differences, 35*, 1331-1349.

Koskelainen, M., Sourander, A., & Kaljonen, A. (2000). The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. *European Child and Adolescent Psychiatry, 9*, 277-284.

Koskelainen, M., Sourander, A., & Vauras, M. (2001). Self-reported strengths and difficulties in a community sample of Finnish adolescents. *European Child and Adolescent Psychiatry, 10*, 180-185.

Lai, K. Y. C., Luk, E. S. L., Leung, P. W. L., Wong, A. S. Y., Law, L., & Ho, K. (2010). Validation of the Chinese version of the strengths and difficulties questionnaire in Hong Kong. *Social Psychiatry and Psychiatric Epidemiology, 45*, 1179-1186. doi:10.1007/s00127-009-0152-z

Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology, 45*, 163-191.

Lien, L., Green, K., Welander-Vatn, A., & Bjertness, E. (2009). Mental and somatic health complaints associated with school bullying between 10th and 12th grade students; results from cross sectional studies in Oslo, Norway. *Clinical Practice and Epidemiology in Mental Health, 5*, 1-8.

Liu, S.-K., Chien, Y.-L., Shang, C.-Y., Lin, C.-H., Liu, Y.-C., & Gau, S. S.-F. (2013). Psychometric Properties of the Chinese version of Strength and Difficulties Questionnaire. *Comprehensive Psychiatry, 54*, 720-730.

Mathai, J., Anderson, P., & Bourne, A. (2004). Comparing psychiatric diagnoses generated by the Strengths and Difficulties Questionnaire with diagnoses made by clinicians. Australian and New Zealand. *Journal of Psychiatry, 38*, 639-643.

Matsuishi, T., Nagano, M., Araki, Y., Tanaka, Y., Iwasaki, M., Yamashita, Y., . . . Kakuma, T. (2008). Scale properties of the Japanese version of the Strengths and Difficulties Questionnaire (SDQ): A study of infant and school children in community sample. *Brain and Development, 30*, 410-415.

Mazur, J., Tabak, I., & Kololo, H. T. (2007). Toward a better assessment of child and adolescent mental health status. Polish version of strengths and difficulties questionnaire. *Developmental Period Medicine, 11*, 13-24.

Mellor, D. (2004). Furthering the Use of the Strengths and Difficulties Questionnaire: Reliability With Younger Child Respondents. *Psychological Assessment, 16*, 396-401.

Mellor, D., & Stokes, M. (2007). The factor structure of the Strengths and Difficulties Questionnaire. *European Journal of Psychological Assessment, 23*, 105-112.

Mieloo, C., Raat, H., van Oort, F., Bevaart, F., Vogel, I., Donker, M., & Jansen, W. (2012). Validity and reliability of the Strengths and Difficulties Questionnaire in 5-6 year olds: Differences by gender or by parental education? *PLoS ONE, 7*(5). doi:10.1371/journal.pone.0036805

Mieloo, C. L., Bevaart, F., Donker, M. C., van Oort, F. V., Raat, H., & Jansen, W. (2014). Validation of the SDQ in a multi-ethnic

population of young children. *The European Journal of Public Health, 24*, 26-32.

Mojtabai, R. (2006). Serious emotional and behavioral problems and mental health contacts in American and British children and adolescents. *Journal of the American Academic of Child and Adolescence Psychiatry, 45*, 1215-1223.

Moscoso, A., Jovanovic. N, & Rojnic, M. (2015). Transition from adolescent to adult mental health services in Europe from the provider's perspective. *Lancet Psychiatry, 2*, 779 780.

Mulloy, M., Evangelista, N., Betkowski, J., & Weist, M. D. (2011). Mental health screening for adolescents. In R. J. R. Levesque (Ed.), *Encyclopedia of Adolescence* (pp. 1707-1714). New York: Springer.

Muris, P., Meesters, C., Eijkelenboom, A., & Vincken, M. (2004). The self-report version of the Strengths and Difficulties Questionnaire: Its psychometric properties in 8- to 13- year-old non-clinical children. *British Journal of Clinical Psychology, 43*, 437-448.

Muris, P., Meesters, C., & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ). Further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child and Adolescent Psychiatry, 12*, 1-8.

Najman, J. M., Heron, M. A., Hayatbakhsh, M. R., Dingle, K., Jamrozik, K., Bor, W., . . . Williams, G. M. (2008). Screening in early childhood for risk of later mental health problems: A longitudinal study. *Journal of Psychiatric Research, 42*, 694-700.

Niclasen, J., Skovgaard, A. M., Andersen, A. M., Somhovd, M. J., & Obel, C. (2013). A confirmatory approach to examining the factor structure of the Strengths and Difficulties Questionnaire. *Journal of Abnormal Child Psychology, 41*, 355-365.

Niclasen, J., Teasdale, T. W., Andersen, A.-M., Skovgaard, A. M., Elberling, H., & Obel, C. (2012). Psychometric properties of the Danish Strengths and Difficulties Questionnaire: The SDQ assessed for more than 70,000 raters in four different cohorts. *PloS ONE, 7*(2).

Olfson, M., Blanco, C., Wang, S., Laje, G., & Correll, C. U. (2014). National Trends in the Mental Health Care of Children, Adolescents, and Adults by Office-Based Physicians. *JAMA Psychiatry, 71*, 81-90.

Ortuño-Sierra, J., Chocarro, E., Fonseca-Pedrero, E., Sastre i Riba, S., & Muñiz, J. (2015a). The assessment of emotional and behavioural problems: Internal structure of The Strengths and Difficulties Questionnaire. *International Journal of Clinical and Health Psychology, 15*, 265-273.

Ortuño-Sierra, J., Fonseca-Pedrero, E., Paino, M., Sastre i Riba, S., & Muñiz, J. (2015b). Screening mental health problems during adolescence: Psychometric properties of the Spanish version of the Strengths and Difficulties Questionnaire. *Journal of Adolescence, 38*, 49-56.

Ortuño-Sierra, J., Fonseca-Pedrero, J., Aritio-Solana, R., Moreno, A., Chocarro de Luis, E., Schumann, G., . . . and the IMAGEN consortium. (2015c). New Evidence of Factor Structure and Measurement Invariance of the SDQ Across Five European Nations. *European Child and Adolescent Psychiatry*. doi:10.1007/s00787-015-0729-x

Ortuño, J., Fonseca-Pedrero, E., Paino, M., & Aritio-Solana, R. (2014). Prevalencia de síntomas emocionales y comportamentales en adolescentes españoles [The prevalence of emotional and behavioural symptoms in Spanish adolescents]. *Revista de Psychiatría y Salud Mental 7*, 121-130.

Palfrey, J., Tonniges, T., Green, M., & Richmond, J. (2005). Addressing the millennial morbidity-the context of community pediatrics. *Pediatrics, 115*, 1121-1123.

Palmieri, P. A., & Smith, G. C. (2007). Examining the structural validity of the Strengths and Difficulties Questionnaire (SDQ) in a U.S. sample of custodial grandmothers. *Psychological Assessment, 19*, 189-198.

Percy, A., McCrystal, P., & Higgins, K. (2008). Confirmatory factor analysis of the adolescent self-report Strengths and Difficulties Questionnaire. *European Journal of Psychological Assessment, 24*, 43-48.

Petermann, U., Petermann, F., & Schreyer, I. (2010). The German Strengths and Difficulties Questionnaire (SDQ) validity of the teacher version for preschoolers. *European Journal of Psychological Assessment, 26*, 256-262.

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry, 56*, 345-365.

Prior, M., Virasinghe, S., & Smart, D. (2005). Behavioural problems in Sri Lankan schoolchildren. Associations with socio-economic status, age, gender, academic progress, ethnicity and religion. *Social Psychiatry and Psychiatric Epidemiology, 42*, 654-662.

Ravens-Sieberer, U., Wille, N., Erhart, M., Bettge, S., Wittchen, H.-U., Rothenberger, A., . . . Dopfner, M. (2008). Prevalence of mental health problems among children and adolescents in Germany: results of the BELLA study within the National Health Interview and Examination Survey. *European Child and Adolescent Psychiatry, 17*, 22-33.

Reinholdt-Dunne, M. L., Esbjorn, B. H., Hoyer, M., Dahl, S., Boesen, A. J., Hansen, S. G., & Leth, I. (2011). Emotional difficulties in seventh grade children in Denmark. *Developmental and Aging, 52*, 433-439.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behaviour Assesment System for Children (BASC)*. Circle Pines, MN: American Guidance Service.

Richter, J., Sagatun, A., Heyerdahl, S., Oppedal, B., & Røysamb, E. (2011). The Strengths and Difficulties Questionnaire (SDQ) – Self-Report. An analysis of its structure in a multiethnic urban adolescent sample. *Journal Child and Psychology Psychiatry, 52*(9), 1002-1011.

Rodríguez-Hernández, P. J., Betancor, M., Ramírez-Santana, G. M., García, R., Sanz-Álvarez, E. J., & De las Cuevas-Castresana, C. (2012). Psychometric properties of the parent and teacher versions of the Strength and Difficulties Questionnaire (SDQ) in a Spanish sample. *International of Clinical and Health Psychology, 12*(2), 265-279.

Rønning, J. A., Helge Handegaard, B. H., Sourander, A., & Mørch, W.-T. (2004). The Strengths and Difficulties Self-Report Questionnaire as a screening instrument in Norwegian community samples. *European Child and Adolescent Psychiatry, 13*, 73-82.

Rothenberg, A., Becker, A., Erhart, M., Wille, N., & Ravens-Sieberer, U. (2008). Psychometric properties of the parent strengths and difficulties questionnaire in the general population of German children and adolescents: results of the BELLA study. *European Child and Adolescent Psychiatry, 17*, 99-105.

Ruchkin, V., Jones, S., Vermeiren, R., & Schwab-Stone, M. (2008). The Strengths and Difficulties Questionnaire: The self-report version in American urban and suburban youth. *Psychological Assessment, 20*, 175-182.

Ruchkin, V., Koposov, R., & Schwab-Stone, M. (2007). The strength and difficulties questionnaire: Scale validation with Russian adolescents. *Journal of Clinical Psychology, 63*, 861-869.

Ruchkin, V., Koposov, R. A., Vermeiren, R., & Schwab-Stone, M. (2012). The Strengths and Difficulties Questionnaire: Russian validation of the teacher version and comparison of teacher and student reports. *Journal of Adolescence, 35*, 87–96.

Rutter, M., & Graham, P. (1966). Psychiatric disorder in 10- and 11-year-old children. *Proceedings of the Royal Society of Medicine, 69*, 382-387.

Sanne, B., Torsheim, T., Heiervang, E., & Stormark, K. M. (2009). The Strengths and Difficulties Questionnaire in the Bergen Child Study: A Conceptually and Methodically Motivated Structural Analysis. *Psychological Assessment, 21*, 352-364.

Shevlin, M., Murphy, S., & McElearney, A. (2012). Confirmatory factor analysis of adolescent self- and informant ratings of the Strengths and Difficulties Questionnaire. *The Irish Journal of Psychology, 33*, 17-28.

Silva, T. B., Osorio, F. L., & Loureiro, S. R. (2015). SDQ: discriminative validity and diagnostic potential. *Frontiers in Psychology, 6, 811*. doi:10.3389/fpsyg.2015.00811

Simpson, G. A., Bloom, B., Cohen, R. A., Blumberg, S., & Bourdon, K. H. (2005). U.S. Children with emotional and behavioral difficulties: Data from the 2001, 2002, and 2003 National Health Interview Surveys. *Advance Data, 360*, 1-13.

Steinberg, L., & Morris, A. S. (2001). Adolescent development. *Annual Review of Psychology, 52*, 83-110.

Stevanovic, D., Urbán, R., Atilola, O., Vostanis, P., Singh Balhara, Y. P., M. Avicenna, M., . . . Petrov, P. (2014). Does the Strengths and Diffculties Questionnaire – self report yield invariant measurements across different nations? Data from the International Child Mental Health Study Group. *Epidemiology and Psychiatric Sciences, 30*, 1-12.

Stockings, E. A., Degenhardt, L., Dobbins, T., Lee, Y. Y., Erskine, H. E., Whiteford, H. A., & Patton, G. (2015). Preventing depression and anxiety in young people: a review of the joint efficacy. *Psychological Medicine, 28*, 1-16.

Stone, L. L., Otten, R., Engels, R. C. M., Vermults, A. A., & Janssens, J. M. (2010). Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: a review. *Clinical Child and Family Psychology Review, 13*, 254-274.

Svedin, C. G., & Priebe, G. (2008). The Strengths and Difficulties Questionnaire as a screening instrument in a community sample of high school seniors in Sweden. *Nordic Journal of Psychiatry, 62*, 225-232.

Sveen, T. H., Berg-Nielsen, T. S., Lydersen, S., & Wichstrøm, L. (2013). Detecting psychiatric disorders in preschoolers: Screening with the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*, 728-736.

Thabet, A. A., Stretch, D., & Vostanis, P. (2000). Child mental health problems in Arab children: application of the Strengths and Difficulties Questionnaire. *International Journal of Social Psychiatry, 46*, 266-280.

Thabet, A. Z. M., Karim, K., & Vostanis, P. (2006). Trauma exposure in pre-school children in a war zone. *British Journal of Psychiatry, 188*, 154-158.

Theunissen, M. H. C., Vogels, A. G. C., De Wolff, M. S., & Reijneveld, S. A. (2013). Characteristics of the strengths and difficulties questionnaire in preschool children. *Pediatrics, 131*(2), e446-e454.

Ullah-Syed, E. U., Abdul-Hussein, S. A., & Haidry, S.-e.-Z. (2009). Prevalence of Emotional and Behavioural Problems Among Primary School Children in Karachi, Pakistan – Multi Informant Survey. *Indian Journal of Pediatrics, 76*, 623-628.

van de Looij-Jansen, P. M., Goedhart, A. W., De Wilde, E. J., & Treffers, P. D. (2011). Confirmatory factor analysis and factorial invariance analysis of the adolescent self-report Strengths and Difficulties Questionnaire: How important are method effects and minor factors? *British Journal of Clinical Psychology, 50*, 127-144.

Van Leeuwen, K., Meerschaert, T., Bosmans, G., De Medts, L., & Braet, C. (2006). The Strengths and Difficulties Questionnaire in a Community Sample of Young Children in Flanders. *European Journal of Psychological Assessment, 22*, 189-197.

Van Roy, B., Grøholt, B., Heyerdahl, S., & Clench-Aas, J. (2006). Self-reported strengths and difficulties in a large Norwegian population 10-19 years: age and gender specific results of the extended SDQ-questionnaire. *European Child and Adolescent Psychiatry, 15*, 189-198.

Van Roy, B., Grøholt, B., Heyerdahl, S., & Clench-Aas, J. (2010). Understanding discrepancies in parent-child reporting of emotional and behavioural problem: Effects of relational and socio-demographic factors. *BMC Psychiatry, 10*(56). doi:10.1186/1471-244X-10-56

Van Roy, B., Veenstra, M., & Clench-Aas, J. (2008). Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-, early, and late adolescence. *Journal of Child Psychology and Psychiatry, 49*, 1304-1312.

van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child and Adolescent Psychiatry, 12*, 281-289.

Vostanis, P. (2006). Strengths and Difficulties Questionnaire: research and clinical applications. *Current Opinion in Psychiatry, 19*, 367-372.

Welham, J., Scott, J., Williams, G. M., Najman, J. M., Bor, W., O'Callaghan, M. J., & McGrath, J. (2009). Emotional and behavioural antecedents of young adults who screen positive for non-affective psychosis: a 21 year birth cohort study. *Psychological Medicine, 39*, 625-634.

Wichstrøm, L., Berg-Nielsen, T. S., Angold, A., Egger, H. L., Solheim, E., & Sveen, T. H. (2012). Prevalence of psychiatric disorders in preschoolers. *Child Psychology and Psychiatry, 53*, 695-705.

Widiger, T. A., De Clercq, B., & De Fruyt, F. (2009). Childhood antecedents of personality disorder: An alternative perspective. *Development and Psychopathology, 21*, 771-791.

Williamson, A., McElduff, P., Dadds, M., D'Este, C., Redman, S., Raphael, B., . . . Eades, S. (2014). The construct validity of the strengths and difficulties questionnaire for Aboriginal children living in urban New South Wales, Australia. *Australian Psychologist, 49*, 163-170.

Yao, S., Zhang, C., Zhu, X., Jing, X., McWhinnie, C. M., & Abela, J. R. Z. (2009). M Measuring adolescent psychopathology: psychometric properties of the self-report strengths and difficulties questionnaire in a sample of Chinese adolescents. *Journal of Adolescent Health, 45*, 55-62.

Papeles del Psicólogo, 2016. Vol. 37(1), pp. 27-35
http://www.papelesdelpsicologo.es

THE ASSESSMENT OF CYBERBULLYING: THE PRESENT SITUATION AND FUTURE CHALLENGE

Beatriz Lucas-Molina[1], Alicia Pérez-Albéniz[2] and Marta Giménez-Dasí[1]
[1]*Universidad de Valencia.* [2]*Universidad de La Rioja*

En la última década se ha asistido a un notable incremento del interés de la comunidad educativa y científica por el cyberbullying, una nueva forma de maltrato e intimidación entre iguales. A pesar de la amplia proliferación de estudios y de instrumentos de evaluación sobre el fenómeno, siguen existiendo importantes lagunas conceptuales y metodológicas. Este trabajo ofrece una revisión general y actualizada de los resultados de la investigación sobre la definición del constructo, su prevalencia y su impacto en las personas implicadas. Finalmente, se centra de manera específica en la evaluación del constructo y proporciona una breve revisión de las características generales y psicométricas de aquellos instrumentos utilizados en algunos de los estudios nacionales e internacionales más relevantes realizados sobre el tema. El trabajo hace especial hincapié en los retos presentes y futuros y finaliza con algunas recomendaciones generales que pretenden guiar en la selección y/o construcción adecuada de instrumentos de evaluación en este campo de estudio.
Palabras clave: *Cyberbullying, Bullying tradicional, Definición, Evaluación, Instrumento.*

In the last decade there has been a significant increase in the interest of the educational and scientific community on cyberbullying, a new form of peer abuse and intimidation. Despite the widespread proliferation of studies and assessment tools on the phenomenon, there are still major conceptual and methodological gaps. This paper offers a comprehensive and updated review of the results of research on the definition of the construct, its prevalence and its impact on the people involved. Finally, it focuses specifically on the assessment of the construct and provides a brief review of the general and psychometric characteristics of the instruments used in some of the most relevant national and international studies conducted on the subject. This work places special emphasis on the present and future challenges and concludes with a number of general recommendations intended to guide the selection and/or construction of assessment instruments in this field of study.
Key words: *Cyberbullying, Traditional bullying, Definition, Measurement, Instrument.*

Since the first study on cyberbullying in 2000 by Finkelhor, Mitchell, and Wolak in the United States, there have been numerous investigations conducted on the phenomenon both outside and inside our country (e.g., Álvarez-García et al., 2011; Beran & Li, 2007; Buelga, Calva & Musitu, 2010; Calvete, Orue, Estévez, Villardón & Padilla, 2010; Hinduja & Patchin, 2008; Ortega, Calmaestra & Mora-Merchán, 2008; Williams & Guerra, 2007; Ybarra & Mitchell, 2008). Proof of this is in the many special issues that various journals, both national and international, have devoted to the subject (e.g., *Journal of Adolescent Health, Journal of Community and Applied Psychology, Psicothema*).

However, despite this widespread proliferation of studies focused mainly on understanding the prevalence of cyberbullying and its correlates with other psychosocial variables, there is still no universally agreed definition (Álvarez-García et al., 2011; Stewart, Drescher, Maack, Ebesutani & Young, 2014; Ybarra, Boyd, Korchmaros & Oppenheim, 2012). This has meant that different methodologies have been used in the evaluation of the construct, which has hampered both the comparison of the results obtained by different studies and the advancement of research in the area (Hanewald, 2013; Ybarra et al., 2012).

This paper aims to present a synthesis of the state of the question regarding the assessment of cyberbullying. To this end, firstly the construct is defined, followed by a presentation of the prevalence rates and the impact on the development of the people involved. In the second stage, focusing more specifically on assessment, this study addresses some of the challenges faced when assessing this construct today and presents some of the most important instruments nationally and internationally. Finally, a number of guidelines and recommendations are offered, which should guide decisions when choosing an existing tool to assess cyberbullying or, alternatively, designing one's own tool, and also when embarking on future research in this field of study.

Correspondence: Beatriz Lucas-Molina. *Departamento de Psicología Evolutiva y de la Educación. Universidad de Valencia. Avenida Blasco Ibáñez 21. 46010 Valencia. España.*
E-mail: beatriz.lucas@uv.es

CONCEPTUAL DEFINITION, PREVALENCE AND IMPACT ON DEVELOPMENT

Cyberbullying has been defined as the kind of harassment committed by an individual or group who, using new information and communications technology (ICT) (mobile phones, email, social networks, blogs, websites, etc.), deliberately and repeatedly attacks someone who cannot easily defend him or herself (Patchin & Hinduja, 2006; Smith et al, 2008). This new form of peer abuse has received other denominations, such as *online bullying* (Nansel et al., 2001), *electronic bullying* (Kowalski & Limber, 2007; Raskauskas & Stoltz, 2007), *online harassment* (Filkelhor et al, 2000), *internet bullying* (Williams & Guerra, 2007), and *cyber aggression* (Pornari & Wood, 2010). This diversity of names illustrates the existing terminological and conceptual confusion in this area of study, which sometimes leads to different terms being used for the same concept or the same term being used with different meanings (Tokunaga, 2010; Ybarra et al., 2012). This paper will use the above definition proposed by Smith et al. (2008) and the term *cyberbullying*, which is the most widely used in the scientific literature.

According to this definition, cyberbullying shares the three characteristics of traditional bullying as it deals with aggressive behaviours that are intentional, repeated and based on an asymmetrical

relationship of power and control over/submission to another person (Kowalski & Limber, 2007; Olweus, 1993). However, some authors have questioned the feasibility of evaluating these three aspects in cyberspace (Menesini & Nocentini, 2009). For example, while it is easy to calibrate the imbalance of power in traditional bullying, either due to the greater physical or psychological strength of the attacker or due to a purely numerical criterion (several aggressors compared with one victim), it is more complicated in cyberspace. Some authors suggest that the use of this imbalance of power criterion could be justified by the greater reach of the attacks that occur using the new technologies, as they transcend to a larger virtual audience compared with traditional bullying, which reaches a much smaller group (Garaigordobil, 2011; Williamson, Lucas-Molina & Guerra, 2013). In other words, the imbalance of power would be determined by the public nature of cyberbullying compared with the private nature of traditional bullying (Thomas et al., 2015). Moreover, the need has also been questioned for the cyberbullying to be of a repetitive nature to be considered as such (Gairagordobil & Martinez-Valderrey, 2015). A single act, such as the publication of a compromising photo (whether real or the result of a montage) in a social network can result in immediate dissemination and thus meet the criterion of being repetitive and frequent (Menesini & Nocentini, 2009).

Despite these similarities, which have led some authors to argue that cyberbullying is a traditional form of bullying (such as physical or relational bullying; Li, 2007), cyberbullying differs from traditional bullying in a number of aspects (Álvarez-García et al., 2011; Buelga et al., 2010; Gairagordobil, 2011; Stewart et al., 2014). The first has already been mentioned before: the greater scope of cyberbullying. With a single click, a student can spread a false rumour to hundreds and thousands of people on the internet, whereas in traditional bullying, because it occurs in person, the scope of the rumour is much more restricted. The second difference is the victim's inability to escape the situation of intimidation. Traditional bullying is essentially limited to the time that the student victim spends in the school environment and its surroundings; in cyberbullying, however, the harassment can continue 24 hours a day, 7 days a week, whether or not the student is at school, as he or she can still receive messages via mobile or computer. The third difference is that, unlike traditional bullying, cyberbullying is not a "face to face" experience; the aggressor does not have to expose him or herself physically to the victim. Besides this, pseudonyms can be used on the internet. All of this gives a certain invisibility and allows the cyberbully to act anonymously. Finally, the contents of electronic bullying can be permanent or difficult to remove, so victims may relive the situation of victimisation over and over again, placing them in a more vulnerable situation (Buelga et al., 2010).

In line with this, several authors propose that the public nature (reaching a large audience) and the anonymous nature (the fact that the aggressor is not known) should be included in the definition of cyberbullying, relegating to the background both the repetitive nature and the imbalance of power (Nocentini et al, 2010; Thomas et al, 2015). However, it is worth noting that anonymity does not occur in all situations of cyberbullying (Tokunaga, 2010). Regardless of the place the criteria occupy in the definition, what the studies focused on how adolescents perceive cyberbullying scenarios do disclose (Menesini et al., 2012) is the need to include specific criteria on cyberbullying that go beyond the intentionality, repetition and imbalance of power.

This new technological form of bullying includes a wide range of behaviours that are usually classified into the following categories (Gairagordobil, 2014; Kowalski, Limber & Agatson, 2010; Willard, 2007): social exclusion (not letting the victim participate in a specific social network), denigration (spreading rumours and false information about the victim), harassment (sending and disseminating offensive messages), impersonation (sending malicious messages in forums or chat rooms posing as the victim), violation of privacy (disseminating secrets or images of the victim); persecution (sending threatening messages) and "happy slapping" (physically assaulting the victim in order to record and disseminate the aggression within their environment). Moreover, these forms can vary, and indeed they do, with the rapid development of ICT, as well as between different cultures (Menesini, Nocentini & Calussi, 2011).

The inconsistency in the conceptualisation and consequently the operationalisation of the construct have led to the use of different assessment methodologies and to the obtaining of different degrees of prevalence. In this regard, it should be noted that studies conducted outside our country have found prevalence rates ranging from 9% (Ybarra, Mitchell, Wolak & Finkelhor, 2006) to 72% (Juvonen & Gross, 2008). In Spain, the studies that have been carried out have also produced varied results (Álvarez-García et al., 2011). For example, the study conducted nationally by the Observatorio Estatal de la Convivencia Escolar [National Observatory of School Life] on 23,100 secondary school students, between 2.5% and 7% of the students admitted to having been a victim and 2.5-3.5% admitted to having been an aggressor of any of the four types of cyberbullying in the last two months (Díaz-Aguado, Martínez-Arias & Martín, 2013). These results are far from those obtained by Buelga et al. (2010), in a sample of 2,101 students aged between 11 and 17 years from Valencia, according to which 24.6% had been bullied by mobile and 29% by internet in the last year. These percentages are consistent with the review by Tokunaga (2010), according to which between 20% and 40% of teens experience cyberbullying. Moreover, several studies have shown the rapid increase in this new form of bullying among adolescents. Wolak, Mitchell and Finkelhor (2006) found that the prevalence rates had doubled five years after their first study on cyberbullying (Finkelhor et al., 2000). Also, it is worth noting that these prevalence rates are higher than those found in traditional bullying (Nansel et al., 2001).

As was the case with traditional bullying, the two variables most analysed in detecting the students involved in this technological form of bullying have been age or educational level, and gender. With regard to the first variable, the different studies within and outside Spain on the prevalence of cyberbullying seem to point to the same pattern detected in face-to-face bullying: a rise in pre-adolescence or during the first years of secondary education, with a subsequent decline in the final years of this educational stage (Buelga et al, 2010; Williams & Guerra, 2007). However some studies have not found differences regarding the age of the pupils (Smith et al., 2006). Regarding gender differences, the literature agrees that girls are bullied more than boys (Burgess-Proctor et al, 2009; Calvete et al, 2010; Félix-Mateo et al, 2010; Kowalski & Limber, 2007; Li, 2007; Ortega, Elipe, Mora-Merchán, Calmaestra & Vega, 2009; Smith et al., 2006; Stewart et al, 2014). For example, a recent study in the Basque Country on a sample of 3,026 adolescents aged between 12 and 18, Garaigordobil and Aliri (2013) found a significantly higher percentage of female victims (17.6% girls, 12.5% boys). However, other investigations have not found these differences (Álvarez-García et al, 2011; Buelga et al, 2010; Hinduja & Patchin,

2008; Juvonen & Gross, 2008; Williams & Guerra, 2007). Interestingly, these results are different from those found in traditional bullying, where there was a higher percentage of boys both in the role of victim and that of aggressor (Tokunaga, 2010).

Cyberbullying has harmful consequences for everyone involved, regardless of the role played, as they are at greater risk for psychosocial maladjustment and psychopathological disorders in adolescence and adulthood (Gairagordobil, 2011; Gradinger, Strohmeier & Spiel, 2009). In general, the research shows that cyberbullying has similar effects to traditional bullying both at the time that it happens, and in the medium and long term (Kowalski et al., 2010). However, some authors suggest that its effects can be more devastating, especially among student victims. Smith et al. (2006) found that victims perceived electronic forms of bullying as more serious than traditional forms, especially when the harassment was of a more public and menacing nature.

As a result of cyberbullying, victims often have feelings of anxiety, depression, helplessness, sadness, low self-esteem and self-confidence as well as a poor psychosocial adjustment (Kowalski et al, 2010; Ybarra & Mitchell, 2004). They also display poor academic performance, poor concentration and truancy (Beran & Li, 2007; Raskauskas & Stolz, 2007) and show higher levels of stress, fear and suicidal ideation (Hinduja & Patchin, 2010). Therefore, as with traditional bullying, cyberbullying has significant effects on victims at the emotional, psychosocial and academic levels (Kowalski et al, 2010; Tokunaga, 2010).

The aggressors are more likely to present moral disengagement and a lack of empathy (Ortega, Sánchez & Menesini, 2002) and they often exhibit problems in compliance with rules as well as aggressive behaviour (Ybarra & Mitchell, 2007). Furthermore, they are also at a higher risk of using drugs and displaying criminal conduct, social isolation and dependence on technologies (Ybarra, Diener-West & Finkelhor, 2007).

ASSESSMENT OF CYBERBULLYING

The results presented earlier report the existence and current importance of the problem. They also underline the urgent need for further research on the topic and for developing valid and reliable assessment measures that enable not only the comparison of results between studies, but especially the correct identification of this form of harassment in order to provide the appropriate prevention and intervention (Dredge, Gleeson & de la Piedad, 2013; Tokunaga, 2010).

In this sense, although at first the objective of this area of research was to determine the presence of the phenomenon and its impact on the personal, social and academic development of adolescents, in recent years efforts have focused on creating new tools for its assessment, as well as studying the psychometric properties of the existing tools (Berne et al, 2013; Dredge et al, 2013; Menesini et al, 2011; Tokunaga, 2010).

The aim of this second section is, firstly, to identify the main difficulties in assessing cyberbullying today, some of which have already been glimpsed in the previous section and are mostly inherited from the study of traditional bullying. Secondly, the instruments used in some of the most important studies carried out on cyberbullying within and outside our borders will be presented, although there will be special emphasis on the national ones, noting their general characteristics and psychometric properties. For a closer look at the international instruments, the reader is recommended to refer to the recent review by Berne et al. (2013) in which the characteristics of 44 cyberbullying assessment tools were analysed exhaustively.

As discussed above, regarding the difficulties faced in the assessment of cyberbullying, a distinction can be made between those that are characteristic of the construct and others that were already present in the study of traditional bullying. Among the former, the most significant is the aforementioned lack of consensus regarding the definition of the cyberbullying construct. This lack of conceptual definition is, according to some authors (Tokunaga, 2010), the most widespread methodological problem in the investigation of cyberbullying.

Another difficulty associated with the cyberbullying construct itself is the enormous variety of behaviours it includes, which have been categorised into different classifications (Gairagordobil, 2011; Kowalski et al, 2010; Willard, 2007). These behaviours and classifications are changing along with the rapid development of ICT (Menesini et al., 2011), leading to the quick obsolescence of the existing categorisations and the continuous inclusion and exclusion of new forms of electronic aggression.

As well as the above, the assessment of cyberbullying also has to deal with a number of problems that were already present in the study of traditional bullying. Thus, even using the same definition of cyberbullying, there are instruments that choose to include it explicitly in the presentation of the questionnaire while others do not. Furthermore, even when including the same definition, two instruments can operationalise the construct in very different ways. For example, some instruments pose one single question after the definition of the frequency with which the respondent has perpetrated or suffered the phenomenon, while others have a list of behavioural descriptors in which different forms of cyberbullying appear.

One of the great debates in the assessment of traditional bullying and cyberbullying, has been to clarify the appropriateness of whether to use a single general question after the definition (e.g., "Have you suffered from / carried out this kind of bullying?") or only to include a list of the different experiences of cyberbullying without a previous definition. The studies show the pros and cons of the two types of formats. Among the advantages of the former option is its practical application as it is based on a single item (Solberg & Olweus, 2003). The disadvantages include, firstly, that the definition may be interpreted differently depending on the student's age or culture (Ybarra et al., 2012), and secondly, that the student's response may be influenced by social desirability, since students are very likely to be reluctant to label themselves as victims or perpetrators of cyberbullying (Menesini et al., 2009). As for the model based on behavioural descriptors, its strengths include that it provides a more reliable, valid and accurate measure compared with the estimation provided by a single item (Menesini et al., 2011). Moreover, this range of items can more accurately represent the complexity of the construct. Its limitations include that it cannot cover all situations of cyberbullying and that this format may result in higher prevalence rates, because students may be considering as cyberbullying acts that in fact are not (Gradinger et al., 2009; Ybarra et al., 2012). In this sense, the studies that use both measurement strategies have highlighted the inconsistency between the responses to the global item and the individual descriptors related to participation in cyberbullying situations (Burgess-Proctor et al., 2009), with the affirmative percentages for isolated behaviours being higher than those for the global item. Regardless of whether or not the definition or the behavioural descriptors are included, studies show

that the formats that lead to fewer errors in the classification of student victims and bullies are those that incorporate the specific criteria of bullying and cyberbullying (intentional, repetitive and power imbalance) (Menesini et al, 2010; Ybarra et al, 2012). Moreover, in the case of cyberbullying it would be interesting to include more specific criteria (public, anonymous nature, etc.) (Menesini et al, 2011; Tokunaga, 2010). It is therefore important to keep this in mind when choosing or designing the assessment instrument.

The variations in the format do not end here, we also find questionnaires which, while based on uniform definitions and behavioural listings, employ different time intervals. Thus, some studies ask to what extent the respondent has perpetrated or suffered such incidents since the start of the school year, others in the past year, others in the last two or three months, and some do not establish any kind of time limit. The combination of these variants results in a multitude of assessment tools that can lead to very different prevalence rates as we have already seen (e.g., between 9% and 72% in the US and between 2.5% and 24.6% in Spain).

Another difficulty in assessing cyberbullying is estimating the discriminative power of the items in distinguishing different levels of severity in cyberbullying, because it is not the same to make an offensive comment via a text message as it is to publish a compromising photograph in a social network. To this end, Menesini et al. (2012), using item response theory (IRT), found that the visual forms of cyberbullying (photographs and videos) were the most serious. However, they found some differences with respect to the previous studies (e.g., Smith et al., 2008), especially in less serious items, concluding that it is important to take cultural differences into account in both the conceptualisation of cyberbullying and the use of new technologies.

Tables 1 and 2 show the most relevant instruments in Spain, and in the European and US contexts, respectively, for evaluating cyberbullying.

In Table 1 we can observe how only one of the six national questionnaires includes the definition of cyberbullying (Ortega et al., 2008), the remaining five incorporate a number of items related to various experiences of cyberbullying (Álvarez-García et al., 2011; Buelga et al., 2010; Calvete et al., 2010; Díaz-Aguado et al., 2013; Gairagordobil & Aliri, 2013), two of which differ between the roles of victim and perpetrator (Díaz-Aguado et al., 2013; Gairagordobil & Aliri, 2013), and one differentiates the electronic medium used (Buelga et al, 2010). Also, while all the instruments include the electronic and intentional dimension of the behaviour evaluated, only two studies incorporate the repetitive nature (Buelga et al, 2010; Ortega et al., 2008). None of them considered the imbalance of power, or other characteristic criteria of cyberbullying (e.g., the public or anonymous

TABLE 1
NATIONAL TOOLS FOR ASSESSING CYBERBULLYING: CONCEPTUAL AND PSYCHOMETRIC CHARACTERISTICS

Authors and year/Region	Instrument	N	Age/Level of education	Subscales (nº items) and how they are obtained	Definition	Forms/Device	Reliability
Álvarez-García et al. (2011)/Asturias	Cuestionario de Violencia Escolar – Revisado [Questionnaire of School Violence - Revised] (CUVE-R)	638	1st-4th year secondary	The questionnaire includes 31 items grouped into 8 factors. One of them: Violence through ICT (6 items) How often does the teacher / student in the class experience the acts [1 = NEVER, 5 = ALWAYS] EFA/CFA	E, I Does not include definition	Harassment, Violation of privacy Mobile/Social Networks	CUVE-R: α=0.924 Does not provide data on the 'Violence through ICT' factor.
Buelga et al. (2010)/ Valencia	Escalas de Victimización (EV) a través del teléfono móvil y a través de Internet [Scales of Victimisation (SV) via Mobile and via Internet] Includes two questions that evaluate the intensity and duration of the harassment	2,101	1st-4th year secondary	SV Mobile (8 items) SV Internet (10 items) Harassment experienced within the last year [1=NEVER, 4= MANY TIMES/ALWAYS]	E, I, R, Does not include definition	Harassment, Persecution, Denigration, Violation of privacy, Social exclusion, Impersonation Mobile/Internet	SV Mobile: α=0.76 SV Internet: α=0.84
Calvete et al. (2010)/Vizcaya	Cuestionario Cyberbullying [Cyberbullying Questionnaire] (CBQ)	1,431	12-17 years of age	CB (16 items) How often has any of the 16 behaviours been carried out [0=NEVER, 2=OFTEN] CFA	E,I Does not include definition	Harassment, Persecution, Denigration, Violation of privacy, Social exclusion, Impersonation, Happy slapping Mobile/Internet	α=0.96

TABLE 1
NATIONAL TOOLS FOR ASSESSING CYBERBULLYING:
CONCEPTUAL AND PSYCHOMETRIC CHARACTERISTICS (Continued)

Authors and year/Region	Instrument	N	Age/Level of education	Subscales (nº items) and how they are obtained	Definition	Forms/Device	Reliability
Díaz-Aguado et al. (2013)/Spain	Acoso con nuevas tecnologías [Bullying with new technologies] (The instrument also evaluates traditional bullying)	23,100	1st-4th year secondary (12-18 years of age)	Victim (4 items) Aggressor (4 items) Frequency with which the respondent has suffered or committed the 4 behaviours within the last two months. [1=NEVER, 5= ALWAYS] EFA	E, I The definition of traditional bullying appears but not that of cyberbullying	Harassment, Persecution, Violation of privacy, Mobile/Internet	Victim: α=0.83 Aggressor: α=0.91
Gairagordobil & Aliri (2013)/Basque Country	Cyberbullying: Screening de acoso entre iguales [Cyberbullying: Screening of peer harassment] (Edited by TEA)	3,026	12-18 years of age	Victim (15 items) Aggressor (15 items) Observer (15 items) Inform of the frequency with which the 15 behaviours have been suffered, perpetrated or seen in the last year [0=NEVER, 3=ALWAYS] EFA	E, I Does not include definition	Harassment, Persecution, Denigration, Violation of privacy, Social exclusion, Impersonation, Happy slapping Mobile//Internet	α=0.91
Ortega et al. (2008)/ Cordoba[1]	Cuestionario Cyberbullying [Cyberbullying questionnaire] (adaptation of the instrument by Smith et al., 2006) Also asks about feelings, coping strategies, etc.	830	1st-4th year secondary	Mobile (2) Internet (2) The global definition appears, which includes several examples and the student has to indicate if s/he has been bullied or has bullied someone in this way via mobile and/or internet in the last two months. [NEVER, ONCE or TWICE, ONCE A WEEK, VARIOUS TIMES A WEEK, OTHER]	E,I,R Includes definition	Harassment, Persecution, Denigration, Violation of privacy, Social exclusion, Impersonation Mobile / Internet	—

Note. The double hyphen (—) is used when no information is given about this in the study. EFA = Exploratory Factor Analysis; CFA = Confirmatory Factor Analysis; CB = Cyberbullying. The following initials represent the defining elements of cyberbullying proposed in the scientific literature (Tokunaga, 2010) and which have been considered in the particular instrument (even if there is no definition included): Electronic medium = E; Intentionality = I; Repetition = R; Imbalance of Power = IP; Anonymity = A; Public/Private = P.

[1] The questionnaire is available from: http://www.uco.es/laecovi/img/recursos/RFUY4MDDVCZWHkm.pdf

nature). All of the instruments reviewed include actions that occurred via mobile and the internet. The types and classifications of behaviours varied in each instrument, although all of the included the form 'harassment' (e.g., insulting or ridiculing with messages or calls), and students had to indicate how often they suffered and/or perpetrated each of the behaviours (generally on a Likert scale of 4 points). Two of the instruments do not impose a time interval (Álvarez-García et al, 2011; Calvete et al., 2010), two others specify 'in the last year' (Buelga et al, 2010; Gairagordobil & Aliri, 2013) and the remaining two 'during the last two months' (Diaz-Aguado et al, 2013; Ortega et al, 2008). From the above, we can see the enormous variability of formats used in the different instruments highlighting what we discussed previously in relation to the difficulties in the conceptualisation and operationalisation of cyberbullying.

As for the psychometric aspects of the instruments reviewed, we can see that, in their construction, exploratory factor analysis (EFA) were conducted in three of the studies (Álvarez-García et al, 2011; Díaz-Aguado et al, 2013; Gairagordobil & Aliri, 2013) and confirmatory factor analysis (CFA) were conducted in two (Álvarez-García et al, 2011; Calvete et al, 2010), in order to validate the construct evaluated. With the exception of one study (Ortega et al., 2008), all of the works provided data of internal consistency as indicators of the reliability of the instruments used. Without undervaluing the efforts of the Spanish researchers to analyse and ensure the psychometric properties of the instruments they have developed (e.g., according to the review by Berne et al., 2013, only 18 of the 44 international instruments analysed reported internal consistency data) it would be interesting in the future to complement these results with other measures of reliability (e.g., test-

retest) and to provide other evidence of validity (e.g., convergent and discriminant validity). Although possibly the first step is, as already mentioned, to agree upon both the definition and the operationalisation of the cyberbullying construct in order to ensure the content validity of the instruments developed.

In Table 2, concerning the international instruments, we can see some similarities with the comments made regarding the national instruments. For more information about the different questionnaires used internationally, again we recommend reading the study by Berne et al. (2013), in which 44 instruments are analysed. Here we have selected the studies that are most cited in the literature and those works not included in Berne's review due to having been published afterwards (e.g., Stewart et al., 2014).

DISCUSSION AND CONCLUSIONS

Based on the above it can be concluded that, despite the large number of investigations carried out in the last decade on cyberbullying, it still seems to be an embryonic field of study. In the future, the experts in the field should work together in order to reach a consensus on the conceptualisation and operationalisation of the phenomenon and to continue to research the validity and reliability of the existing instruments.

Specifically, based on what was stated in the previous sections, a number of future challenges arise in the assessment of cyberbullying and these are presented below:

1. It should be noted that the construction of new assessment tools should be based on the analysis of the advantages and disadvantages of the questionnaires already developed by other researchers to avoid the current situation, in which it is unusual to find the same instrument being used in different studies except those written by the same author (Berne et al, 2013; Tokunaga, 2010).
2. The instruments should be based on a definition of cyberbullying and this should appear explicitly in the instrument along with the defining criteria that are to be evaluated. In the future, it would be interesting for the instruments to include, as well as the three criteria of traditional bullying (intentional, repetitive and imbalance of power), the differentiating criteria of cyberbullying, at least the ones that refer to its anonymous and public nature.
3. The instruments should include different behavioural descriptors covering the current classifications in cyberbullying (e.g., Willard, 2007). If a general item is opted for (e.g., after the definition of cyberbullying, asking the question have you suffered/committed this kind of bullying?), it is important to include these more specific descriptors in order to identify and differentiate the different types of cyberbullying.
4. A specific time interval should also be specified in the instructions or in the drafting of the items (e.g., within the last two months), especially in studies aimed at prevention or intervention. The use of general or ambiguous terms should be avoided (e.g., "ever" without specifying a time period or "in the past year", which could be interpreted in various ways), as they do not provide data on active cases during a given time interval. This is vital in comparing the prevalence rates among different studies.
5. The instruments should have sufficient validity indicators. It is necessary to develop valid instruments to ensure that they are all measuring the same phenomenon. In this sense, studies are required that provide evidence of the validity of the assessment instruments. Since one of the main problems is the conceptual definition of the construct, it would be advisable to assess the content validity of the instruments by groups of experts to assess whether the items represent the content domain (Sireci & Faulkner-Bond, 2014). Also in relation to the construct validity, the internal structure of the instruments should be studied (e.g., exploratory and confirmatory analysis) and validity evidence provided in relation to other tests that measure the

TABLE 2
INTERNATIONAL INSTRUMENTS FOR ASSESSING CYBERBULLYING: CONCEPTUAL AND PSYCHOMETRIC CHARACTERISTICS

Authors and year/Region	Instrument	N	Age/Level of education	Subscales (nº items) and how they are obtained	Definition	Forms/Device	Reliability
Beran & Li (2007)/Canada	Cyber-harassment student survey	432	7th-9th grade (12-15 years of age)	Based on the definition of 'harassment' the pupil has to indicate how often s/he has suffered this situation **(does not include time limit).** [1=NEVER, 5= ALWAYS]	E, I, R, IP Includes definition of 'harassment'	Mobile/Internet/Computer /Voice mail/Video cameras	—
Hinduja & Patchin (2008)/USA	General cyberbullying measure	1,378	10-17 years of age	2 items: If respondent has ever been bullied/bullied others online.	E, I, R Includes definition of 'online bullying'	Social exclusion, Harassment, Persecution Mobile/Internet	—
Menesini et al. (2011)/Italy	Cyberbullying Scale	1,092	11-18 years of age	Victim (10 items) Aggressor (10 items) Frequency with which the behaviours have been suffered/committed **within the last 2 months.** [1=NEVER, 5= ALWAYS] CFA	E, I, R, IP Does not include definition	Harassment, Violation of privacy, Denigration, Persecution, Happy slapping Mobile/Internet	Male victims: α=0.87 Female victims: α=0.72 Male aggressors: α=0.86 Female aggressors: α=0.67

TABLE 2
INTERNATIONAL INSTRUMENTS FOR ASSESSING CYBERBULLYING: CONCEPTUAL AND PSYCHOMETRIC CHARACTERISTICS (Continued)

Authors and year/Region	Instrument	N	Age/Level of education	Subscales (nº items) and how they are obtained	Definition	Forms/Device	Reliability
Ortega et al. (2009)/ Cordoba[2]	DAPNHE Questionnaire[3] European Cyberbullying Research Project (ECRP) (Also evaluates traditional bullying)	1,671	1st-3rd Secondary 1st Bacc.	Mobile (12 items) Internet (12 items) 2 items: How often has s/he suffered/ committed this type of bullying via mobile/internet in the **last 2 months** [1=NEVER, 5=A NUMBER OF TIMES A WEEK OR MORE] The rest of the items (10) are related to feelings, coping strategies, etc.	E, I, R, IP Includes definition	Harassment, Persecution, Denigration, Violation of privacy. Mobile/Internet	—
Smith et al. (2008)/England	Cyberbullying questionnaire (Also evaluates traditional bullying with Olweus Bullying/Victim questionnaire)	(1) 92 (2) 533	11-16 years of age	If s/he has suffered bullying (1 item) or bullied someone (1 item) through 7 different means. They are also asked since when. [1=NEVER, 5=A NUMBER OF TIMES A WEEK]	E, I, R, IP Includes definition	Media: text message, photos or videos, telephone calls, email, chat rooms, instant messaging and websites. Mobile/Internet.	—
Stewart et al. (2014) /USA	Cyberbullying Scale	736	6th-12th grade (11-18 years of age)	Whether s/he has suffered bullying or has bullied someone through 8 means (2 items, the pupil must indicate the means). Victim (14 items) Frequency with which s/he has suffered any of the 14 behaviours in the past few months. [1=NEVER, 5= A NUMBER OF TIMES A WEEK] EFA/CFA	E, I, Does not include definition	Media: email, text/Twitter messages, images, instant messaging, online videos, social networks, chat rooms, virtual world (The Sims). Harassment, Persecution, Denigration, Social exclusion, Impersonation, Violation of privacy. Mobile/Internet	Victim (14 items): Boys α=0.94 Girls α=0.93
Ybarra et al. (2006)/ USA	Internet Harassment/Youth Internet Safety Survey	1,501	10-17 years of age	Victim (2 items) Aggressor (2 items) Indicate whether s/he has suffered/ committed any of the 2 behaviours within the last year.	E,I Does not include definition	Harassment, Denigration Internet	—
Ybarra & Mitchell (2008)/USA	Growing up with media (GuwM): youth- reported internet harassment	1,588	10-15 years of age	Victim (3 items) Aggressor (3 items) Frequency with which they have suffered/perpetrated the 3 behaviours **in the past year** [1=NEVER, 5= ALWAYS] CFA	E, I Does not include definition	Harassment, Denigration, Persecution Internet	Victim: α=0.79 Aggressor: α=0.82
Williams & Guerra (2007)/USA		3,339	5th-8th grade (10-14 years of age)	1 item ("I tell lies about my classmates via email or text messages")	E, I Does not include definition	Denigration Mobile/Internet	

Note. The double hyphen (—) is used when no information is given about this in the study. EFA = Exploratory Factor Analysis; CFA = Confirmatory Factor Analysis. The following initials represent the defining elements of cyberbullying proposed in the scientific literature (Tokunaga, 2010) and which have been considered in the particular instrument (even if there is no definition included): Electronic medium = E; Intentionality = I; Repetition = R; Imbalance of Power = IP; Anonymity = A; Public/Private = P.

[2] Although this study was carried out by Spanish authors and in the Spanish population, it has been considered international because the instrument was developed within a European project.

[3]The instrument is available in its English version from: http://www.bullyingandcyber.net/media/cms_page_media/44/Questionario%20EQCB%20english_4.pd

same construct or a different one (AERA, APA & NMCE, 2014). In this line and considering that one of the objectives is diagnosis, it would also be advisable to have an external criterion to serve as a gold standard in assessing, for example, the sensitivity and specificity of the instrument.

6. The instruments should have reliability indicators. Among the instruments presented, it has been observed that few of them provide data on reliability and the ones that do so, refer only to the internal consistency. In the future it would be interesting to conduct longitudinal studies that would allow us to obtain information on the test-retest reliability of the instruments.
7. In connection with the above, we would propose the use of the information function (IF) models from item response theory (Muniz, 1997) as an alternative to Cronbach's alpha. It is of particular interest in this context, since the IF would enable us to know the degree of precision with which the instrument is measuring people with high scores on cyberbullying.
8. If we consider that the participants in these types of situation tend to hide this fact, it would be interesting to highlight the importance of using proxies in addition to assessment using self-reports (Benítez, Padilla & Ongena, 2012). It could be very useful to complement the assessment with information obtained from the parents, friends and teachers.
9. Finally, it would be necessary in future studies to incorporate a cultural perspective that would allow us to make cross-cultural comparisons, as suggested by Menesini et al (2012).

REFERENCES

Álvarez-García, D., Núñez, J. C., Álvarez, L., Dobarro, A., Rodríguez, C. & González Castro, P. (2011). Violencia a través de las tecnologías de la información y la comunicación en estudiantes de secundaria [Violence through information and communications technology in secondary education]. *Anales de Psicología,* 27, 221-230.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Benítez, I., Padilla, J.L. & Ongena, Y. (2012). Evaluation of the convergence between "self-reporters" and "proxies" in a disability questionnaire by means of behaviour coding method. *Quality and Quantity 46*(4), 1311-1322. doi: 10.1007/s11135-011-9443-z

Beran, T. & Li, Q. (2007). The relationship between cyberbullying and school bullying. *Journal of Student Wellbeing, 1*(2), 15-33.

Berne, S., Frisén, A., Schultze-Krumbholz, A., Scheithauer, H., Naruskov, K., Luik, P., et al. (2013). Cyberbullying assessment instruments: A systematic review. *Aggression and Violent Behavior, 18(2),* 320–334. doi:10.1016/j.avb.2012.11.022

Buelga, S., Cava, M.J. & Musitu, G. (2010). Cyberbullying: victimización entre adolescentes a través del teléfono móvil y de Internet [Cyberbullying: victimisation among adolescents via the mobile telephone and internet]. *Psicothema, 22,* 784-789.

Burgess-Proctor, A., Patchin, J. W. & Hinduja, S. (2009). Cyberbullying and online harassment: Reconceptualizing the victimization of adolescent girls. In V. García & J. Clifford (Eds.), *Female crime victims: Reality reconsidered* (pp. 153-175). Upper Saddle River, NJ: Prentice Hall.

Calvete, E., Orue, I., Estévez, A., Villardón, L. & Padilla, P. (2010). Cyberbullying in adolescents: Modalities and aggressors' profile. *Computers in Human Behavior, 26*(5), 1128-1135. doi:10.1016/j.chb.2010.03.017

Díaz-Aguado, M. J., Martínez-Arias, R. & Martín, J. (2013). El acoso entre adolescentes en España. Prevalencia, papeles adoptados por todo el grupo y características a las que atribuyen la victimización [Bullying among adolescents in Spain. Prevalence, roles adopted by the entire group and characteristics attributed to victimisation]. *Revista de Educación, 362*, 348-379. doi: 10-4438/1988-592X-RE-2011-362-164

Dredge, R., Gleeson, J. & de la Piedad García, X. (2013). Progress and limitations in the measurement of cyber bullying. In R. Hanewald (Ed.), *From cyber bullying to cyber safety. Issues and approaches in educational contexts* (pp. 271-288). New York: Nova Science Publishers.

Félix-Mateo, V., Soriano-Ferrer, M., Godoy-Mesas, C. & Sancho-Vicente, S. (2010). El ciberacoso en la enseñanza obligatoria [Cyberbullying in compulsory education]. *Aula Abierta*, *38*(1), 47-58.

Garaigordobil, M., & Aliri, J. (2013). Ciberacoso (cyberbullying) en adolescentes y jóvenes del País Vasco: Diferencias de sexo en víctimas, agresores y observadores [Cyberbullying in adolescents and young people in the Basque Country: Sex differences in victims, aggressors and observers]. *Psicologia Conductual, 21,* 461-474.

Garaigordobil, M. (2011). Prevalence and consequences of cyberbullying: A review. *International Journal of Psychology and Psychological Therapy, 11*(2), 233-254.

Garaigordobil, M., & Martínez-Valderrey, V. (2015). Effects of Cyberprogram 2.0 on "face-to-face" bullying, cyberbullying and empathy. *Psicothema, 27, 45-51.* doi: 10.7334/psicothema2014.78

Gradinger, P., Strohmeier, D. & Spiel, C. (2009). Traditional bullying and cyberbullying: Identification of risk groups for adjustment problems. *Zeitschrift Fur Psychologie/Journal of Psychology, 217,* 205-213. doi:10.1027/0044-3409.217.4.205

Hanewald, R. (2013). Cyber Bullying: Emergence, current status and future trends. In R. Hanewald (Ed.), *From cyber bullying to cyber safety. Issues and approaches in educational contexts* (pp. 3-17). New York: Nova Science Publishers.

Hinduja, S. & Patchin, J.W. (2010). Bullying, Cyberbullying, and Suicide. *Archives of Suicide Research, 14,* 206-221. doi: 10.1080/13811118.2010.494133

Hoff, D.L. & Mitchell, S.N. (2009). Cyberbullying: Causes, effects, and remedies. *Journal of Educational Administration, 47*, 652-655. doi: 10.1108/09578230910981107

Juvonen, J. & Gross, E. F. (2008). Extending the School Grounds? - Bullying Experiences in Cyberspace. *Journal of School Health*, *78*, 496-505.

Kowalski, R. & Limber, S. (2007). Electronic bullying among middle school students. *Journal of Adolescent Health, 41,* 22-30. doi:10.1016/j.jadohealth.2007.08.017.

Kowalski R, Limber, S. & Agatston, P. (2010). Cyber Bullying: El acoso escolar en la era digital [Cyber Bullying: School bullying in the digital age]. Bilbao: Desclée de Brower.

Li, Q. (2007). New bottle but old wine: A research of cyberbullying in schools. *Computers in Human Behavior, 23,* 1777-1791. doi:10.1016/j.chb.2005.10.005

Menesini E., Nocentini, A. (2009). Cyberbullying definition and measurement: Some critical considerations. *Zeitschrift Fur Psychologie/Journal of Psychology, 217*, 230–232. doi:10.1027/0044-3409.217.4.230

Menesini, E., Nocentini, A. & Calussi, P. (2011). The measurement of cyberbullying: Dimensional structure and relative item severity and discrimination. *Cyberpsychology, Behavior, and Social Networking, 14*, 267-274. DOI: 10.1089/cyber.2010.0002

Menesini, E., Nocentini, A., Palladino, B. E., Frisén, A., Berne, S., Ortega-Ruiz, R., et al. (2012). Cyberbullying definition among adolescents: A comparison across six European countries. *Cyberpsychology, Behavior and Social Networking, 15(9), 455–*463. doi:10.1089/cyber.2012.0040

Muñiz, J. (1997). *Introducción a la teoría de respuesta a los ítems [Introduction to item response theory].* Madrid: Pirámide.

Nansel, T. R., Overpeck, M., Pilla, R., Ruan, J., Simons-Morton, B. & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychological adjustment. *Journal of the American Medical Association, 285*, 2094-2100. doi:10.1001/jama.285.16.2094

Nocentini, A., Calmaestra, J., Schultze-Krumbholz, A., Scheithauer, H., Ortega, R. & Menesini, E. (2010). Cyberbullying: Labels, behaviors and definition in three European countries. *Australian Journal of Guidance and Counselling, 20(02),* 129–142. doi:10.1375/ajgc.20.2.129

Olweus, D. (1993). *Bullying at school: What we know and what we can do.* Oxford: Blackwell.

Ortega, R., Calmaestra, J. & Mora-Merchán, J. A. (2008). Cyberbullying. *International Journal of Psychology and Psychological Therapy, 8, 183-192.*

Ortega, R., Elipe, P., Mora-Merchán, J. A., Calmaestra, J. & Vega, E. (2009). The emotional impact on victims of traditional bullying and cyberbullying. A study of Spanish Adolescents. *Zeitschrift Fur Psychologie/Journal of Psychology, 217*, 197-204. doi: 10.1027/0044-3409.217.4.197

Ortega, R., Sánchez, V. & Menesini, E. (2002) Violencia entre iguales y desconexión moral: un análisis transcultural [Violence among peers and moral disconnection: a transcultural analysis]. *Psicothema, 14,* 50-62.

Patchin, J. & Hinduja, S. (2006). Bullies move beyond the schoolyard: A preliminary look at cyberbullying. *Youth Violence and Juvenile Justice, 4*, 148-169. doi: 10.1177/1541204006286288

Pornari, C. D. & Wood, J. (2010). Peer and cyber aggression in secondary school students: The role of moral disengagement, hostile attribution bias, and outcome expectancies. *Aggressive Behavior, 36*, 81–94. doi: 10.1002/ab.20336.

Raskauskas, J. & Stoltz, A. D. (2007). Involvement in traditional and electronic bullying among adolescents. *Developmental Psychology, 43, 564-575.* doi:10.1037/0012-1649.43.3.564

Sireci, S. & Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psicothema, 26*, 100-107.

Slonje, R. & Smith, P. K. (2008). Cyberbullying: Another main type of bullying. *Scandinavian Journal of Psychology*, *49*, 147-154. doi:10.1111/j.1467-9450.2007.00611.x

Smith, P.K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S. & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry, 49*, 376-385. doi:10.1111/j.1469-7610.2007.01846.x

Solberg, M. E. & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire. *Aggressive Behavior, 29(3)*, 239–268. doi:10.1002/ab.10047

Stewart, R. W., Drescher, C. F., Maack, D. J., Ebesutani, C. & Young, J. (2014). The development and psychometric investigation of the Cyberbullying Scale. *Journal of Interpersonal Violence, 29*, 2218-2238. doi: 10.1177/
0886260513517552

Thomas, H. J., Connor, J. P. & Scott, J. G. (2015). Integrating traditional bullying and cyberbullying: Challenges of definition and measurement in Adolescents – a review. *Educational Psychological Review, 27*, 135-152. doi: 10.1007/s10648-014-9261-7

Tokunaga, R.S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior*, *26*, 277-287. doi:10.1016/j.chb.2009.11.014

Willard, N.E. (2007). The authority and responsibility of school officials in responding to cyberbullying. *Journal of Adolescent Health, 41*, S64-S65.

Williams, K. & Guerra, N. (2007). Prevalence and predictors of Internet bullying. *Journal of Adolescent Health*, *41*, S14–S21. doi:10.1016/j.jadohealth.2007.08.018

Williamson, A., Lucas-Molina, B., & Guerra, N. (2013). What can high schools do to prevent cyberbullying? In R. Hanewald (Ed.), *From cyber bullying to cyber safety. Issues and approaches in educational contexts* (pp. 225-239). New York: Nova Science Publishers.

Wolak, J., Mitchell, K.J., Finkelhor, D. (2006). Online victimization of youth: 5 years later. Alexandria, VA: National Center for Missing and Exploited Children.

Ybarra, M. L., Boyd, D., Korchmaros, J. D., & Oppenheim, J. (2012). Defining and measuring cyberbullying within the larger context of bullying victimization. *Journal of Adolescent Health, 51(1*), 53–58. doi:10.1016/j.jadohealth.2011.12.031

Ybarra, M., Diener-West, M. & Leaf, P. (2007). Examining the overlap in Internet harassment and school bullying: Implications for school intervention. *Journal of Adolescent Health, 41,* 42-50. doi:10.1016/j.jadohealth.2007.09.004

Ybarra, M.L. & Mitchell, K. (2004). Online aggressor/targets, aggressors, and targets: A comparison of associated youth characteristics. *Journal of Child Psychology and Psychiatry, 45,* 1308-1316. doi:10.1111/j.1469-7610.2004.00328.x

Ybarra, M.L. & Mitchell, K.J. (2007). Prevalence and frequency of Internet harassment instigation: implications for adolescent health. *Journal of Adolescent Health, 41*, 189-195. doi: 10.1016/j.jadohealth.2007.03.005

Ybarra, M.L., Mitchell, K.J., Wolak, J. & Finkelhor, D. (2006). Examining characteristics and associated distress related to Internet harassment: Findings from the second youth Internet safety survey. *Pediatrics, 118,* 1169-1177. doi: 10.1542/peds.2006-0815

Ybarra, M. L. & Mitchell, J. K. (2008). How risky are social networking sites? A comparison of places online where youth sexual solicitation and harassment occurs. *Pediatrics*, *121*, e350–e357. doi: 0.1542/peds.2007-0693

Papeles del Psicólogo, 2016. Vol. 37(1), pp. 36-44
http://www.papelesdelpsicologo.es

ASSESSMENT OF EARLY TRAUMATIC EXPERIENCES IN ADULTS

Nuria Ordóñez-Camblor[1], Eduardo Fonseca-Pedrero[2, 3], Mercedes Paino[1, 3], Leticia García-Álvarez[1, 3], Juan Pablo Pizarro-Ruiz[4] and Serafín Lemos-Giráldez[1]

[1]*Universidad de Oviedo.* [2]*Universidad de La Rioja.* [3]*Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM).* [4]*Universidad de Burgos*

La presencia de experiencias traumáticas tempranas se ha asociado con una gran variedad de alteraciones psicopatológicas en la edad adulta. Esto hace que en la práctica clínica sea importante la evaluación y el abordaje de eventos traumáticos previos en los pacientes con diferentes trastornos mentales. El objetivo fundamental de este trabajo fue realizar una breve descripción y aproximación a los principales instrumentos de medida para la evaluación de las experiencias traumáticas tempranas, preferentemente en adultos; y presentar un nuevo instrumento de medida para la evaluación de las experiencias traumáticas tempranas en pacientes con trastorno mental grave. Los resultados mostraron que la mayoría de los instrumentos existentes proporcionan escaso o ningún apoyo psicométrico y se centran en evaluar un solo tipo de experiencia traumática, lo que reduce su aplicabilidad en la práctica clínica. Del mismo modo, son escasos los instrumentos que han sido adaptados y validados en población española. Futuros estudios deberían seguir examinando las propiedades psicométricas de este conjunto de autoinformes, con la finalidad de mejorar la evaluación de las experiencias traumáticas en la edad adulta.
Palabras clave: *Psicosis, Experiencias traumáticas tempranas, Psicopatología, Autoinfome, Evaluación.*

The presence of early traumatic experiences has been associated with a variety of psychiatric disorders in adulthood. This means that in clinical practice the assessment of and approach to previous traumatic events is important in patients with different mental disorders. The main purpose of this paper is, firstly, to provide a brief description and approximation of the main measurement instruments for the assessment of early traumatic experiences; and secondly to present the new self-report for the assessment of early traumatic experiences in patients with severe mental illness. The results show that most of the existing tools for assessing early traumatic experiences have little or no psychometric support and a number of instruments are designed to measure only one type of trauma, which reduces their clinical applicability. Similarly, there are few instruments that have been adapted to and validated in the Spanish population. Future studies should continue to examine the psychometric properties of this group of self-reports with the aim of improving the evaluation of early traumatic experiences in adulthood.
Key words: *Psychosis, Early traumatic experiences, Psychopathology, Self-report, Assessment.*

The term *early traumatic experiences* refers to different events that occur in childhood and adolescence, which are characterised by being out of the child's control, preventing or disrupting normal development, and causing stress and suffering (Burgermeister, 2007). The most common are emotional abuse, physical abuse, sexual abuse, emotional neglect and physical neglect (Bernstein et al., 2003).

It is difficult to know the true prevalence of abuse in children, since the majority of cases are usually not detected. The fact that most of these traumatic experiences occur in the family environment, the shame experienced by the victim, the early age at which they happen and the dependence on the adult, and the criminal sanctions involving the reporting of such cases are some of the factors that impede them from being known about (Goldman & Padayachi, 2000). In our country, the study conducted by the Reina Sofia Centre (Sanmartín, 2011) on child abuse in the family is noteworthy. The results showed that 4.54% of boys and 3.94% of girls between 8 and 17 years reported suffering abuse by a family member. The most prevalent types of abuse in this age group are: psychological (2.35%), followed by physical (2.24%), sexual (0.89%) and finally, negligence (0.78%). With regards to sex, boys are at greater risk of suffering physical abuse (2.41%), while girls suffer higher rates of psychological abuse (2.72%), sexual abuse (1.13%) and negligence (0.91%). Similarly, the results indicate that experiences of abuse decrease as the age of the child increases. Thus, the prevalence of abuse between 8 and 11 years stood at 5.05%, while between 12 and 14 years it was 4.65%.

The presence of stressful events or occurrences has important implications for people's physical and emotional health (Thabrew, de Sylva & Romans, 2012). During childhood, these traumas can have a more significant impact, since they do not affect a biological, psychological and socially mature human being, but a human being in a phase of development that requires certain external conditions of stability and protection (López-Soler, 2008). Thus, the fact that the human brain continues to develop during childhood and adolescence, and even during the adult period, makes it especially vulnerable to traumatic situations or chronic stress and can cause damage, sometimes irreversible, of a physical, emotional and cognitive nature (Mesa-Gresa & Moya-Abiol, 2011). In general, the following reactions to a seriously stressful event are considered normal: sadness, anxiety, anger, altered behaviour and other minor difficulties that disrupt for a short time (López-Soler, 2008). However, sometimes the difficulties are more intense and longer lasting, resulting in significant problems in the personal functioning and psychosocial adaptation of the child. Thus, negative occurrences, such as physical abuse,

Correspondencie: Nuria Ordóñez Camblor. *Facultad de Psicología. Universidad de Oviedo. Plaza Feijóo, s/n 33003. Oviedo. España. Email: nuorcapsico@hotmail.com*

emotional neglect, sexual abuse and others may produce negative psychological effects, both in the short and long term (Pereda, 2009, 2010) (see Figure 1).

Among the short and medium term psychological effects, it has been found that episodes of abuse cause changes in the emotional development and personality of the child. Different studies show the profile of the child with experience of abuse as insecure, with low self-esteem and difficulties in social relations (Flynn, Cicchetti, & Rogosch 2014; Young & Widom, 2014), presenting great difficulty in expressing and recognising emotions and having more negative emotions (Shenk, Putnam & Noll, 2013). Similarly, often they lack essential positive beliefs about themselves and their world, showing less skills in recognising and responding to the distress of others (Sanmartin, 2011). It has also been found that these children are more likely to exhibit delays in cognitive development, showing gaps in the development of language, low scores on intelligence tests and generally poor school performance (Merritt & Klein, 2015; Viezel, Freer, Lowell & Castillo, 2014).

These deficits, far from abating over time, can sometimes persist into adulthood, leading to different psychopathologies. The presence of early traumatic experiences has been associated with increased vulnerability to psychopathology and a worse physical condition in adulthood (Shonkoff et al., 2012). At the psychiatric level, numerous studies have found a relationship between childhood trauma and various psychopathological disorders, such as mood disorders and anxiety, post-traumatic stress disorder, dissociative disorders, psychotic disorders, and substance use disorders, among others (Agorastos, et al, 2014; Park et al, 2014; Van Nierop et al, 2015). It is estimated that between 34 and 53% of patients with mental health problems have a history of physical and sexual abuse during childhood (Alarcón, Araujo, Godoy & Vera, 2010). Similarly, the existence of early trauma has been associated with various complications in the course of mental disorders, such as increased comorbidity and severity, worse response to drug treatment and worse prognosis (Álvarez et al, 2011; Nanni, Uher & Danese, 2012; Teicher & Smamson, 2013).

Within the trauma and psychopathology relationship, interest in the study of the relationship between psychosis and early traumatic experiences has increased in recent years. Different studies show that traumatic experiences can play a causal role in the development of psychotic disorders (Falukozi & Addington, 2012; Lataster, Myin-Germeys, Lieb, Wittchen & van Os, 2012; Thompson et al, 2014.) Recently, Varese and colleagues (2012) conducted a comprehensive meta-analysis, finding a significant relationship between the presence of different types of traumatic experiences and psychosis (odds ratio (OR) = 2.78, $p \leq 0.001$). In Spain, there are few studies measuring this relationship, with the results being consistent with those found in international samples; and between 40 and 75% of the patients with psychosis studied reported the previous presence of traumatic experiences (Álvarez et al, 2011; Ordóñez-Camblor et al, 2014). Similarly, recent studies highlight the role of traumatic experiences before the development of psychosis, as a possible trigger factor; a history of traumatic experiences has been linked with the transition to psychosis in people at high risk (Bechdolf et al, 2010; Thompson et al, 2014; Tikka et al, 2014).

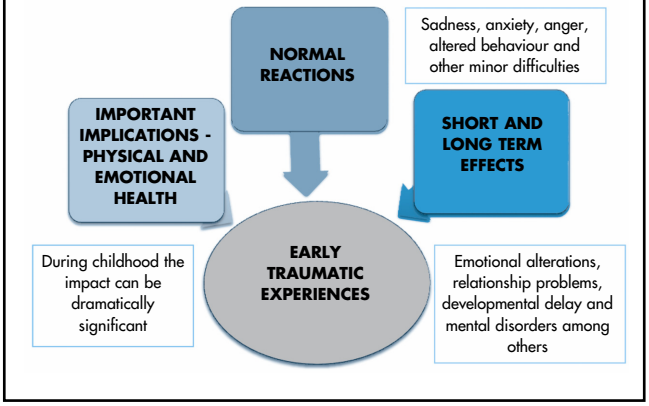
The study of the relationship between psychosis and traumatic experiences has also focused on the effect that these experiences can have on the course and outcome of psychotic disorders. Overall, the patients who report a history of trauma present a more severe clinical profile, with an earlier onset of symptoms, more positive symptoms, especially hallucinations, a higher number of admissions and more suicide thoughts and attempts (Conus, Cotton, Schimmelmann, McGorry & Lambert, 2010; Daahlman et al, 2012). Similarly, they have lower adherence to treatment and increased comorbid symptomatology (Bendall, Álvarez-Jiménez, Nelson & McGorry, 2013; Schäfer et al, 2012).

Within this context, the purpose of this study was to perform a brief description and approach to the main measuring instruments for the assessment of early traumatic experiences, preferably in adults. The assessment and treatment of traumatic experiences is of utmost importance for the prevention and intervention of psychopathological alterations. In this sense, it is relevant to have short and simple measuring instruments to identify the presence of different early traumatic experiences, in order to carry out a more precise and thorough psychological evaluation of these experiences and to guide the intervention and the possible therapeutic process.

INSTRUMENTS FOR ASSESSING EARLY TRAUMATIC EXPERIENCES: A SELECTIVE REVIEW

Due to the growing interest in the study of the presence of child traumatic experiences, and given their frequency and their impact on the psychological adjustment of the people who experience them, in recent years the development of different questionnaires for assessing the presence of such experiences has increased (Donald, 2012). In Spain, there are few studies that have carried out a review of the various measuring instruments currently used for the assessment of early traumatic experiences. At the international level, various authors have recently undertaken reviews of the instruments for assessing early trauma (see Burgermeister, 2007; Roy & Perry, 2004; Thabrew et al, 2012). Such reviews agree in affirming that most of the instruments do not provide information about their psychometric properties, which

FIGURE 1
THE IMPLICATIONS OF EARLY TRAUMATIC EXPERIENCES



makes them less useful in both clinical practice and research (Thabrew et al., 2012). Similarly, many of the instruments focus on evaluating a single type of traumatic experience (Roy & Perry, 2004). Tables 1 and 2 show the main interviews and self-reports for assessing children's traumatic experiences that exist internationally.

If we consider the tools for evaluating various types of traumatic experiences, the most used ones that have provided adequate psychometric properties and have been adapted and validated into Spanish are the *Childhood Trauma Questionnaire* (CTQ) (Bernstein, Ahluvalia, Pogge & Handelsman, 1994) and the *Early Trauma Inventory* (ETI) (Bremner, Vermetten & Mazure, 2000).

The *Childhood Trauma Questionnaire* (CTQ) (Bernstein et al., 1994) is the most widely used self-report used in the assessment of traumatic childhood experiences (Thabrew et al., 2012). It includes five types of childhood trauma: emotional abuse, physical abuse, sexual abuse, physical neglect and emotional neglect. The CTQ consists of 70 items of Likert format with 5 categories (0 = "never" and 5 = "very often"). The internal consistency in international studies, is high, both for the total scale (0.98), and for the different factors, ranging from 0.79 to 0.94. Similarly, the test-retest reliability for the whole scale is 0.88, while for the individual factors it was between 0.80 and 0.83. Later, Bernstein et al. (2003) developed a short version, *The Childhood Trauma Questionnaire-Short Form* (CTQ-SF), composed of 28 items. The CTQ-SF has been translated and adapted to several languages, showing adequate psychometric properties (Grassi-Oliveira et al., 2014). Recently the CTQ-SF has been adapted and translated into Spanish by Hernández et al. (2013) in a sample of 185 women with various mental disorders, showing adequate psychometric properties. The Cronbach's alpha coefficient is between 0.66 and 0.94. The factor analysis supported a five-factor structure originally proposed by Bernstein et al., (2003). Similarly, the correlation factor is high, ranging between 0.29 and 0.50.

The Early Trauma Inventory (ETI) (Bremner et al., 2000) has been developed for the evaluation of different types of abuse –such as sexual, physical and emotional abuse– and other traumas –such as witnessing violent acts. The ETI has 56 items in dichotomous Yes / No format, through which it evaluates the different types of trauma, the frequency of the abuse, the age at which the abuse began, the perpetrator of the abuse and the impact of the abuse. The internal consistency, in studies carried out with foreign samples, is between 0.86 and 0.92, with the test-retest reliability oscillating between 0.51 and 0.99. There is also an abridged version, *The Early Trauma Inventory-Short Form* (SF-TSI) (Bremmer et al., 2007), consisting of 27 items in dichotomous Yes / No response format, which, like the original version, evaluates four

TABLE 1
INTERVIEWS FOR THE ASSESSMENT OF CHILDREN'S TRAUMATIC EXPERIENCES

Name of the scale	Reference	Abbrev.
Abuse History	(Soloff, Lynch & Kelly, 2002)	AH
Brief Physical and Sexual Abuse Questionnaire	(Marshall et al., 1998)	BPSAQ
Childhood Experience of Care and Abuse	(Bifulco, Brown & Harris, 1994).	CECA
Childhood Life Events and Family Characteristics Questionnaire	(Byrne, Velamoor, Cernovsky, Cortese & Losztyn, 1990)	CLEFCQ
Childhood Maltreatment Interview Schedule	(Briere, 1992, Briere, Elliott, Harris & Cotman, 1995)	CMIS y CMIS-SF
Childhood Trauma Interview	(Fink, Bernstein, Hadelsman, Foote & Lovejoy, 1995)	CTI
Developmental Interview	(Paris, Zweig-Frank & Guzder, 1994)	DI
Early Home Environment Interview	(Lizardi et al., 1995)	EHEI
Early Trauma Inventory	(Bremner et al., 2000)	ETI
Family Experience Interview	(Ogata et al., 1990)	FEI
Family Interview for Protectiveness and Empathy	(Laporte & Guttman, 2001)	FIPE
History of Physical and Sexual Abuse Questionnaire	(Meyer, Muenzenmaier, Cancienne & Struening, 1996)	HPSAQ
Interview for Traumatic Events in Childhood	(Lobbestael, Arntz, Harkema-Schouten & Bernstein, 2009)	ITEC
Instrument on child sexual abuse	(Russell, 1986)	
Retrospective Assessment of Traumatic Experience	(Gallaghe, Flye, Hurt, Stone & Hull, 1992)	RATE
Retrospective Childhood Experience Questionnaire	(Zanarini, Gunderson, Marino, Schwartz & Frankenburg, 1989)	RCEQ
Retrospective Family Pathology Questionnaire	(Zanarini et al., 1989)	RFPQ
Retrospective Separation Experience Questionnaire	(Zanarini et al., 1989)	RSEQ
Sexual Abuse Severity Scale	(Silk, Lee, Hill & Lohr, 1995)	SASS
Structured Trauma Interview	(Draijer & Langeland, 1999)	STI
Traumatic Antecedents Interview	(Herman, Perry & Van der Kolk, 1989)	TAI
Trauma History Screen	(Allen, Huntoon & Evans, 1999)	THS
Unwelcome Childhood Sexual Events	(Russ, Shearin, Clarkin, Harrison & Hull, 1993)	UCSE

Note. Abbrev: Abbreviation

dimensions (general trauma, sexual abuse, physical abuse and emotional abuse). The ETI-SF has been adapted and translated into different languages, presenting adequate psychometric properties (Osóiro et al., 2013).

In our country, Plaza et al. (2011) have carried out the validation and adaptation into Spanish of the ETI and ETI-SF in a sample of 227 postpartum women. The results indicate that the Spanish version of the ETI presents adequate psychometric properties. The reliability of the scores relating to the internal consistency (estimated using Cronbach's alpha) for the overall scale was 0.79, while the values for the subscales ranged between 0.58 and 0.76). Meanwhile the test-retest was 0.92 for the global scale and between 0.76 and 0.95 for the different subscales. Similarly, the results indicate that the ETI-SF had adequate psychometric properties. The Cronbach's alpha coefficient for the total score was 0.72, while the values for the subscales ranged from 0.42 to 0.72. Obtaining evidence of validity of the two instruments was performed by analysis of the ROC curve. The results indicate that both instruments have adequate evidence of validity, although the ETI-SF is slightly weaker in the detection of physical abuse.

DEVELOPMENT AND VALIDATION OF AN ASSESSMENT INSTRUMENT FOR THE DETECTION OF EARLY TRAUMATIC EXPERIENCES IN PATIENTS WITH SEVERE MENTAL DISORDERS: THE EXPTRA-S

Within this research context, at the national level, as yet no instrument has been specifically built, weighted and validated for the assessment of traumatic experiences in clinical population, particularly in patients with severe mental illness (SMI) (e.g., psychosis). Also, the vast majority of assessment instruments developed for this purpose do not incorporate new developments in psychological measurement (e.g., the construction, translation and/or adaptation of the test, or the construction of items) or new statistical procedures, such as item response theory (IRT) or differential item functioning (DIF). Similarly, there are few instruments that are available in the specialist literature, that evaluate the distress associated with these experiences. This is important since, from a clinical point of view, the way in which the patient processes and manages the trauma is critical in working through the traumatic experiences.

Recently our research team developed the *Cuestionario de Screening*

TABLE 2
SELF-REPORTS FOR ASSESSING CHILDREN'S TRAUMATIC EXPERIENCES

Name of the scale	Reference	Abbrev.
Adverse Childhood Experiences Study Questionnaire	(Dube, Williamson, Thompson, Felitti & Anda, 2004)	AEQ
Assessing Environments III	(Berger, Knuston, Mehm & Perkins, 1988)	AEII
Computer Assisted Maltreatment Inventory	(DiLillo et al., 2010)	CAMI
Childhood Abuse and Trauma Scale	(Sanders & Becker-Lausen, 1995)	CATS
Childhood Experiences Questionnaire	(Ferguson & Dacey, 1997)	CEQ
Childhood Unwanted Sexual Events	(Lange, Kooiman, Huberts & van Oostendorp, 1995)	CHUSE
Child Maltreatment History Self-Report	(MacMillan et al., 1997)	CMHSR
Childhood Traumatic Events Scale	(Pennebaker & Susman, 1988)	CTES
Childhood Trauma Questionnaire	(Bernstein et al.,1994)	CTQ
Childhood Violence Scale	(Riggs, O'Leary & Breslin, 1990)	CVS
Comprehensive Child Maltreatment Scales for Adults	(Higgins y McCabe, 2001)	CCMS-A
Life Experience Questionnaire	(Bryer, Nelson, Miller & Krol, 1987)	LEQ
Neglect Scale	(Harrington, Zuravin, DePanfilis, Ting & Dubowitz, 2002)	NS
Parental Physical Maltreatment Scale	(Briere & Runtz, 1990)	PHY
Psychological Maltreatment Inventory	(Engels & Moisan, 1994)	PMI
Physical and Sexual Abuse Questionnaire	(Nagata, Kiriike, Iketani, Kawarada & Tanaka, 1999)	PSA
Parental Psychological Maltreatment Scale	(Briere & Runtz, 1990)	PSY
Revised Childhood Experiences Questionnaire	(Zanarini et al., 1997)	RCEQ
Sexual Abuse Exposure Questionnaire	(Ryan, Rodríguez, Rowan y Foy, 1992)	SAEQ
Sexual Abuse Questionnaire	(Finkelhor, 1979)	SAQ
Sexual Events Questionnaire	(Calam & Slade, 1989)	SEQev
Sexual Experience Questionnaire	(Wagner & Linehan, 1994)	SEQex
Sexual Life Events Inventory	(Palmer,Chaloner & Oppenheimer, 1992)	SLEI
Sexual and Physical Abuse History Questionnaire of Leserman andcolleagues	(Leserman & Drossman, 1995)	SPAHQ
Stressful Life Events Screening Questionnaire	(Goodman, Corcoran, Turner, Yuan & Green, 1998)	SLESQ
Traumatic Experiences Questionnaire	(Nijenhuis, Spinhoven, van Dyck, van de Hart & Vanderlinden, 1998)	TEQ

Note. Abbrev: Abbreviation

de Experiencias Traumáticas [*Questionnaire for the Screening of Traumatic Experiences*] (ExpTra-S), a short, simple and useful instrument for assessing, through screening, the frequency and distress of early traumatic experiences frequently found in patients with SMI. For further detail please consult Camblor Ordoñez (2015). This is not an assessment instrument that covers all of the possible traumatic experiences; however, it does cover the traumatic experiences that are considered most frequent in childhood (Bernstein et al., 2003). Furthermore, it is intended for use as a screening method, so the information found must be completed using other methods of assessment and different informants.

The construction process of the ExpTra-S, was conducted according to the international guidelines for the construction of assessment instruments (American Educational Research Association et al., 1999; Downing, 2006; Schmeiser & Welch, 2006; Wilson, 2005), following a series of steps that would ensure that the construction process was carried out in a systematic and rigorous way (Muñiz & Fonseca-Pedrero, 2008). The item bank was built based on a comprehensive review of the existing tools for assessing early traumatic experiences in adults and the judgement of experts in the field. The items that made up the bank were selected or modified from different scales and/or newly created ones. All of the items were constructed and drafted based on the principles of simplicity, clarity, comprehensibility and relevance to the population of interest. The translation, adaptation and construction of the items was conducted in accordance with international guidelines for the translation and adaptation of tests (Hambleton, Merenda & Spielberg, 2005; Muñiz & Bartram, 2007; Muñiz, Elosua & Hambleton, 2013), and the construction of multiple-choice items (Haladyna, 2002; Moreno, Martínez & Muñiz, 2006).

The ExpTra-S, has two scales, one of frequency and another of distress. The frequency scale is composed of a total of 18 items in Likert response format of four categories (0 "never", 1 "sometimes", 2 "frequently", 3 "almost always"). The presence of early traumatic experiences is evaluated through 17 questions regarding different types of child abuse: sexual abuse, physical abuse, emotional abuse, and emotional and physical neglect, adding a final item that refers to any other type of traumatic event that may have occurred which has not been covered in previous questions and which has caused the participant distress. Similarly, the scale of distress is also made up of 18 items in Likert format with four categories (1, "no distress", 2 "slight distress", 3 "considerable distress" and 4 "great distress"), where the distress associated with these experiences is evaluated. The scale of distress should be answered only if the trauma is present at least "sometimes" on the frequency scale. An example of an item could be: "When you were a child, did a family member regularly and repeatedly insult you?"

The construction and validation of the ExpTra-S, was conducted with a sample of 114 patients with psychotic disorders and 153 young non-clinical adults, and presented adequate psychometric properties. The estimation of the reliability showed an internal consistency of 0.96, with all indices of discrimination greater than 0.30. The reliability of the scores was also estimated using IRT. Validity studies allowed us to collect enough evidence that could serve as the scientific basis for the interpretation of the scores of participants of the ExpTra-S. Similarly, validity evidence was obtained with other self-reports that evaluated psychotic symptoms and subjective complaints of cognitive deficits in patients with psychosis (Ordóñez-Camblor, 2015). It would be interesting for future studies to apply ExprTra-S in patients with other serious mental disorders.

BY WAY OF CONCLUSION

A topic of growing interest is the study of the relationship between early adversity and psychological difficulties in later phases of life. In this regard, recent research indicates poorer mental health in general in people who have suffered abuse, with greater presence of symptoms and psychiatric disorders (Carr, Martins, Stingel, Lembruber & Juruena, 2013; Sala, Goldstein, Wang & Blanco, 2014; Subica, 2013). The importance of the trauma and psychopathology relationship has led to the creation of different assessment instruments.

In this sense, the purpose of this research was to conduct a review of the main existing self-reports for assessing early traumatic experiences in adults; and to present a new measuring instrument for assessing early traumatic experiences in patients with SMI. The results show that, although a large number of instruments have been built over the past 30 years focusing on the assessment of adult traumatic experiences, more studies are still needed to further facilitate development and knowledge in the field of assessing traumatic experiences. There is great heterogeneity among the instruments, not only in the formats and methods of administration, but also in the kinds of traumatic experiences that they focus on, which makes comparison between them difficult. At the same time, not all of the instruments developed have provided information about their psychometric properties. Similarly, there are few instruments that have been adapted and validated in the Spanish population. These and other limitations reduce the clinical applicability of these instruments (Thabrew et al., 2012). There is no doubt that the use of assessment instruments with adequate metric quality, upon which solid and well-founded decisions can be based, is a must from both the clinical and research perspectives (Fonseca et al., 2011).

On the other hand, until now none of the existing instruments had been developed specifically for assessing early traumatic experiences in patients with SMI. In this sense, the ExpTra-S is an instrument that can facilitate the brief and simple assessment of early traumatic experiences in patients with SMI.

The assessment of early traumatic experiences is extremely important from the clinical point of view. The early identification and intervention of abusive experiences may decrease the development of mental disorders in adulthood. In the same vein, the presence of early traumatic experiences in a patient with a mental disorder may hinder the therapeutic process and it may be an indicator of poor prognosis as well as influencing the evolution of the clinical condition and the therapy or prophylactic treatment. Thus, participants that report traumatic experiences with some distress should be subject to monitoring as well as a specific intervention during the therapy in order to work through the early traumatic experiences and to reduce the associated distress.

For future work in this line, it is important to continue to obtain validity evidence of the ExpTra-S and to continue to examine the psychometric properties of the measuring instruments available for assessing early traumatic experiences.

REFERENCES

Agorastos, A., Pittman, J. O., Angkaw, A. C., Nievergelt, C. M., Hansen, C. J., Aversa, L. H., … Baker, D. G. (2014). The cumulative effect of different childhood trauma types on self-reported symptoms of adult male depression and PTSD, substance abuse and health-related quality of life in a large active-duty military cohort. *Journal of Psychiatric Research, 58,* 46-54. doi: 10.1016/j.jpsychires.2014.07.014

Alarcón, L. C., Araujo, P. A., Godoy, A. P. & Vera, M. E. (2010). Maltrato infantil y sus consecuencias a largo plazo [Child abuse and its long-term consequences]. *MedUnab, 13*(2), 103-115.

Alle, J. G., Huntoon, J. & Evans, R. B. (1999). A self-report measure to screen for trauma history and its application to women in inpatient treatment for trauma-related disorders. *Bulletin of the Menninger Clinic, 63*, 429-442.

Álvarez, M. J., Roura, P., Osés, A., Foguet, Q., Solà, J. & Arrufat, F. X. (2011). Prevalence and clinical impact of childhood trauma in patients with severe mental disorders. *Journal of Nervous and Mental Disease, 199*(3), 156-161. doi: 10.1097/NMD.0b013e31820c751c

American Educational Research Association, American Psychological Association and National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing.* Washington, D.C.: Author.

Arruabarrena, M. I. & De Paúl, J. (1996). *Maltrato a los niños en la familia. Evaluación y Tratamiento [Child abuse in the family. Evaluation and Treatment.].* Madrid: Pirámide.

Bechdolf, A., Thompson, A., Nelson, B., Cotton, S., Simmons, M. B., Amminger, G. P., … Yung, A. R. (2010). Experience of trauma and conversion to psychosis in anultra-high-risk (prodromal) group. *Acta Psychiatrica Scandinavica, 121*(5), 377-384. doi: 10.1111/j.1600-0447.2010.01542.x

Bendall, S., Álvarez-Jiménez, M., Nelson, B. & McGorry, P. (2013). Childhood traumaand psychosis: new perspectives on aetiology and treatment. *Early Intervention in Psychiatry, 7*(1), 1-4. doi: 10.1111/eip.12008

Berger, A. M., Knutson, J. F., Mehm, J. G. & Perkins, K. A. (1988). The self-report of punitive childhood experiences of young adults and adolescents. *Child Abuse & Neglect, 12,* 251-262.

Bernstein, D. P., Ahluvalia, T., Pogge, D. & Handelsman, L. (1997).Validity of the CTQ in adolescent psychiatric population. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,*340-348.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., …

Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect, 27*(2), 169-190. doi:10.1016/s0145-2134(02)00541-0

Bifulco, A., Brown, G. W. & Harris, T. O. (1994). Childhood experiences of care and abuse (CECA): A retrospective interview measure. *Journal of Child Psychology and Psychiatry, 35,* 1419-1435.

Bremner, J. D., Bolus, R. & Mayer, E. A. (2007). Psychometric properties of the Early Trauma Inventory-Self Report. *Journal of Nervous and Mental Disease, 195*(3), 211-218. doi:10.1097/01.nmd.0000243824.84651.6c

Bremner, J. D., Vermetten, E. & Mazure, C. M. (2000). Development and preliminary psychometric properties of an instrument for the measurement of childhood trauma: The early trauma inventory. *Depression and Anxiety, 12*, 1-12.

Briere, J. (1992). Child abuse trauma: Child maltreatment interview schedule. In J. R. Conte (Series Ed), *Theory and treatment of the lasting effects: Interpersonal violence: The practice series* (pp. 165–177). Newbury Park, CA: Sage Publications.

Briere, J. & Runtz, M. (1990). Differential adult symptomatology associated with three types of child abuse histories. *Child Abuse & Neglect, 14*, 357-364.

Bryer, J. B., Nelson, B. A., Miller, J. B. & Krol, P. A. (1987). Childhood sexual and physical abuse as factors in adult psychiatric illness. *American Journal of Psychiatric, 144,* 1426-1430.

Burgermeister, D. (2007). Childhood adversity: a review of measurement instruments. *Journal of Nursing Measurement, 15*(3), 163-176.

Byrne, C. P., Velamoor, V. R., Cernovsky, Z. Z., Cortese, L. & Losztyn, S. (1990). A comparison of borderline and schizophrenic patients for childhood life events and parent-child relationships. *Canadian Journal of Psychiatry, 35,* 590-595.

Calam, R. M. & Slade, P. D. (1989). Sexual experiences and eating problems in female undergraduates. *International Journal of Eating Disorder, 8*, 391-397.

Carr, C. P., Martins, C. P., Stingel, A. M., Lembruber, V. B. & Juruena, M. F. (2013). The role of early life stress in adult psychiatric disorders: A systematic review according to childhood trauma subtypes. *Journal of Nervous and Mental Disease, 201*(12), 1007-1020. doi:10.1097/NMD.0000000000000049

Centre International de L´enfance. (1980). *Los niños víctimas de los servicios y maltratos [The child victims of services and abuse].* Document for the ICC Project. Paris.

Conus, P., Cotton, S., Schimmelmann, B. G., McGorry, P. D. & Lambert, M. (2010). Pretreatment and outcome correlates of sexual and physical trauma in an epidemiological cohort of first-episode psychosis patients. *Schizophrenia Bulletin, 36*(6), 1105-1114. doi: 10.1093/schbul/sbp009

Daahlman, K. M., Diederen, M. J., Derks, E. M., van Lutterveld, R., Kahn, R. S. & Sommer, I. E. (2012). Childhood trauma and auditory verbal hallucinations. *Psychological Medicine, 42*(12), 2475-2484. doi: 10.1017/S0033291712000761

Davidson, J. R., Hughes, D. & Blazer, D. G. (1990). Traumatic experiences in psychiatric outpatients. *Journal of Traumatic Stress, 3*, 459-475.

Donald, J. L. (2012). A review of childhood abuse questionnaires and suggested treatment approaches. In E. S. Kalfo lu (Ed.), *Sexual Abuse-Breaking the Silence.* Rijeka, Croatia: In Tech.

Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing and T. M. Haladyna (Eds.), *Handbook of test development* (pp. 3-25). Mahwah, NJ: Lawrence Erlbaum Associates

Draijer, N. & Langeland, W. (1999) Childhood trauma and perceived parental dysfunction in etiology of dissociative symptoms in

psychiatric inpatients. *American Journal of Psychiatry, 156,* 379-385.

Dube, S. R., Williamson, D. F., Thompson, T., Felitti, V. J. & Anda, R. F. (2004). Assessing the reliability of retrospective reports of adverse childhood experiences among adult HMO members attending a primary care clinic. *Child Abuse & Neglect, 28*(7), 729-737. doi:10.1016/j.chiabu.2003.08.009

Engels, M. & Moisan, D. (1994). The psychological maltreatment inventory: development of a measure of psychological maltreatment in childhood for use in adult clinical settings. *Psychological Reports, 74*, 595-604.

Falukozi, E. & Addington, J. (2012). Impact of trauma on attenuated psychotic symptoms. *Psychosis, 4*(3), 203-212.

Ferguson, K. S. & Dacey, C. M. (1997). Anxiety, depression and dissociation in women health care providers reporting a history of childhood psychological abuse. *Child Abuse & Neglect, 21,* 941-952.

Fink, L. A., Bernstein, D., Hadelsman, L., Foote, J. & Lovejoy, M. (1995). Initial reliability and validity of the childhood trauma interview. *The American Journal of Psychiatry, 152,* 1329-1335.

Finkelhor, D. (1979). *Sexually Victimized Children.* New York, NY: Free Press.

Flynn, M., Cicchetti, D. & Rogosch, F. (2014). The prospective contribution of childhood maltreatment to low self-worth, low relationship quality, and symptomatology across adolescence: A developmental-organizational perspective. *Developmental Psychology, 50*(9), 2165-2175.doi:10.1037/ a0037162

Foa, E. B., Cashman, L., Jaycox, L. & Perry, K. (1997). The validation of a self-report measure of posttraumatic stress disorder: The Postraumatic Diagnostic Scale. *Psychological Assessment, 9*, 445-451.

Fonseca-Pedrero, E., Lemos-Giráldez, S., Paíno, M., Santarén-Rosell, M., Sierra-Baigrie, S. & Ordóñez-Camblor, N. (2011). Instrumentos de medida para la evaluación del fenotipo psicótico [Assessment Instruments for Psychosis Phenotype]. *Papeles del Psicólogo, 32,* 129-151.

Gallagher, R. E., Flye, B. L., Hurt, S. W., Stone, M. H. & Hull, J. W. (1992). Retrospective assessment of traumatic experiences. *Journal of Personality Disorder, 6*, 99-108.

Goldman, J. D. G. & Padayachi, U. K. (2000). Some methodological problems in estimating incidence and prevalence in child sexual abuse research. *The Journal of Sex Research, 37*(4), 305–314.

Goodman, L. A., Corcoran, C., Turner, K., Yuan, N. & Green, B. L. (1998). Assessing traumatic event exposure: general issues and preliminary findings for the Stressful Life Events Screening Questionnaire. *Journal of Traumatic Stress, 11,* 521-542.

Grassi-Oliveira, R., Cogo-Moreira, H., Salum, G. A., Brietzke, E., Viola, T. W., Manfro, G. G., … Arteche, A. X. (2014). Childhood Trauma Questionnaire (CTQ) in Brazilian Samples of Different Age Groups: Findings from Confirmatory Factor Analysis. *PLoS ONE, 9*(1), e87118. doi:10.1371/ journal.pone.0087118

Haladyna, T. M. (2002). A review of multiple-choice item-writing guidelines. *Applied Measurement in Education, 15*(3), 309-334.

Hambleton, R. K., Merenda, P. F. & Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment.* London: Lawrence Erlbaum Associates.

Herman, J. L., Perry, J. C. & Van der Kolk, B. A. (1989). Childhood trauma in borderline personality disorder. *American Journal of Psychiatry, 146*, 490-495.

Hernández, A., Gallardo-Pujol, D., Pereda, N., Arntz, A., Bernstein, D. P., Gaviria, A. M., … Gutierrez-Zotes, J. A. (2013). Initial Validation of the Spanish Childhood Trauma Questionnaire-Short Form: Factor Structure, Reliability and Association With Parenting. *Journal of Interpersonal Violence, 28*(7), 1498-1518. doi:10.1177/ 0886260512468240

Higgins, D. J. & McCabe, M. (2001). The development of the comprehensive child maltreatment scale. *Journal of family studies, 7*(1), 7-28.

Lange, A., Kooiman, K., Huberts, L. & van Oostendorp, E. (1995). Childhood unwanted sexual events and degree of psychopathology of psychiatric patients: Research with a new anamnestic questionnaire (the CHUSE). *Acta Psychiatrica Scandinavica, 92,* 441-446.

Laporte, L. & Guttman, H. (2001). Abusive relationships in families of women with borderline personality disorder, anorexia nervosa and a control group. *Journal of Nervous and Mental Disease, 189*, 522-531.

Lataster, J., Myin-Germeys, I., Lieb, R., Wittchen, H. U. & van Os, J. (2012). Adversity and psychosis: a 10-year prospective study investigating synergism between early and recent adversity in psychosis. *Acta Psychiatrica Scandinavica, 125*(5), 388-399. doi: 10.1111/j.1600-0447.2011.01805.x

Lizardi, H., Klein, D. N., Ouimette, P. C., Riso, L. P., Anderson, R. L. & Donaldson, S. K. (1995). Reports of the childhood home environment in early-onset dysthymia and episodic major depression. *Journal of Abnormal Psychology, 104,* 132-139.

Lobbestael, J., Arntz, A., Harkema-Schouten, P. & Bernstein, D. (2009). Development and psychometric evaluation of a new assessment method for childhood maltreatment experiences: The interview for traumatic events in childhood (ITEC). Child Abuse & Neglect, 33(8), 505–517. doi:10.1016/j.chiabu. 2009. 03.002

López-Soler, C. (2008). Las reacciones postraumáticas en la infancia y adolescencia maltratada: el trauma complejo [Posttraumatic reactions in child and adolescent abuse: the complex trauma]. *Revista de Psicopatología y Psicología clínica, 13*(3), 159-179.

MacMillan, H. L., Fleming, J. E., Trocme, N., Boyle, M. H., Wong, M., Racine, Y. A., … Offord, D. R. (1997). Prevalence of child physical and sexual abuse in the community: Results from the Ontario health supplement. *JAMA*, 278, 131-135.

Marshall, R. D., Schneider, F. R., Fallon, B. A., Knight, C. B. G., Abbate, L. A., Goetz D, C., … Liebowtiz, M. R. (1998). An open trial of paroxetine in patients with noncombat-related, chronic posttraumatic stress disorder. *Journal of Clinical Psychopharmacology, 18,* 10-18.

Merritt, D. H. & Klein, S. (2015). Do early care and education services improve language development for maltreated children? Evidence from a national child welfare sample. *Child Abuse & Neglect, 39,* 185-196. doi: 10.1016/j.chiabu. 2014.10.011

Mesa-Gresa, P & Moya-Albiol, L. (2011). Neurobiología del maltrato

infantil: el "ciclo de la violencia" [Neurobiology of child abuse: the "cycle of violence"]. *Revista de Neurología, 52*(8), 489-503.

Meyer, I. H., Muenzenmaier, K., Cancienne, J. & Struening, E. (1996). Reliability and validity of a measure of sexual and physical abuse histories among women with serious mental illness. *Child Abuse & Neglect, 20*, 213-219.

Moreno, R., Martínez, R. & Muñiz, J. (2006). New guidelines for developing multiple choice items. *Methodology, 2*(2), 65-72. doi:10.1027/1614-2241.2.2.65

Muñiz, J. & Bartram, D. (2007).Improving international tests and testing. *European Psychologist, 12*(3), 206-219. doi:10.1027/1016-9040.12.3.206

Muñiz, J., Elosua, P. & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición [International Test Commission Guidelines for test translation and adaptation: Second edition]. *Psicothema, 25,*151-157.

Muñiz, J. & Fonseca-Pedrero, E. (2008). Construcción de instrumentos de medida para la evaluación universitaria [The construction of measuring instruments for university evaluation]. *Revista de Investigación en Educación, 5,* 13-25.

Nagata, T., Kiriike, N., Iketani, T., Kawarada, Y. & Tanaka, H. (1999). History of childhood sexual or physical abuse in Japanese patients with eating disorders: Relationship with dissociation and impulsive behaviours. *Psychological Medicine, 29*, 935-942.

Nanni, V., Uher, R. & Danese, A. (2012). Childhood Maltreatment Predicts Unfavorable Course of Illness and Treatment Outcome in Depression: A Meta-Analysis. *The American Journal of Psychiatry, 164*(4), 141-151.

Nijenhuis, E. R. S., Spinhoven, P., van Dyck, R., van de Hart, O. & Vanderlinden, J. (1998). Degree of somatoform and psychological dissociation in dissociative disorder is correlated with reported trauma. *Journal of Traumatic Stress, 11*, 711-730.

Ogata, S. N., Silk, K. R., Goodrich, S., Lohr, N. E., Westen, D. & Hill, E. M. (1990). Childhood sexual and physical abuse in adult patients with borderline personality disorder. *The American Journal of Psychiatry, 147*, 1008-1013.

Ordóñez-Camblor, N. (2015). *Experiencias traumáticas tempranas y procesos de recuperación en los trastornos psicóticos [Early traumatic experiences and recovery processes in psychotic disorders*. (Unpublished doctoral thesis). University of Oviedo. Oviedo.

Ordóñez-Camblor, N., Lemos-Giráldez, S., Paino, M., Fonseca-Pedrero, E., García-Álvarez, L. & Pizarro-Ruiz, J. P. (2014). Relación entre psicosis y experiencias traumáticas tempranas [Relationship between psychosis and early traumatic experiences]. *Anuario de Psicología/The UB Journal of Psychology, 44*(3), 283-294.

Organización Mundial de la Salud [World Health Organization] (1999). *WHO: Report of the Consultation on Child Abuse Prevention*. Geneva, 29-31 March 1999.

Osóiro, F. L., Salum, G. A., Donadon, M. F., Forni-dos-Santos, L., Loureiro, S. R. & Crippa, J. A. S. (2013). Psychometrics properties of Early Trauma Inventory Self Report- Short Form (ETISR-SR) for the Brazilian Context. *PLoS ONE, 8*(10), e76337. doi:10.1371/journal.pone.0076337

Palmer, R. L., Chaloner, D. A. & Oppenheimer, R. (1992). Childhood sexual abuse with adults reported by female psychiatric patients. *British Journal of Psychiatry, 160*, 261-265.

Paris, J., Zweig-Frank, H. & Guzder, J. (1994). Risk factors for borderline personality in male outpatients. *Journal of Nervous and Mental Disease, 182,* 375-380.

Park, S. G., Bennett, M. E., Couture, S. M. & Blanchard, J. J. (2013). Internalized stigma in schizophrenia: relations with dysfunctional attitudes, symptoms, and quality of life. *Psychiatry Research, 205*(1-2), 43-47. doi: 10.1016/j.psychres. 2012.08.040

Pennebaker, J. W. & Susman, J. R. (1988). Disclosure of traumas and psychosomatic processes. *Social Science and Medicine, 26,* 327-332.

Pereda, N. (2009). Consecuencias psicológicas iniciales del abuso sexual infantil [Short-term psychological consequences of child sexual abuse]. *Papeles del Psicólogo, 30*(2), 135-144.

Pereda, N. (2010). Consecuencias psicológicas a largo plazo del abuso sexual infantil [Long-term psychological consequences of child sexual abuse]. *Papeles del Psicólogo, 31*(2), 191-201.

Plaza, A., Torres, A., Martin-Santos, R., Gelabert, E., Imaz, M. L., Navarro, P., … Garcia-Esteve, L. (2011). Validation and Test-Retest Reliability of Early Trauma Inventory in Spanish Postpartum Women. *The Journal of Nervous and Mental Disease, 199*(4), 280-285. doi:10.1097/nmd.0b013e31821245b9

Riggs, D. S., O'Leary, K. D. & Breslin, F. C. (1990). Multiple correlates of physical aggression in dating couples. *Journal of Interpersonal Violence, 5*, 61-73.

Roy, C. A. & Perry, C. (2004). Instruments for the assessment of childhood trauma in adults. *Journal of Nervous and Mental Disease, 192*(5), 343-351.

Russ, M. J., Shearin, E. N., Clarkin, J. F., Harrison, K. & Hull, J. W. (1993). Subtypes of self-injurious patients with borderline personality disorder. *American Journal of Psychiatry, 150,* 1869-1871.

Russell, D. E. H. (1986). *The secret trauma: Incest in the lives of girls and women.* New

York: Basic Books.

Ryan, S. W., Rodríguez, N., Rowan, A. & Foy, D. W. (1992). Psychometric analysis of the Sexual Abuse Exposure Questionnaire (SAEQ). In *Proceedings of the100th Annual Convention of the American Psychological Association*. Washington, DC: American Psychological Association.

Sala, R., Goldstein, B. I., Wang, S. & Blanco, C. (2014). Childhood maltreatment and the course of bipolar disorders among adults: Epidemiologic evidence of doseresponse effects. *Journal of Affective Disorders, 165,* 74-80. doi: 10.1016/ j.jad.2014.04.035.

Sanders, B. & Becker-Lausen, E. (1995). The measurement of psychological maltreatment: Early data on the child abuse and trauma scale. *Child Abuse & Neglect, 19*, 315-323.

Sanmartín, J. (2011). *Maltrato Infantil en la familia de España [Child Abuse in the family in Spain].* Madrid: Ministry of Health, Social Policy and Equality.

Schmeiser, C. B. & Welch, C. (2006). Test development. In R. L. Brennan (Ed.), *Educational Measurement (4th ed.)* (pp. 307-353). Westport, CT: American Council on Education/Praeger.

Schäfer, I., Fisher, H. L. Aderhold, V., Huber, B., Hoffmann-Langer, L., Golks, D., …Harfst, T. (2012). Dissociative symptoms in patients with schizophrenia: Relationships with childhood trauma and psychotic symptoms. *Comprehensive Psychiatry, 53*(4), 364-371. doi: 10.1016/j.comppsych.2011. 05. 010

Shenk, C. E., Putnam, F. W. & Noll, J. G. (2013). Predicting the accuracy of facial affect recognition: The interaction of child maltreatment and intellectual functioning. *Journal of Experimental Child Psychology, 114*(2), 229-242. doi: 10.1016/j.jecp.2012.08.007

Shonkoff, J. P., Garner, A. S., Siegel, B. S., Dobbins, M. I., Earls, M. F., Garner, A. S., … Wood, D. L. (2011). The Lifelong Effects of Early Childhood Adversity and Toxic Stress. *Pediatrics, 129*(1), e232–e246. doi:10.1542/peds.2011-2663

Silk, K. R., Lee, S., Hill, E. M. y Lohr, N. E. (1995). Borderline personality disorder symptoms and severity of sexual abuse. *American Journal of Psychiatry, 152*, 1059-1064.

Simón, C., López, J. L. & Linaza, J. L. (2000). *Maltrato y desarrollo infantil [Child abuse and development].* Madrid: Universidad Pontificia de Comillas.

Soloff, P. H., Lynch, K. G. & Kelly, T. M. (2002). Childhood abuse as a risk factor for suicidal behavior in borderline personality disorder. *The Journal of Personality Disorders, 16,* 201-214.

Subica, A. M. (2013).Psychiatric and physical sequelae of childhood physical and sexual abuse and forced sexual trauma among individuals with serious mental illness. *Journal of Traumatic Stress, 26*(5), 588-596. doi: 10.1002/jts.21845

Teicher, M. H. & Smamson, J. A. (2013). Childhood maltreatment and psychopathology: A case for ecophenotypic variants as clinically and neurobiologically distinct subtypes. *The American Journal of Psychiatry, 170*(10), 1114-1133. doi: 10.1176/appi.ajp.2013.12070957

Thabrew, H., de Sylva, S. & Romans, S. E. (2012). Evaluating childhood adversity. *Advances in psychosomatic medicine, 32,* 35-57.doi: 10.1159/000330002

Thompson, A. D., Nelson, B., Yuen, H. P., Lin, A., McGorry, P. D., Wood, S. J. & Yung, A. R. (2014). Sexual trauma increases the risk of developing psychosis in an ultra high-risk "prodromal" population. *Schizophrenia Bulletin, 40*(3), 697-706. doi: 10.1093/schbul/sbt032

Tikka, M., Luutonen, S., Ilonen, T., Tuominen, L., Kotimäki, M., Hankala, J. & Salokangas, R. K. (2013). Childhood trauma and premorbid adjustment among individuals at clinical high risk for psychosis and normal control subjects. *Early Intervention in Psychiatric, 7*, 51-57. doi:10.1111/j.1751-7893.2012. 00391.x

Van Nierop, M., Viechtbauer, W., Gunther, N., van Zelst, C., de Graaf, R., ten Have, M., … van Winkel, R. (2014). Childhood trauma is associated with a specific admixture of affective, anxiety, and psychosis symptoms cutting across traditional diagnostic boundaries. *Psychological Medicine, 45*(6), 1277–1288. doi:10.1017/s0033291714002372

Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., … Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: a metaanalysis of patient-control, prospective- and cross-sectional cohort studies. *Schizophrenia Bulletin, 38,* 661-671. doi: 10.1093/schbul/sbs050

Viezel, K. D., Freer, B. D., Lowell, A. & Castillo, J. A. (2014). Cognitive abilities of maltreated children. *Psychology in the Schools, 52*(1), 92-106. doi:10.1002/pits.21809

Wagner, A. W. & Linehan, M. M. (1994). Relationship between childhood sexual abuse and topography of parasuicide among women with borderline personality disorder. *Journal of Personality Disorder, 8*, 1-9.

Wilson, M. (2005). *Constructing measures: An item response modeling approach.* Mahwah, NJ: Lawrence Erlbaum Associates.

Young, J. C. & Widom, C. S. (2014). Long-term effects of child abuse and neglect on emotion processing in adulthood. *Child abuse & Neglect, 38*(8), 1369-1381.doi: 10.1016/j.chiabu.2014.03.008

Zanarini, M. C., Gunderson, J. G., Marino, M. F., Schwartz, E. O. & Frankenburg, F. R. (1989). Childhood experiences of borderline patients. *Comprehensive Psychiatry, 30,* 18-25

Papeles del Psicólogo, 2016. Vol. 37(1), pp. 45-51
http://www.papelesdelpsicologo.es

PSYCHOLOGICAL ASSESSMENT OF OPIOID DRUG ABUSE

José Luis Carballo, Ainhoa Coloma-Carmona, Dana Mrozowicz-Gaudyn, Verónica Vidal-Arenas, Carlos van-der Hofstadt and Jesús Rodríguez-Marín
Universidad Miguel Hernández de Elche

El incremento de la preinscripción de fármacos analgésicos opioides se ha asociado a un aumento de las tasas de abuso y las consecuencias negativas asociadas a su uso inadecuado. Diversos organismos sanitarios internacionales recomiendan la realización de una evaluación global y multidisciplinar del paciente durante todo el proceso terapéutico con opiáceos, con el fin de identificar un posible abuso. Ante la ausencia de guías específicas en el ámbito sanitario español, el objetivo de este artículo ha sido presentar una propuesta de evaluación psicológica, en base a los principales instrumentos de evaluación, disponibles actualmente para evaluar el abuso de opioides. Se establecen pautas para el proceso de evaluación en función de las variables psicológicas que puedan predecir y mantener dicho abuso, clasificando todo ello en función del momento del proceso terapéutico en el que los pacientes se encuentren. Aunque existen instrumentos con buenas propiedades psicométricas, son necesarias futuras investigaciones para la adaptación, traducción y validación de éstos a población española. Al mismo tiempo son necesarios futuros estudios que profundicen en estrategias de prevención e intervención para reducir la probabilidad de abuso en pacientes tratados con fármacos opioides.
Palabras clave: *Evaluación psicológica, Abuso de fármacos, Opioides, Dolor crónico, Factores de riesgo, Guías clínicas.*

The increase in the prescription of opioid analgesics is related to increased rates of opioid abuse and the negative consequences of medication misuse. Several international health organisations recommend comprehensive and multidisciplinary patient assessment for the duration of the opioid treatment in order to identify and prevent medication abuse. Due to the lack of specific clinical guidelines in the Spanish National Health System, the aim of this paper is to present a proposal for psychological assessment based on the main psychological tools currently available for assessing opioid abuse. The assessment guidelines have been established based on the psychological variables that can predict and prolong the abuse, classifying all of the variables depending on the current stage of the therapeutic process for each patient. Although there are instruments with good psychometric properties, further research is necessary to adapt, translate and validate these instruments for use in the Spanish population. Future studies are also needed to investigate intervention and prevention strategies in depth in order to reduce the likelihood of abuse in patients treated with opioids.
Key words: *Psychological assessment, Prescription drug abuse, Opioids, Chronic pain, Risk factors, Clinical guidelines.*

The inappropriate use of psychotropic drugs is a public health problem worldwide. In 2012, the National Survey on Drug Use and Health (SAMHSA, 2013b) reported that 2.6% of the general population had consumed psychoactive drugs without a prescription in the last month. This figure rose to 5.3% for people between 18 and 25 years of age. Regarding the situation in Spain, the Spanish Observatory of Drugs and Drug Addiction (OEDT, 2011) has detected an increase in the use of hypnotic drugs during the last 30 days in the general population rising from 3.7% in 2005, to 5.2% in 2009; also, around 2% of the surveyed population had consumed hypnotic drugs without a prescription during the past year.

Special attention should be paid to analgesic opioids, given their highly addictive power (Manchikanti et al., 2012) and the considerable increase in their prescription over the last two decades, both in Spain (Garcia del Pozo, Carvajal, Viloria, Velasco & Garcia del Pozo, 2008) and in the rest of the world (Dhalla et al, 2009; Edlund et al, 2010; Gomes et al, 2011; Leong, Murnion, & Haber, 2009), largely due to the fact that their prescription has established itself as the treatment of choice for patients with medium-high levels of chronic pain (Liebschütz, Beers & Lange, 2014). Associated with this increase in prescriptions of opioid drugs, there has been an increase in the rates of abuse (Atluri, Sudarshan, & Manchikanti, 2014; Turk, Swanson, & Gatchel, 2008); although there are few data on its prevalence, some studies indicate abuse rates between 20 and 24% of people receiving this treatment (Sullivan et al., 2010).

All of this is associated with an increase in the number of negative consequences associated with the inappropriate use of opioid drugs. In the United States, the rate of overdose deaths from opioid analgesics has tripled since 1999, to the point that, since 2003, there have been more deaths from overdoses related to these drugs than heroin and cocaine together (Centers for Disease Control and Prevention, 2011, 2013). Every year more than one million people visit the emergency department for problems related to drug misuse, defined as taking more than the prescribed dose, consuming drugs prescribed for someone else, voluntary poisoning or documented drug abuse. Of these emergency room visits, almost 40% are related to opioid analgesics, a percentage that represents almost half a million people each year (SAMHSA, 2013a).

In addition, the rate of admission to treatment for the abuse of opioid drugs has also skyrocketed, increasing every year since 2001 and reaching a 300% increase since then (SAMHSA, 2013c).

This situation highlights the need for strategies to identify the abuse of opioid drugs in patients who receive them. The main health institutions

Correspondencia: José Luis Carballo. *Universidad Miguel Hernández de Elche. Avenida de la Universidad s/n. 03202 Elche. España. E-mail: jcarballo@umh.es*

in the field of addiction and pain, the American Pain Society (APS), the American Academy of Pain Medicine (AAPM), the Grading of Recommendations Assessment (GRADE), the National Institute on Drug Abuse (NIDA) and the National Institute of Mental Health (NIMH) as well as the World Health Organization (WHO) emphasise the importance of a comprehensive and multidisciplinary patient assessment for the duration of the therapeutic process with opioids (Chou, 2009).

The development and use of valid and reliable assessment tools is not only useful for identifying cases of abuse but also for planning preventive strategies and specific treatments for addiction to opioids (Chang & Compton, 2013). In the absence of specific guidelines for the Spanish population, the present study aims to present a proposal for psychological assessment addressing the main strategies and psychological tools currently available to assess the abuse of opioid drugs, as well as the psychological variables that predict and maintain it. To this end, a narrative review was conducted of the clinical guidelines for instruments of detection and assessment of the Opioid Risk project (funded by the National Institute on Drug Abuse) and the clinical guidelines of the American Pain Society (APS) and the American Academy of Pain Medicine (AAPM) for the use of opioid therapy in patients with chronic non-cancer pain. From these guides, the instruments with the best psychometric properties in terms of reliability and validity were selected, as well as the most used ones according to the guidelines themselves, after reviewing the original articles of each of these assessment tools.

THE PSYCHOLOGICAL ASSESSMENT OF ADDICTION TO OPIOID DRUGS

The psychological assessment of opioid addiction and the factors that may be involved in it is useful for health care settings (e.g., pain units of hospitals) where patients with chronic pain problems are treated by medical specialists in order to improve their adaptation to their daily activities and to improve their quality of life (Chang & Compton, 2013). These assessment procedures have several purposes within these health care settings: (1) on the one hand, they are aimed at identifying those patients who may be abusing and/or dependent on these drugs, (2) on the other hand, the intention is to examine the medical, psychological and social factors that can predict the risk that a person may develop an addictive process, (3) in line with the previous purpose, in cases where the probability of developing an addiction is high, this assessment would enable alternative pain management interventions to be sought, (4) to develop specific preventive strategies to reduce the likelihood of abuse and/or dependence appearing, (5) to plan guidelines for opioid drug use (e.g., drug dosage and route of administration) according to the patient's risk of developing an addiction, and (6) finally, the aim is to plan interventions in cases where addiction appears, based on the characteristics and circumstances of each patient.

General considerations of the assessment

This type of clinical assessment has certain peculiarities, common to the assessment of addictive behaviours, which should be evaluated at the time of carrying out the assessment, such as: (1) that the patient is under the influence of the drug when being assessed, which will impact the validity of the results, (2) low motivation to change and to recognise or identify that they may be using the drugs inappropriately; it could be contradictory for the patient to consider that something that is "alleviating" them and has been prescribed by a specialist can create an addiction, and (3) finally, as mentioned above, the recommendations of healthcare organisations emphasise the need for the assessment to be multidisciplinary since there are many factors involved in the risk of abuse and the consequences that these may entail can affect many areas of the patient's life.

CLASSIFICATION OF THE ASSESSMENT INSTRUMENTS

The international assessment guidelines emphasise the importance of the assessment being made throughout the whole of the therapeutic process and the importance is stressed of differentiating between two moments of assessment, with specific methods and procedures (see Table 1): the Initial Assessment (before starting to use the drug) and the Control Assessment (after starting to use it).

Initial Assessment

When considering starting treatment with opioid drugs, the risk of abuse or their problematic use should be evaluated carefully, in order to be able to identify the probability that the patient has of developing these issues (Passik, 2009). This is why conducting an initial assessment is essential, since the definition of alternative treatments for pain, in the case of likelihood of abuse, is essential and necessary for an appropriate intervention and for preventing the abuse of opioids (Chou et al., 2009). Thus, this assessment would facilitate the establishing of a prescription drug appropriate to the patient's condition, limiting, for example, the dose and the maximum duration of the prescription, as well as selecting the most appropriate drug for each case (Thorson et al., 2014).

To carry out the initial assessment, it would be appropriate to assess the following aspects:

Assessment of socio-demographic characteristics and general state of health

Different studies show differences in the consumption of psychoactive drugs according to socio-demographic variables such as sex, age, type of family life or employment status. These studies suggest a higher prevalence of abuse of psychotropic drugs in women and at older ages, as well as in people with a low educational level who are unemployed and living alone (Secades Villa et al., 2003).

The sociodemographic variables can be assessed by administering questionnaires and/or clinical interviews to collect data on age, sex, marital status, employment status, educational and economic level.

In relation to the assessment of the general state of health, given the multidisciplinary nature of this assessment, the psychologist must have information of the state of health assessment contained in the patient's clinical record.

Assessment of consumption of psychoactive substances and opioid drugs prior to treatment

The history of personal and family substance abuse appears to be significantly related to the risk of abuse of opiates in pain patients (Chou et al, 2009; Matteliano, St Marie, Oliver, & Coggins, 2012; Sehgal,

Manchikanti, & Smith, 2012). Therefore it is important, before initiating treatment with opioid drugs, to evaluate specifically, the possible existence of substance abuse and to intervene, if abuse is detected, at the same time as treating the pain with opioids (Passik, Kirsh, & Casper, 2008). The following are some of the most used instruments for the evaluation of psychoactive substance consumption, due to their simplicity and good psychometric properties:

On the one hand, screening instruments such as the *Alcohol Use Disorders Identification Test* (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993) in its Spanish version (Rubio et al., 1998) and the *Drug Abuse Screening Test* (DAST-10; Bohn, Babor & Kranzler, 1991) for illegal drugs, which also has a Spanish version (Gálvez et al., 2010).

The Screening Instrument for Substance Abuse Potential (SISAP; Coambs, Larry, Santhiapillai, Abrahamsohn et al, 1996) could also be used. This is an interview composed of 5 items, through which both the history of substance abuse and the risk of dependence or abuse of opioid drugs are assessed. This interview is recommended by the guidelines for the assessment of abuse of opioid drugs, and shows good sensitivity and specificity, 0.91 and 0.78 respectively (Coambs et al. 1996).

All of these self-reports could be accompanied by assessments by family members or individuals living with the patient, as well as biochemical tests for detecting use (e.g., markers in urine samples) in the event that the presence of consumption of one or several psychoactive substances is suspected, given the high risk of developing abuse behaviours of opioid drugs in polydrug-using patients.

Evaluation of other psychosocial variables related to the risk of abuse

A personal and family history of abuse of alcohol and other drugs, along with a personal history of physical and sexual abuse and the presence of psychiatric disorders are the main risk factors identified for the abuse of psychotropic drugs (Chou, et al., 2009; Matteliano, et al, 2012; Sehgal et al, 2012). SAMHSA (2012) determined the risk in terms of these variables, classified as low (e.g., no history of substance abuse); medium (e.g., having a personal and family history of substance abuse) and high (e.g., presenting current substance abuse and a history of previous abuse of opioid drugs).

The questionnaire *Opioid Risk Tool* (ORT; Webster & Webster, 2005), developed specifically for pain patients, enables us to assess the risk of abuse of psychotropic drugs. It is a self-report composed of 5 items in which the following dimensions are included: personal and family history in relation to substance abuse, age, episodes of sexual abuse in preadolescence and presence of psychological disorders. The higher the score, the greater the risk, which can be classified as follows: 0-3 points (low risk), 4-7 points (moderate risk) and more than 8 points (high risk). This instrument provides excellent discrimination between patients with high and low risk, and between men and women in the analyses showing a capacity of 90.9% for predicting abuse of opioid drugs in high risk patients and 94.4% for predicting no abuse in patients with low risk (Webster & Webster, 2005).

On the other hand, the *Screener and Opioid Assessment for Patients With Pain - Revised* (SOAPP-R; Butler et al, 2007), is a self-report developed specifically to predict the abuse of psychotropic drugs in pain patients (Butler, Fernandez, Benoit, Budman, & Jamison, 2008). It consists of 24 items with a Likert scale, which ranges from O (never) to 4 (very often). The dimensions evaluated are as follows: history of consumption of alcohol or other substances, psychological state and stress. The higher the score, the greater the risk of abuse of psychotropic drugs. The SOAPP-R is the only questionnaire of this type that has undergone cross-validation. The test-retest reliability analysis shows an

TABLE 1
PROPOSED MULTIDIMENSIONAL ASSESSMENT OF OPIOID DRUG ABUSE

Dimension	Instrument	Evaluación Entrada	Evaluación Control
Abuse of other psychoactive substances	*Alcohol Use Disorders Identification Test* (AUDIT) (Rubio, Bermejo, Caballero, & Santo-Domingo, 1998)[a]	X	
	Drug Abuse Screening Test (DAST-10)(Gálvez, Fernández, Manzanaro, Valenzuela, & Lafuente, 2010)[a]	X	
	Screening Instrument for Substance Abuse Potential (SISAP) (Coambs et al., 1996)[b]	X	
Other risk factors for abuse of opioid drugs	*Opioid Risk Tool* (ORT)(Webster & Webster, 2005)[c]	X	
	Screener and Opioids Assessment for Patients with Pain – Revised (SOAPP-R)(Butler et al., 2008)[b]	X	
Perceived pain	*Brief Pain Questionnaire (BPQ)*(Llach et al., 2003)[a]	X	X
	Visual Analogue Scale (VAS)[a]	X	X
Psychological state	*Hospital Anxiety and Depression Scale (HADS)* (Quintana et al., 2003)[a]	X	X
	Symptom Checklist-90-Revised (SCL-90-R) (Vallejo, Jordán, Díaz, Comeche, & Ortega, 2007)[a]	X	X
Compliance with the prescriptions of the opioid treatment	*Pain Assessment and Documentation Tool* (PADT)(Passik et al., 2004)[b]		X
Use and abuse of opioid drugs	*Prescription Opioid Misuse Index* (POMI)(Knisely et al., 2008)[d]		X
	Current Opioid Misuse Measure (COMM)(Butler et al., 2007)[b]		X

[a] This is a version that has been adapted and translated for the Spanish population.
[b] There is not a version that has been adapted and translated for the Spanish population.
[c] ORT: There is a Spanish translation. Further information can be requested from the authors of this manuscript on the adaptation and translation of the instrument, as authorisation has been obtained from the authors of the instrument. See: http://www.lynnwebstermd.com/risk-tool-download/
[d] POMI: There is a Spanish translation. Further information can be requested from the authors of this manuscript on the adaptation and translation of this instrument, as authorisation has been obtained from the authors of the instrument.

intra-class index of ICC = 0.94 (CI 95%: 0.90 - 0.97) with an alpha of 0.86, indicating a very good reliability. In addition, it has a sensitivity and specificity of 79% and 52%, respectively (Butler, Budman, Fernandez, Fanciullo, & Jamison, 2009).

Pain assessment and impact on daily activities

Adequate pain assessment enables us to evaluate both the effectiveness of treatment and possible adverse effects thereof (Ibáñez, Morales, Calleja, Moreno & Gálvez, 2001). Thus, if a patient requests a dose increase, through this evaluation it is possible to determine whether the pain has increased, there is tolerance or other effects are being sought, such as sedation or anxiety reduction (Center for Substance Abuse Treatment, 2012). The most commonly used instruments for this evaluation are:

The Spanish version (Llach et al., 2003) of the *Brief Pain Questionnaire* (BPQ; Cleeland & Ryan, 1994), which is a self-administered questionnaire that includes two dimensions: the intensity of the pain and its impact on the patient's daily activities through 9 items using Likert scales and dichotomous answers (yes / no). The reliability analysis shows a Cronbach's alpha greater than 0.70 for each of the dimensions (Badia et al., 2003).

The *Visual Analogue Scale* (VAS), graded numerically, is a unidimensional scale that measures the severity of the pain, representing the subjective feeling of the patient in numbers. It is a 100mm horizontal line ranging from "no pain" to "the worst pain imaginable" in which the patient must indicate the intensity of their pain. This tool allows us to compare pain scores at different times. It is a simple instrument to use, which has shown good properties of test-retest reliability of r = 0.947 and intra-class index ICC = 0.97 (Grupo Valoración, 2009).

Assessment of psychological state

Psychopathological comorbidity is one of the main risk factors in the development of abuse of opioid drugs (Chou, 2009) and therefore it requires a specific assessment. Given the hospital setting, where these patients generally receive care, the validated Spanish version (Quintana et al., 2003) of the *Hospital Anxiety and Depression Scale* (HADS; Zigmond & Snaith, 1983), could be used to assess symptoms of anxiety and depression. It has two subscales, each consisting of 7 items that are valued from 0 to 3 and a score of higher than 10 is considered indicative of morbidity. The scale has high internal consistency, with a Cronbach's alpha of 0.86 and 0.86, for the scales of anxiety and depression respectively; and high test-retest reliability, with a correlation coefficient above 0.85 (Quintana et al., 2003).

Other instruments which can be used, in addition to the clinical history of psychological and psychiatric treatment, for a brief, general assessment of psychopathological symptoms include, for example, the *Symptom Checklist - Revised* (SCL-90-R; Derogatis, 1975) in its Spanish version (Vallejo, Jordán, Díaz, Comeche, & Ortega, 2007). This is a self-report questionnaire of 90 items with a Likert scale, which assesses psychological symptoms and distress.

Control Assessment

Once the treatment with opioids for pain management has started, patients require periodic inspection and monitoring to determine and ensure compliance with the guidelines set by the specialist doctor, so as to ensure the effectiveness of the treatment, and identify and reduce the potential risk of abuse (Morgan et al, 2013; Sehgal et al, 2012).

In this sense, the Control Assessment involves the continuous monitoring of the response to treatment with opioids and the current use of the drug. Within this assessment, the following aspects should be taken into account:

Assessment of response to treatment with opioids

It is recommended that, on one hand, the presence of side effects and symptoms of tolerance to the drug should be recorded, and on the other hand, the current use of the drug (e.g., the number and frequency of doses) and the degree of perceived pain as well as the functional capacity of patients (Chou, 2009; Passik et al, 2004). In this line, other factors that may be interfering with treatment response would also be evaluated (Morasco, Duckart, & Dobscha, 2011; Sehgal et al, 2012; Sullivan et al., 2010), such as the consumption of drugs without medical supervision or the parallel use of alternatives for pain reduction (e.g., the use of medicinal plants and physiotherapy). Some of the tools that are used include:

The *Pain Assessment and Documentation Tool* (PADT; Passik et al., 2004). This is a structured clinical interview (or set of notes), lasting about 10 minutes, consisting of 41 items to be completed by the psychologist with the help of specialist doctors, which assesses the progress of the patient during long-term treatment with opioids, based on four dimensions: 1) analgesia or perceived pain, 2) the patient's functional capacity (e.g., mood or social and family relationships), 3) side effects of the treatment (e.g., nausea, vomiting or constipation) and 4) presence of risk behaviours of abuse (e.g., excessive sedation, reports of lost or stolen prescriptions). Also, at the end of the questionnaire, there is a section aimed at performing a clinical assessment of the treatment benefit for the patient. The reliability analysis indicates good internal consistency of 0.86 and good interrater reliability (Passik et al., 2004b).

Assessment of the use of the opioid drug

Different assessment strategies are proposed, such as using self-reports, which are a significant source of information for behavioural assessment of patients, collecting information on the proper use of the drug (e.g., dose, route, frequency of administration and circumstances surrounding its use). In addition, it may be useful for the clinician and the patient to identify higher risk situations, where it is more likely that the subject will consume skipping the established patterns (e.g., the time of day or where it is taken).

In this regard, and in order to corroborate the information recorded by the patient, other methods of evaluation can be used, such as:

On the one hand, conducting interviews with family members or caregivers providing information regarding the patient's functional capacity and help in identifying problematic behaviours related to drug use (e.g., asking for help to obtain or borrow more medication). And, on the other, the use of biochemical markers (e.g., in urine), which are recommended and are especially important for high-risk patients and those who are suspected of drug misuse (Chou, 2009).

Assessment of abuse and/or dependence on opioid drugs

In recent decades, self-reports have been developed to specifically assess the abuse of opioid drugs, among which are the following:

The *Prescription Opioid Misuse Index* (POMI; Knisely, Wunsch, Cropsey, & Campbell, 2008b) is a clinical interview composed of 6 items of dichotomous response (Yes / No), which records aspects such as the characteristics of drug use (dose, frequency of consumption), the need to shorten the time between doses, or the feeling of euphoria and/or pleasure after taking the drug. The POMI is a sensitive and specific instrument for identifying patients who misuse opioid drugs (score > 1). The reliability analysis indicates good internal consistency, with an alpha equal to 0.85, and presenting sensitivity and specificity of 82% and 92.3%, respectively (Knisely et al., 2008). (Further information may be requested from the authors of this paper regarding the adaptation and translation of this instrument, as authorisation has been obtained from the authors).

In the same vein, the *Current Opioid Misuse Measure* (COMM; Butler et al, 2007) is a Likert-type scale consisting of 17 items, specifically aimed at the population experiencing chronic pain, whereby the problematic use of the drug is evaluated, taking into consideration the following dimensions: 1) signs or symptoms of problematic use of the psychotropic drug, 2) emotional/psychiatric problems, 3) failure to follow medical guidelines, 4) use of the opioid drug and 5) problematic use of the psychotropic drug. Obtaining a score equal to or greater than nine (≥ 9) identifies patients that are at a high risk of presenting a pattern of problematic use or abuse of opioids, presenting sensitivity and specificity of 77% and 66%, respectively (Butler et al., 2007; Chou et al, 2009). The analysis indicates a very good internal consistency (alpha = 0.86) and high test-retest reliability with ICC = 0.86 (CI 95%: 0.77 to 0.92).

CONCLUSIONS

The aim of this paper was to present an assessment proposal covering the psychological strategies and tools currently available to assess the abuse of opioid drugs, as well as the psychological variables that can predict and maintain it. Following the recommendations of the international guidelines, the assessment instruments have been classified and described according to the moment where the patient currently is in the therapeutic process, thus establishing an Initial and Control Assessment.

This article has presented a set of tools that may be useful for health professionals, especially psychologists working in health care settings with non-cancer population suffering from chronic pain. It seeks to go a step further in improving interventions with opioid drugs, and although there is still a lack of studies regarding their efficacy in the long term, they show very good results in patients' short and medium term adaptation to daily activities (Chang & Compton, 2013). Therefore, given the addictive power of these drugs, it is necessary to prevent their inappropriate use, in cases where it is considered that they will be effective. It is necessary to assess patients in a multidimensional and multidisciplinary way before starting to use these drugs and during interventions with them (Manchikanti et al., 2012).

In this sense, faced with the increased prevalence of opioid drug abuse worldwide and the resulting health consequences, both SAMHSA (2013b) and NIDA (2012) state that is necessary to carry out studies on the efficacy of multicomponent psychological interventions to help to reduce the likelihood of addiction to opioid drugs, and this is certainly not possible without comprehensive and personalised assessments for each patient.

Although, throughout this article, specific instruments to measure addiction to opioid drugs have been presented which also have good psychometric properties in the Spanish population, there is little evidence of assessment instruments for this problem. Further research is needed both in the area of the adaptation, translation and validation of this set of tools to the Spanish population, as well as research that addresses and analyses the efficacy and efficiency of intervention strategies to reduce the likelihood of abuse among patients with chronic pain that receive pharmacological treatment with opioids.

To conclude, we make some final reflections considering this proposed psychological assessment of the abuse of opioid drugs: (1) we stress the need for the assessment to be multidisciplinary, with the participation of all health professionals involved in the treatment of non-cancer pain; (2) the evaluation procedures used in both the Initial and Control Assessment would be the collection of self-reports, biochemical samples and the application of self-reports (pencil and paper interviews and instruments); (3) it could be carried out in two assessment sessions lasting 30 to 40 minutes; in any case, it is important to track usage following the prescriptions of the opioid drug; (4) finally, it should be noted that this type of assessment generally has a low cost and is less intrusive, given the nature of the procedures applied, and at the same time, it can be a great benefit to patients' health, as it may prevent the development of abuse and dependence on opioid drugs by identifying risk factors and problematic consumption.

ACKNOWLEDGEMENTS

This work was funded by the call for aid for conducting research to improve the care of the chronically ill complex patient and those susceptible to palliative care (2013) of the Department of Health of Valencia (ref. PCC18 / 13).

REFERENCES

Atluri, S., Sudarshan, G., & Manchikanti, L. (2014). Assessment of the trends in medical use and misuse of opioid analgesics from 2004 to 2011. *Pain physician*, 17(2), E119-128.

Badia, X., Muriel, C., Gracia, A., NunezOlarte, J. M., Perulero, N., Galvez, R., Carulla, J., et al. (2003). [Validation of the Spanish version of the Brief Pain Inventory in patients with oncological pain]. *Medicina Clínica*, 120(2), 52-59.

Barry, D. T., Irwin, K. S., Jones, E. S., Becker, W. C., Tetrault, J. M., Sullivan, L. E., Hansen, H., et al. (2010). Opioids, Chronic Pain, and Addiction in Primary Care. *The Journal of Pain: Official Journal of the American Pain Society*, 11(12), 1442-1450.

Butler, S. F., Budman, S. H., Fernandez, K. C., Fanciullo, G. J., & Jamison, R. N. (2009). Cross-Validation of a Screener to Predict Opioid Misuse in Chronic Pain Patients (SOAPPR). *Journal of Addiction Medicine*, 3(2), 66-73.

Butler, S. F., Budman, S. H., Fernandez, K. C., Houle, B., Benoit, C., Katz, N., & Jamison, R. N. (2007). Development and Validation of

the Current Opioid Misuse Measure. *Pain*, 130(1-2), 144-156.

Butler, S. F., Fernandez, K., Benoit, C., Budman, S. H., & Jamison, R. N. (2008). Validation of the revised Screener and Opioid Assessment for Patients with Pain (SOAPPR). *The Journal of Pain: Official Journal of the American Pain Society*, 9(4), 360.

Centers for Disease Control and Prevention. (2011). Vital signs: overdoses of prescription opioid pain relievers - United States, 1999-2008. *MMWR. Morbidity and Mortality Weekly Report, 60*(43), 1487-1492.

Centers for Disease Control and Prevention. (2013). Vital signs: overdoses of prescription opioid pain relievers and other drugs among women United States, 1999-2010. *MMWR. Morbidity and Mortality Weekly Report, 62*(26), 537542.

Chang, Y.P., & Compton, P. (2013). Management of chronic pain with chronic opioid therapy in patients with substance use disorders. *Addiction Science & Clinical Practice, 8*(1), 21.

Chou, R. (2009). 2009 Clinical Guidelines from the American Pain Society and the American Academy of Pain Medicine on the use of chronic opioid therapy in chronic noncancer pain: what are the key messages for Clinical practice? *Polskie Archiwum Medycyny Wewn trznej, 119*(7-8), 469-477.

Chou, R., Fanciullo, G. J., Fine, P. G., Adler, J. A., Ballantyne, J. C., Davies, P., Donovan, M. I., et al. (2009). Clinical guidelines for the use of chronic opioid therapy in chronic noncancer pain. *The Journal of Pain: Official Journal of the American Pain Society, 10*(2), 113-130.

Chou, R., Fanciullo, G. J., Fine, P. G., Miaskowski, C., Passik, S. D., & Portenoy, R. K. (2009). Opioids for Chronic Noncancer Pain: Prediction and Identification of Aberrant Drug-Related Behaviors: A Review of the Evidence for an American Pain Society and American Academy of Pain Medicine Clinical Practice Guideline. *The Journal of Pain, 10*(2), 131-146.e5.

Cleeland, C. S., & Ryan, K. M. (1994). Pain assessment: global use of the Brief Pain Inventory. *Annals of the Academy of Medicine, Singapore*, 23(2), 129-138.

Dhalla, 1. A., Mamdani, M. M., Sivilotti, M. L. A., Kopp, A., Qureshi, O., & Juurlink, D. N. (2009). Prescribing of opioid analgesics and related mortality before and after the introduction of longacting oxycodone. *Canadian Medical Association Journal, 181*(12), 891-896.

Edlund, M. J., Martin, B. C., Fan, M.Y., Braden, J. B., Devries, A., & Sullivan, M. D. (2010). An analysis of heavy utilizers of opioids for chronic noncancer pain in the TROUP study. *Journal of Pain and Symptom Management, 40*(2), 279-289.

Gálvez, B. P., Fernández, L. G., Manzanaro, M. P. de V., Valenzuela, M. A. O., & Lafuente, M. L. (2010). Validación española del Drug Abuse Screening Test (DAST-20 y DAST-10) [Spanish validation of the Drug Abuse Screening Test (DAST-20 and DAST-10)]. *Health and Addictions/Salud y Drogas, 10*(1).

Garcia del Pozo, J., Carvajal, A., Viloria, J. M., Velasco, A., & Garcia del Pozo, V. (2008). Trends in the consumption of opioid analgesics in Spain. Higher increases as fentanyl replaces morphine. *European Journal of Clinical Pharmacology, 64*(4), 411-415.

Gomes, T., Juurlink, D. N., Dhalla, 1. A., MailisGagnon, A., Paterson, J. M., & Mamdani, M. M. (2011). Trends in opioid use and dosing among socioeconomically disadvantaged patients. *Open Medicine, 5*(1), e13-e22.

Grupo "Valoración". Consejería de Salud y Servicios Sanitarios. Principado de Asturias [Ministry of Health and Healthcare Services, Asturias] (2009). *Proceso de cuidados: Valoración enfermera [Care process: Nurse Rating]*. Asturias: Asturgraf S. L.

Knisely, J. S., Wunsch, M. J., Cropsey, K. L., & Campbell, E. D. (2008). Prescription Opioid Misuse Index: A brief questionnaire to assess misuse. *Journal of Substance Abuse Treatment, 35*(4), 380-386.

Leong, M., Mumion, B., & Haber, P. S. (2009). Examination of opioid prescribing in Australia from 1992 to 2007. *Internal Medicine Journal, 39*(10), 676-681.

Liebschutz, J., Beers, D., & Lange, A. (2014). Managing Chronic Pain in Patients With Opioid Dependence. *Current Treatment Options in Psychiatry, 1*(2), 204-223.

Llach, X. B., Cleeland, C. S., Muriel, C., Gracia, A., Perulero, N., Carulla, J., Olarte, J. M. N., et al. (2003). Validación española del cuestionario Brief Pain Inventory en pacientes con dolor de causa neoplásica [Spanish Validation of the Brief Pain Inventory in pain patients with neoplastic cause]. *Medicina Clínica, 120*(2), 52-59.

Manchikanti, L., Abdi, S., Atluri, S., Balog, C. C., Benyamin, R. M., Boswell, M. V., Brown, K. R., et al. (2012). American Society of Interventional Pain Physicians (ASIPP) guidelines for responsible opioid prescribing in chronic noncancer pain: Part 2guidance. *Pain Physician, 15*(3 Suppl), S67-116.

Matteliano, D., St Marie, B. J., Oliver, J., & Coggins, C. (2012). Adherence Monitoring with Chronic Opioid Therapy for Persistent Pain: A Biopsychosocial Spiritual Approach to Mitigate Risk. *Pain Management Nursing*. Retrieved June 22, 2014, from http://www.sciencedirect.com/science/article/pii/S1524904212001427

Morasco, B. J., Duckart, J. P., & Dobscha, S. K. (2011). Adherence to Clinical Guidelines for Opioid Therapy for Chronic Pain in Patients with Substance Use Disorder. *Journal of General Internal Medicine, 26*(9), 965-971.

Morgan, L., Weaver, M., Sayeed, Z., & Orr, R. (2013). The use of prescription monitoring programs to reduce opioid diversion and improve patient safety. *Journal of Pain & Palliative Care Pharmacotherapy, 27*(1), 49.

National Institute on Drug Abuse. (2012). *Los medicamentos de prescripción: Abuso y Adicción [Prescription medicines: Abuse and addiction]*. Washington, DC: NIH.

OEDT. (2011). *Informe 2011: Situación y tendencias de los problemas de drogas en España [2011 Report: Status and trends of drug problems in Spain]*. Madrid: Ministry of Health, Social Policy and Equality.

Passik, S. D. (2009). Issues in Long-term Opioid Therapy: Unmet Needs, Risks, and Solutions. *Mayo Clinic Proceedings, 84*(7), 593-601.

Passik, S. D., Kirsh, K. L, & Casper, D. (2008). Addiction Related Assessment Tools and Pain Management: Instrumenta for Screening, Treatment Planning, and Monitoring Compliance. Pain Medicine, 9, S145-S166.

Passik, S. D., Kirsh, K. L., Whitcomb, L., Portenoy, R. K., Katz, N. P.,

Kleinman, L., Dodd, S. L., et al. (2004). A new tool to assess and document pain outcomes in chronic pain patients receiving opioid therapy. *Clinical Therapeutics, 26*(4), 552-561.

Quintana, J. M., Padiema, A., Esteban, C., Arostegui, I., Bilbao, A., & Ruiz, I. (2003). Evaluation of the psychometric characteristics of the Spanish version of the Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 107*(3), 216-221.

Rubio, G., Bermejo, J., Caballero, M. C., & SantoDomingo, J. (1998). AUDIT. *Revista Clínica Española, 198*(1), 11-14.

SAMHSA. (2013a). *Results from the 2012 National Survey on Drug Use and Health: Summar), of National Findings* (NSDUH Series H46, HHS Publication No. (SMA) 13-4789.). Rockville, MD: Substances Abuse and Mental Health Services Administration.

SAMHSA. (2013b). *Drug Abuse Warning Network, 2011: National Estimates of Drug — Related Emergency Department Visits* (HHS Publication No. (SMA) 134760, DAWN Series D39.). Rockville, MD: Substances Abuse and Mental Health Services Administration. Retrieved April 23, 2014, from http://jflahiff.wordpress.com/2013/07/19/drug-abuse-warning-network-2011-national-estimates-of-drug-related-emergency-department-visits/

SAMHSA. (2013c). *Treatment Episode Data Set (TEDS): 20012011. State Admissions to Substance Abuse Treatment Services*. (BHSIS Series S-68, HHS Publication No. (SMA) 14-4832.). Rockville, MD: Substances Abuse and Mental Health Services Administration.

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption - II. *Addiction (Abingdon, England), 88*(6), 791-804.

Secades Villa, R., Rodríguez García, E., Valderrey Barbero, J., Fernández Hermida, J. R., Vallejo Seco, G., & Jiménez García, J. M. (2003). El consumo de psicofármacos en pacientes que acuden a Atención Primaria en el Principado de Asturias (España) [The consumption of psychotropic drugs in patients attending primary care in the Principality of Asturias (Spain)]. *Psicothema, 15* (4), 650655.

Sehgal, N., Manchikanti, L., & Smith, H. S. (2012). Prescription opioid abuse in chronic pain: a review of opioid abuse predictors and strategies to curb opioid abuse. *Pain Physician, 15*(3), ES67—ES92.

Substance Abuse and Mental Health Services Administration. (2012). *Managing Chronic Pain in Adults With or in Recovery From Substance Use Disorders*. SAMHSA/CSAT Treatment Improvement Protocols. Rockville (MD): Substance Abuse and Mental Health Services Administration (US).

Sullivan, M. D., Edlund, M. J., Fan, M.Y., DeVries, A., Braden, J. B., & Martin, B. C. (2010). Risks for Possible and Probable Opioid Misuse Among Recipients of Chronic Opioid Therapy in Commercial and Medic aid Insurance Plans: the TROUP Study. *Pain, 150*(2), 332-339.

Sullivan, M. D., & Howe, C. Q. (2013). Opioid therapy for chronic pain in the United States: Promises and perils. *Pain, 154*(0 1), S94-100.

Thorson, D., Biewen, P., Bonte, B., Epstein, H., Haake, B., Hansen, C., Hooten, M., et al. (2014). Acute Pain Assessment and Opioid Prescribing Protocol. Retrieved July 20, 2014, from https://www.icsi.org/_asset/dyp5wm/Opioids.pdf

Turk, D. C., Swanson, K. S., & Gatchel, R. J. (2008). Predicting opioid misuse by chronic pain patients: a systematic review and literature synthesis. *The Clinical Journal of Pain, 24*(6), 497-508.

Upshur, C. C., Luckmann, R. S., & Savageau, J. A. (2006). Primary Care Provider Concerns about Management of Chronic Pain in Community Clinic Populations. *Journal of General Internal Medicine, 21*(6), 652-655.

Vallejo, M. A., Jordán, C. M., Díaz, M. I., Comeche, M. I., & Ortega, J. (2007). Psychological assessment via the internet: a reliability and validity study of online (vs paper-and-pencil) versions of the General Health Questionnaire-28 (GHQ-28) and the Symptoms Check-List-90-Revised (SCL-90-R). *Journal of Medical Internet Research, 9*(1), e2.

Webster, L. R., & Webster, R. M. (2005). Predicting aberrant behaviors in opioid-treated patients: preliminary validation of the Opioid Risk Tool. *Pain medicine* (Malden, Mass.), 6(6), 432-442.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*(6), 361-370.

Papeles del Psicólogo, 2016. Vol. 37(1), pp. 52-61
http://www.papelesdelpsicologo.es

ADVANCES IN THE ASSESSMENT OF ADDICTIONS

Sergio Fernández-Artamendi and Sara Weidberg
Universidad de Oviedo

El fenómeno de la adicción al consumo de drogas legales e ilegales supone un grave problema sociosanitario a nivel global. La correcta evaluación de la gravedad de la adicción y de sus consecuencias es crucial para poder ofrecer alternativas terapéuticas adecuadas a las necesidades de los pacientes. A lo largo de las últimas décadas se han desarrollado multitud de entrevistas diagnósticas para facilitar la detección de los trastornos por consumo de drogas, así como entrevistas centradas en la gravedad de la adicción para determinar el alcance de las consecuencias de su uso. También se han elaborado cuestionarios genéricos que evalúan diversos aspectos relacionados con el uso de sustancias de forma transversal, y cuestionarios específicos que se centran en áreas muy concretas o en sustancias puntuales. La evolución de los instrumentos de evaluación para población adulta ha sido muy significativa, pero las herramientas para adolescentes son de aparición relativamente reciente y la disponibilidad, sobretodo de instrumentos validados al castellano, es más escasa. No obstante, en la actualidad se dispone de una amplia variedad de herramientas psicométricamente robustas. En esta revisión se analiza en primer lugar la magnitud del problema de la adicción a nivel global y nacional, su impacto y coste social, y los antecedentes y avances en la evaluación de las adicciones. Finalmente se discuten las necesidades más acuciantes y las líneas futuras en este campo de la psicología.
Palabras clave: *Adicción, Evaluación, Cuestionarios, Entrevistas, Adultos, Adolescentes.*

The phenomenon of addiction to legal and illegal drugs represents a serious social and health problem at a global level. The correct assessment of the severity of the addiction and its consequences is crucial in order to be able to offer suitable therapeutic alternatives adapted to the needs of the patients. In recent decades, numerous diagnostic interviews have been developed to facilitate the detection of substance use disorders, as well as interviews focused on the severity of the addiction which determine the extent of the consequences of drug use. Additionally, generic questionnaires have been created that assess different aspects related to substance use across multiple substances, as well as specific questionnaires focused on particular areas or substances. The evolution of the assessment instruments for adults has been very significant, but tools aimed specifically at adolescents have only appeared relatively recently and their availability, particularly with regard to instruments validated in Spanish, is scarcer. Nevertheless, there is now a wide variety of psychometrically robust instruments available for professionals. The present review firstly analyses the magnitude of the problem of addiction worldwide and in Spain, its impact and social costs, and the background and advances in the assessment of addictions. Finally, we provide a discussion on the most pressing needs and the future lines of development in this field of psychology.
Key words: *Addiction, Assessment, Questionnaires, Interviews, Adults, Adolescents*

According to the American Psychological Association (APA), addiction "is a condition in which the body must consume a certain substance to prevent withdrawal symptoms of a physical and psychological nature" (American Psychological Association, 2015). This definition can be extended and amended to include new concepts of addiction not related to substance use, such as pathological gambling. However, at present the main healthcare problem due to addiction has to do with drug use (Gowing et al., 2015), and therefore this will be the main focus of this review.

When using instruments for assessment and intervention in the field of addictions, the main guidelines suggest an approach specifically adapted to adults and adolescents (National Institute on Drug Abuse, 2014). In this way, the differences in the substances consumed, the patterns of use and the problems deriving from the use can be dealt with more appropriately. However, while the instruments for attending to adults have advanced considerably in recent decades, in working with adolescents there has been a lack of developmentally appropriate assessment tools (and interventions) until recent years (White, Dennis, & Tims, 2002). Therefore, this paper deals separately with the advances in the field of addictions to substances in each of these populations, with special emphasis on the latest advances for working with adolescents.

Correspondence: Sergio Fernández-Artamendi. *Facultad de Psicología. Universidad de Oviedo. Plaza Feijóo s/n. 33003 Oviedo. España. E-mail: fernandezasergio@uniovi.es*

THE MAGNITUDE OF THE PROBLEM

According to the *Plan Nacional Sobre Drogas* [the National Drug Plan] (PNSD, 2013) among Spaniards aged between 15 and 64 years, regular use of alcohol was detected (typically collected in surveys such as consumption in the last 30 days) in 62.3% of respondents, tobacco in 37.6%, cannabis in 7.0%, hypnosedatives without prescription in 8.3% and tranquilisers in 6.9% of respondents. These rates of consumption are considerably high, and in fact Spain holds the first place in Europe for cocaine consumption, the second place in cannabis use, and it is among the first in the use of tobacco and amphetamines (European Monitoring Centre for Drugs and Drug Addiction, 2015a). Among young Spaniards, approximately 74% had consumed alcohol in the last month, 29.7% tobacco, 16.1% cannabis, 6.6% hypnosedatives and 1.5% cocaine (PNSD, 2014). In this age group, Spain's position is

alarming in the consumption of cannabis (the second highest in Europe) and cocaine (the third highest).

SOCIAL IMPACT

These high rates of consumption are an important social and healthcare risk, manifested in the form of direct costs (mainly healthcare), indirect costs (lower productivity, unemployment, loss of years of life, etc.) and intangible costs (such as personal and family suffering) (World Health Organization, 2014). The World Health Organization (WHO) estimates that about 0.7% of the global cost of health problems is due to the use of cannabis and cocaine, with a global cost of using illegal drugs at around 2% of the economy (World Health Organization, 2008). Alcohol consumption costs between 1.3% and 3.3% of Gross Domestic Product (GDP) (World Health Organization, 2014). Other estimates for the particular case of Spain suggest that illegal drugs represent a cost equivalent to 0.2% of GDP (García-Altes, Olle, Antonanzas & Colom, 2002). With regards to legal drugs, according to Camarelles Guillem et al. (2009), smoking is the health problem that causes the greatest healthcare and social costs faced by Spanish society. More specifically, Lievens et al. (2014) estimated the cost of treatments aimed at problems of alcohol and illegal drug consumption in the EU, concluding that these represent a total cost of 7,600 million euros - in hospital treatment alone, amounting to 2.1 per capita in the case of Spain.

With regards to the indirect costs, it is estimated that in the European Union 3.4% of deaths in people between 15 and 39 years of age are due to substance overdose, in 66% of cases as a result of opioid use (European Monitoring Centre for Drugs and Drug Addiction, 2015b). However, tobacco is the leading cause of preventable death and if it continues at the current rates of consumption, by 2030 it will be responsible for 8 million deaths annually worldwide (World Health Organization, 2015). In Spain, it is calculated that in 2006 more than 53,155 deaths (1 in 7 deaths in people aged 35 years) could be attributable to tobacco use (Banegas et al., 2006). With regards to alcohol, the legal drug par excellence, it is responsible for 3.3 million deaths annually, 5.9% of the total, as well as being the cause of some 200 illnesses and medical problems. According to the analysis conducted by Nutt and his colleagues in the UK (Nutt, King, Phillips, & Independent Sci Comm Drugs, 2010), alcohol is in fact the substance that causes the most social damage, ahead of heroin and crack. Although according to this study, alcohol is not the most harmful drug to the individual consumer, it is however the one that causes the most harm to the people around, so considering its effects as a whole, it is the most damaging. Nationally, it is estimated that 10% of all deaths and 30% of deaths from traffic accidents in 2011 were attributable to alcohol consumption (Pulido et al., 2014).

As for the "intangible" costs, drugs are an important source of personal and family problems which lead many users to seek professional help. According to the Observatorio Español de Drogas y Toxicomanías [Spanish Observatory on Drugs and Drug Addiction] (OEDT, 2014), 3,000 new requests were recorded in 2011 for treatment for heroin use, 10,637 for cocaine use and 9,736 for cannabis consumption, to which must be added 29,014 people being treated for alcohol use. Furthermore, the OEDT notes how each year hospital emergencies receive about 10,000 cases related to the non-therapeutic use of drugs, of which more than half are directly related to consumption. Especially noteworthy in this regard is the increase in the number of cases where cannabis is mentioned, which has gone from 7.4% of the total to 44.9% between 1996 and 2011.

BACKGROUND

In the field of psychological assessment, it is difficult to go back in time beyond a few decades, because it was not primarily until the 70s and thereafter when the object of evaluation in psychology began to extend to multiple fields beyond measuring the individual characteristics (Fernández Ballesteros, 2004). It was from this time onwards that specific assessment instruments began to be created and validated in the field of drug addiction. As a result, there is now a multitude of questionnaires and interviews for the specific assessment of the consequences of consumption, as well as other related aspects. Thus, one can find large repositories of instruments for adolescents and adults on websites such as the National Institute on Drug Abuse (NIDA, www.drugabuse.gov/nidamed-medical-health-professionals), or the Bank of Instruments of the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA, www.emcdda.europa.eu/eib), where one can find tools in Spanish.

However, the bulk of this development has taken place in the field of working with adults while the greatest advances in tools for use with adolescents are from more recent years. Below, we review the background and progress made in the assessment of adults and young people, including the generic and specific questionnaires, and the diagnostic interviews and those referring to the severity of the addiction. To this end, we first include the tools that have been used longest in the clinical and research field and which are the main references in the assessment of addictions. Then, we detail the latest progress including updates of several key tools, the use of new technologies and a number of questionnaires and interviews of recent creation or adaptation into Spanish covering important clinical needs with psychometrically robust alternatives.

Questionnaires

The questionnaires in the field of drug addiction are mainly divided into generic questionnaires (applicable to different substances or providing a cross-sectional assessment) and specific questionnaires for a given substance. The first questionnaires for addictions were of this second type. Specifically, they were for evaluating the problems associated with alcohol use in the adult population, determining the presence of alcoholism and assessing its seriousness. Thus, between the 60s and the 80s, a multitude of self-reports emerged, some of which are now widely used in clinical work, such as the *Michigan Alcoholism Screening Test* (MAST) (Selzer, 1971), the *Alcohol Clinical Index* (Skinner & Holt, 1987), the *CAGE* (the acronym corresponding to Cut down –the need to reduce consumption, Annoyed –the feeling of annoyance due to being criticised for drinking, Guilty –feeling guilty for drinking, and Eye-opener –early morning consumption to steady the

nerves or relieve a hangover) by Ewing (1984), the *Münchner Alkoholismus Test* (MALT) (Feuerlein, Küfner, Ringer, & Antons, 1979) or more recently the *Alcohol Use Disorders Identification Test* (Saunders, Aasland, Babor, De La Fuente, & Grant, 1993). Subsequently, questionnaires have been created aimed at evaluating very varied aspects related to the use of different substances. Table 1 lists, as examples, some of the most used. However, it is not the aim of this study to perform a systematic review of the many existing instruments, and for a more detailed review of the classic (and new) instruments, the aforementioned repositories and other more comprehensive reviews and manuals can be consulted (Fernández Hermida, Secades-Villa, & Fernández-Artamendi, in press; García-Portilla & Bobes-Bascarán, 2011; Winters, McLellan & Dembo, 1999).

The creation of these specific tools was followed by the significant development of generic tools applicable to various substances. These questionnaires enable us to assess the problem of drug use in situations of polydrug use, and to make comparisons between different substances. In this regard, it is worth mentioning two main instruments, widely used today and validated in Spanish:

- ✔ The *Drug Abuse Screening Test* (DAST), by Skinner (1982), validated in Spanish by Pérez-Gálvez et al., (2010). There is a version of 10 items and another of 20, and in its Spanish validation they had a high internal consistency of $\alpha = 0.89$ and 0.93 respectively, showing high reliability and sensitivity to the diagnostic criteria of the Diagnostic Statistic Manual- IV TR (DSM-IV TR, American Psychiatric Association, 1994).
- ✔ The *Severity of Dependence Scale* (SDS) by Gossop et al. (1995), validated in Spanish by González-Sáiz et al. (2008). This consists of five items that assess the severity of dependence on any substance. It has good psychometric properties with a test-restest reliability of 0.89 (Gossop, Best, Marsden, & Strang, 1997).

TABLE 1
SUMMARY OF THE SPECIFIC INSTRUMENTS FOR VARIOUS SUBSTANCES

Substance	Test	Authors
Tobacco	Fagerström Test for Nicotine Dependence (FTND)	Heatherton, Kozlowski, Frecker, & Fagerstrom, 1991
	Nicotine Dependence Syndrome Scale (NDSS)	Shiffman, Waters, & Hickcox, 2004
Cocaine	Cocaine High Risk Situations Questionnaire (CHRSQ)	Michalec et al., 1992
	Cocaine Reasons for Quitting (CRFQ)	McBride et al., 1994
Cannabis	Marijuana Craving Questionnaire (MCQ)	Heishman, Singleton, & Liguori, 2001
	Cannabis Problems Questionnaire (CPQ)	Copeland, Gilmour, Gates, & Swift, 2005
Opiates	Situational Confidence Questionnaire (SCQ)	Barber, Cooper, & Heather, 1991

Questionnaires for adolescents

Since adolescents have their own peculiarities in terms of consumption patterns, associated problems and the evolution of disorders, this population requires adapted tools. While it is true that the *Adolescent Alcohol Involvement Scale* (AAIS) (Mayer & Filstead, 1979) was published back in 1979, one of the first questionnaires aimed at the adolescent population, Leccese and Waldron (1994) note how in the mid-90s most practitioners in the US still used questionnaires in clinical practice that had been developed for use in adults, or without proper validation with young people. If this was the state of things in the US, in Spain the situation was not much different, and it was customary to use tools for the adult population or prepared by the practitioners and institutions themselves. Throughout the 90s, a significant number of specific tools to assess problems of drug use in adolescents began to appear, and greater accessibility to these instruments facilitated the incorporation into clinical practice of reliable questionnaires, validated specifically for young people. However, this development has been slow, and it is not until recently that we find the first instrument for assessing problems due to marijuana use among young people, the *Marijuana Problems Inventory* (Vandrey, Budney, Kamon, & Stanger, 2005).

As for generic questionnaires, the first tools appeared in the 90s. At that time, the urgent need for psychometrically valid instruments that were not focused solely on alcohol led Winters (1992) to develop the *Personal Experience Questionnaire* (PESQ) for adolescents. This is a brief self-report instrument to identify young people in need of treatment for various substances. At the same time, Tarter (1990) developed a more extensive instrument (also applicable as an interview), the *Drug Use Screening Inventory* (DUSI) which reviews multiple areas of the life of the adolescent consumer.

Interviews

The use of interviews in addiction assessment is divided into two main types: diagnostic interviews and problem-focused interviews. The former are an assessment guide to determine whether the consumer has symptoms that meet criteria for the diagnosis of abuse or dependence on one or more psychoactive substances according to the main diagnostic manuals, the ICD-10 (International Classification of Diseases - 10; World Health Organization, 1992) and the DSM-IV-TR (or more recently, the DSM-5). One of the first diagnostic interviews is the *Structured Clinical Interview for DSM* (SCID) available since the earlier versions of the DSM (Spitzer et al., 1992).

Problem focused interviews are aimed at assessing the severity of the consequences caused by consumption. The main instrument that has been the model for subsequent tools is the *Addiction Severity Index* (ASI), developed by McLellan, Luborsky, Woody and O'Brien (1980). It was developed in response to the lack of tools that offered a detailed and comprehensive assessment of the substance consumption and its consequences, which went beyond information on the pattern of use of chemical substances. Over the years, this instrument would be the starting point for many similar tools with proven clinical and research utility.

For working with adolescents, the ASI was also a model for subsequent tools. Thus, an adaptation of the ASI was created, the *Teen-Addiction Severity Index* (T-ASI, Kaminer, Burkstein, & Tarter, 1991) as well as other interviews, such as the ADAD (*Adolescent Drug Abuse Diagnosis*) by Friedman and Utada (1989).

LATEST DEVELOPMENTS

Currently, the field of drug dependency evaluation has many questionnaires for assessing consumption problems in adults, and specific interviews focused on the diagnosis and problems of drug abuse. Also, for assessing teenagers, there are reliable and valid interviews and questionnaires that can be used in clinical practice and research. In this regard, what recent developments can be highlighted in the field of the assessment of addictions? The main developments concerning substance use disorders and their diagnostic criteria are discussed below, as well as some new tools for adults and adolescents.

Diagnostic Criteria

One of the most significant recent changes in the field of the assessment of drug addiction is the modification of the diagnostic criteria of the *Diagnostic Statistic Manual* (DSM), which in its DSM-5 version (American Psychiatric Association, 2013) proposes a new classification with regards to the previous versions such as the DSM-IV-TR (American Psychiatric Association, 1994).

Two diagnoses related to substance use that were included in the DSM-IV-TR are widely used today: abuse and dependence. Given the criteria required for these diagnoses, classically the latter was understood as more severe, and with a more pronounced physiological character compared with the psychosocial problems of abuse, considered to be milder or earlier. In the DSM-5 however, the two diagnoses have been unified under a single heading of *substance use disorder* (American Psychiatric Association, 2013). This new diagnosis also includes a continuum of severity according to which the disorder can be considered as mild or severe. While the new diagnosis is mostly a combination of the previous two, the changes made (see Table 2) include the elimination of the criterion related to the legal problems associated with consumption, which corresponded to the diagnosis of abuse. Thus, the cultural problems deriving from legislative differences between countries on this matter are avoided. Additionally, the new list of symptoms includes craving or the urge to consume, and the diagnostic threshold for a mild disorder requires two to three symptoms, compared with the diagnosis of abuse which required only the presence of a single symptom.

Some of the diagnostic criteria of the DSM-IV-TR had shortcomings when used with adolescents. Primarily, the criterion regarding the legal problems in the diagnosis of abuse (when adolescent consumption is generally illegal in most countries), and those concerning the development of tolerance (a common phenomenon caused by the developmental maturing in the population of young people) and physiological dependence (usually a result of continued use for extended periods of time, difficult to observe in adolescence) for the criterion of dependency (Newcomb, 1995; Winters, 2001).

While eliminating the symptom regarding legal issues is a positive development, the symptoms of tolerance and dependence remain in the DSM-5 and are of limited clinical utility with teenagers. Moreover, the addition of the criterion of craving is also of dubious usefulness in adolescents, as well as that relating to consumption in dangerous situations. Craving is rare in young people, as they have usually recently started using the substance and the possibilities of young people consuming in dangerous situations are lower given their limited access to heavy machinery or risk workplaces. One aspect that has been overcome with the DSM-5 is the elimination of what were known as diagnostic orphans: young people that had one or two symptoms of dependence (insufficient for the cut-off point of three symptoms needed

TABLE 2
COMPARISON OF THE DIAGNOSTIC CRITERIA OF DSM-IV-TR AND DSM-V

DSM-IV-TR		DSM-5	
One or more symptoms: **abuse**	1. Problems at work, school or home 2. Consumption in physically hazardous situations **3. Recurrent legal problems*** 4. Consumption despite interpersonal problems	1. More consumption than intended 2. Persistent desire or unsuccessful effort to control it 3. Much time spent obtaining/using/recovering **4. Craving or urge to use **** 5. Problems at work, school or home 6. Consumption despite interpersonal problems 7. Reduction or abandonment of other activities 8. Consumption in dangerous situations 9. Consumption despite physical or psychological problems 10. Tolerance (need for more of the substance or diminished effect) 11. Withdrawal (withdrawal symptoms or consumption of substance to avoid them)	Presence of at least two symptoms: **Substance Use Disorder (SUD)** Severity: **Light:** 2-3 symptoms **Moderate**: 4-5 **Severe**: 6 or more
Three symptoms or more : **dependence**	1. Tolerance (need for increased amount of the substance or diminished effect) 2. Withdrawal (withdrawal symptoms or consumption of the substance to avoid them) 3. More consumption than intended 4. Persistent desire or unsuccessful efforts to cut down or control 5. Much time spent obtaining/using/recovering 6. Reduction or abandonment of other activities 7. Consumption despite physical or psychological problems		

*Excluded in the DSM-5 **New in the DSM-5

for diagnosis regardless of severity) and none of abuse. Ultimately these cases were not diagnosed despite the problems presented. Therefore, although some problems have been resolved with the DSM-5, an alternative has still not been offered that is specially adapted to substance use in adolescence and that includes the particularities of their associated problems and improves the usefulness of the diagnostic criteria.

NEW TOOLS FOR ASSESSING ADDICTIONS

Questionnaires

In recent years, a number of new tools have emerged for assessing addiction in adults, improving previous versions or using new technologies as a means to facilitate their use and dissemination. An interesting tool recently developed by the WHO and launched in Spain by the *Plan Nacional Sobre Drogas* [National Drug Plan] is the "ASSISTete". It is an adaptation of the ASSIST (*Alcohol, Smoking and Substance Involvement Scale*), created by the WHO (WHO ASSIST Working Group, 2002) for the early detection and early treatment of problems related to substance use, with appropriate Kappa levels of interrater reliability (> 0.60). This tool uses a virtual platform for the assessment (available from http://assistete.es/index.php?nuevo= 1), it requires little time and it also provides a detailed report and a guide with links to the main help services that are closest.

The growing interest in offering specific tools for adolescents, on the part of both clinicians and researchers, has led to the creation of various questionnaires, some of which have already been validated in the Spanish population. As for the generic questionnaires, in the international literature the POSIT (Problem Oriented Screening Instrument for Teenagers) is noteworthy, which includes various domains related to substance use, physical and psychological health, social relationships, skills, problem behaviour, and educational and work situation (Dembo, Turner, Borden, Schmeidler, & Manning, 1994). Most of the subscales have acceptable Cronbach's alphas above 0.70 (Knight, Goodman, Pulerwitz, & DuRant, 2001). The DAST-A (*Drug Abuse Screening Test - Adolescents*) is an adaptation of the aforementioned DAST for use in adolescents, which through a rapid assessment (approximately 5 minutes) provides an estimate of the severity of the problems of consumption (Martino, Grilo & Fehon, 2000), with good internal reliability ($\alpha = 0.91$) and concurrent validity. Another of the instruments recently validated with adolescents is the *Severity of Dependence Scale* (SDS) (G. Martin, Copeland, Gates, & Gilmour, 2006), with a good reliability of $\alpha = 0.83$ (and which has been validated in Spanish by the *Plan Nacional Sobre Drogas* [National Plan on Drugs] (Plan Nacional Sobre Drogas, 2009) for young cannabis users.

TABLE 3
AREAS OF THE VERSIONS OF THE ADDICTION SEVERITY INDEX (ASI)

Europ-ASI	ASI-6	T-ASI
✔ General information	✔ General information	✔ Substance use
✔ Medical situationSubstance use	✔ Alcohol consumption	✔ School situation
✔ Employment/Support	✔ Consumption of drugs	✔ Employment/support
✔ Family/social relationships	✔ Physical health	✔ Family relationships
✔ Legal situation	✔ Mental health	✔ Legal situation (involvement in judicial system)
✔ Psychological/ psychiatric status	✔ Training, employment and financial resources	✔ Social/peer relationships
	✔ Legal situation	✔ Psychiatric status
	✔ Personal and social relationships	

Among the questionnaires of a specific nature, we find several that have already been validated with Spanish adolescents, such as the *Rutgers Alcohol Problems Index* (RAPI) by White and Labouvie (1989), adapted by López-Núñez et al. (2012) for problems related with alcohol consumption. In this study a good reliability was obtained ($\alpha = 0.87$), as well as high sensitivity and specificity for detecting alcohol abuse and dependence. In addition, there are validations of the abbreviated Spanish version of the *Cannabis Problems Questionnaire*, the CPQ-A-S (Fernandez-Artamendi, Fernandez-Hermida, Muniz-Fernandez, Secades-Villa & Garcia-Fernandez, 2012) and the CAST (*Cannabis Abuse Screening Test*, Plan Nacional Sobre Drogas, 2009) which evaluates the severity of problems caused by consuming this substance. Both have good psychometric properties with $\alpha = 0.74$ for the CPQ-A-S and $\alpha = 0.84$ for the CAST (Fernandez-Artamendi et al., 2012). Recently, the Expectancy Questionnaire by Leigh and Stacy (1993) has also been validated with Spanish adolescents (CE, Camacho et al., 2013); this evaluates the positive and negative expectancies concerning alcohol, with a reliability between acceptable ($\alpha = 0.75$) and excellent ($\alpha = 0.96$) for the subscales.

Interviews focused on the problem

As described above, the development of assessment tools in the field of addictions has meant that there are a wide range of instruments available for various purposes today. In the field of adult assessment, the semi-structured clinical interview remains a key instrument, in particular ones derived from the ASI (see Table 3). The Europ-ASI (Kokkevi & Hartgers, 1995) has existed for years in Europe. It is a version of the ASI adapted to the European population, and it is the comprehensive assessment instrument par excellence in the field of addictions. It consists of 141 items divided into 6 areas (health status, employment status, alcohol and drug use, legal problems, family and social relations, and psychiatric status), and its implementation takes about 45-60 minutes. The interview collects information on the problems experienced by users, with particular emphasis on the last 30 days. From these data a series of severity scores are obtained in each of the areas by following a standardised protocol which facilitates the good inter-rater reliability of the instrument. The interview and the application manual can be downloaded from the website of Socidrogalcohol (http://www.socidrogalcohol.org/manuales-y-guías-clínicas-de-socidrogalcohol.html).

The most recent development in this field has been the creation of the ASI-6, a new version of the *Addiction Severity Index*, which corrects some aspects relating to the structure and content of the ASI-5, in order to adapt to new circumstances in the field of drug addiction. This version consists of 257 items and, as well as collecting general information

about the patient, it contains 7 scales regarding 1) alcohol use, 2) drugs, 3) physical health and 4) mental health, 5) training, employment and financial resources, 6) legal situation and 7) personal and social relationships. The Spanish validation found some psychometric weaknesses (Diaz Mesa et al., 2010), but the ASI-6 has been proven useful for treatment planning, follow-up assessments, and it provides relevant information on variables related to adherence to treatment (Casares-Lopez et al., 2011). Nevertheless, for now the Europ-ASI remains the reference tool for the majority of clinicians in the European context. For work with adolescents, the T-ASI (Kaminer et al., 1991) was validated in Spanish a few years ago (Diaz et al., 2008) and its use continues to spread, little by little, in clinical practice and research internationally.

Diagnostic interviews

The appearance of the DSM-5 has involved the adaptation of the main diagnostic interviews for assessing addiction in adults. Thus, the update is now available of the main structured interview, which aims to determine the presence of a substance abuse disorder based on the criteria of the DSM-5 (Structured Clinical Interview for DSM-5) by First, Williams, Karg and Spitzer (2015). Also, the *Composite International Diagnostic Interview - WMH-CIDI* (World Mental Health - Composite International Diagnostic Interview) by Kessler and Üstün (2010), uses the ICD-10 criteria for diagnosing addiction.

In the case of adolescents, in recent years clinical interviews have been developed, such as the DISC-IV (Diagnostic Interview Schedule for Children - IV), recently validated in Spanish by Saldivia, Vicente, Validivia and Melipillan (2013). It uses DSM-IV and ICD-10 criteria for the diagnosis of various disorders, including abuse and dependence on alcohol, nicotine, marijuana and other drugs. One of the most used interviews, the SCID-SUDM (*Structured Clinical Interview for the DSM - Substance Use Disorders Module*) by Spitzer, Williams and Gibbon (1987) has also been adapted for use in adolescents (Martin, Pollock, Bukstein, & Lynck, 2000). The versions adjusted to the DSM-5 criteria have yet to be adapted to adolescents.

New areas of interest

The development of research into the consequences of substance use has expanded the field under evaluation in the world of addictions. In recent years, the importance has been demonstrated of studying the neuropsychological consequences of the use and abuse of substances (Verdejo-García, López-Torrecillas, Orozco Giménez, & Pérez-García, 2004) and their mediating role in the possible outcomes of treatment. The abuse of alcohol, cannabis, cocaine, stimulants, opioids and other substances (Fernández-Serrano, Pérez-García, Río Valle, & Verdejo-García, 2010; Fernández-Serrano, Pérez-García & Verdejo-García, 2011; Verdejo-García, Toribio, Orozco, Puente, & Pérez-García, 2005), has negative consequences on the executive functions (abilities such as short and long term memory, processing speed, visuospatial memory, learning, planning, attention, etc.) which must be taken into account in the assessment process due to the implications for intervention. The new assessment protocols in drug addiction must take into account the particularities of these potential deficits to adjust the interventions.

Another of the aspects of interest in the assessment process is impulsivity. Impulsive behaviour is both a cause and a consequence of drug use (de Wit, 2009). Assessing the levels of impulsivity in children has proved very useful because it allows us to predict the subsequent development of substance use disorders and the age of onset of these disorders (Tarter et al., 2003). The assessment of impulsivity is also convenient in the clinical setting, because this construct is strongly associated with the patients' ability to achieve and maintain abstinence after treatment (Jentsch & Pennington, 2014). Behavioural tasks, such as delay discounting, enable us to assess indirectly the consumer's preference for immediate reinforcers (such as those associated with the pharmacodynamic effects of the substance) versus delayed reinforcers, such as improvements in health, family life and working life (García-Rodriguez, Weidberg, Yoon, García-Fernández, & Secades-Villa, 2013). Delay discounting is a predictor of the treatment success of different substances such as cocaine (Washio et al., 2011), tobacco (Sheffer et al., 2014) or cannabis (Stanger et al., 2012). Furthermore, although the delay discount rates have proved stable in the absence of any intervention (Beck & Triplett, 2009), growing evidence indicates that they can be reduced with effective treatments in various populations of drug addicts (Bickel, Yi, Landes, Hill & Baxter, 2011; Black & Rosen, 2011). In the case of young people, the following have proven particularly useful in the area of addictions: the *Barratt Impulsiveness Scale version* for adolescents (BIS-11-A, Martínez-Loredo, Fernández-Hermida, Fernández-Artamendi, Carballo-Crespo & García-Rodriguez, 2015), with good reliability (α = 0.87), and the *Impulsiveness and Sensation-Seeking subscale* (ImpSS) belonging to the Personality Questionnaire by Zuckerman-Kuhlman (ZKPQ-III, Gutiérrez-Zotes, Ramos Brieva, & Sáiz Ruiz, 2001). These instruments allow us to assess the impulsiveness and decision-making of the substance user, which can be crucial in understanding the habits and patterns of consumption.

DISCUSSION AND RECAPITULATION

The use and abuse of drugs is a serious social problem today worldwide. The personal and social damage caused by the use and abuse of these substances poses a serious risk to the users and to the environment, both economically and in terms of healthcare and society (World Health Organization, 2008). Thus, the aim of this study was to analyse the current needs in the field of the assessment of addictions and the new tools available for clinical work and research.

Fortunately, in recent years the quantity and quality of tools at our disposal in order to carry out this assessment task has grown significantly, offering a wide range of psychometrically robust instruments. There are both generic and specific questionnaires for the different drugs, in Spanish, for different areas of life affected by consumption, and adapted to the particularities of the adult and adolescent population (Becoña Iglesias & Tomás Cortés, 2011; Fernández-Artamendi, Fernández-Hermida, & Secades-Villa, in press; Fernández-Hermida, Secades-Villa, & Fernández-Artamendi, in press; García-Portilla & Bobes-Bascarán, 2011; National Institute on Drug

Abuse, 2014). The need to differentiate between adults and adolescents in the assessment (and intervention) process is no longer debatable today, and the instruments offer a high level of specialisation based on the different substances of consumption. In addition, general tools, such as diagnostic interviews or the assessment of the severity of the addiction and the generic questionnaires, allow us to carry out a first global and cross-sectional approach to the problem of addiction, which is very helpful in planning the intervention. This variety of alternatives has clearly benefited both the clinic and research.

However, it is necessary to continue to promote the use of these tools in daily clinical practice in our country. For this, we must create, adapt and validate tools for certain population groups such as adolescents, since the lack of adaptations and validations substantially limits the possibilities of a valid and reliable assessment. Clinical implications cannot (and must not) be drawn from using tools that have not been properly adapted and validated to the context of use (Callegaro Borsa, Figueiredo Damásio, & Ruschel Bandeira, 2012, Gudmundsson, 2009). But today, the variety of instruments available in the English-speaking world far exceeds those that can be found in Spanish.

Moreover, new issues are arising that must be resolved in the assessment of addictions. Psychologists must keep in mind the possible use of what are known as *synthetic drugs* (ecstasy, synthetic cannabinoids, etc.), which are rapidly evolving and increasingly common especially among young people (Weaver, Hopper, & Gunderson, 2015). In the case of our country, in the future it is necessary to develop new tools for the new substances and situations of consumption, to adjust their formats to the new technologies so they can be used on computers and platforms such as the Internet. The inclusion of questionnaires and tests for assessing neuropsychological functions, impulsivity and decision-making, among other aspects, can help to facilitate the adaptation of interventions to the particularities of each consumer.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Economy and Employment of the Principality of Asturias for the help given to the Research Group on Addictive Behaviours to develop research activities (Ref: FC-15-grupin-14-047). They also wish to show their appreciation to Professors José Ramón Fernández Hermida and Roberto Secades Villa, the directors of the Research Group, for their collaboration in this publication.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders - IV - Text Revised* (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and Statistic Manual of Mental Disorders (5th Edition)*. Washington D.C.: American Psychiatric Association.

American Psychological Association. (08/10/2015). Retrieved from: http://www.apa.org/topics/addiction/

Barber, J. G., Cooper, B. K., & Heather, N. (1991). The Situational Confidence Questionnaire (Heroin). *International Journal of the Addictions, 26*(5), 565-575.

Becoña Iglesias, E., & Cortés Tomás, M. (2011). *Manual de adicciones para psicólogos especialistas en psicología clínica [Addictions manual for specialist psychologists in clinical psychology]*. Valencia: Socidrogalcohol.

Banegas, J.R., Díez-Ganan, L., Banuelos-Marco, B., González-Enriquez, J., Villar-Álvarez, F., Martin-Moreno, J.M., Córdoba-García, R., Pérez-Trullen, A., & Jiménez-Ruíz, C. (2011). Smoking-attributable deaths in Spain, 2006. *Medicina Clínica, 136*(3), 97-102.

Beck, R.C., & Triplett, M.F. (2009). Test-retest reliability of a group-administered paper-pencil measure of delay discounting. *Experimental and Clinical Psychopharmacology, 17*. 345-355. doi: 10.1037/a0017078

Bickel, W.K, Yi, R., Landes, R.D., Hill, P.F., & Baxter, C. (2011). Remember the future: Working memory training decreases delay discounting among stimulant addicts. *Biological Psychiatry, 69*, 260-265. doi: 10.1016/j.biopsych.2010.08.017

Black, A.C., & Rosen, M.I. (2011). A money management-based substance use treatment increases valuation of future rewards. *Addictive Behaviors, 36*, 125-128. doi: 10.1016/j.addbeh.2010.08.014

Callegaro Borsa, J., Figueiredo Damásio, B., & Ruschel Bandeira, D. (2012). Cross-cultural adaptation and validation of psychological instruments: some considerations. *Paidéia, 22*(53).

Camacho, L., Mezquita, L., Ibáñez, M.I., Moya, J., Villa, H., Viruela, A.M., & Ortet, G. (2013). Spanish adaptation of the Expectancy Questionnaire (EQ) about alcohol effects in adolescents. *Psicothema, 25*(4), 529-535.

Camarelles Guillem, F., Salvador Llivina, T., Ramón Torell, J.M., Córdoba García, R., Jiménez Ruíz, C., López García-Aranda, V., . . . López de Santiago, A. (2009). Consenso sobre la atención sanitaria del tabaquismo en España [Consensus on healthcare concerning smoking in Spain]. *Revista Española de Salud Pública, 83*(2).

Casares-Lopez, M.J., Diaz-Mesa, E., Garcia-Portilla, P., Saiz, P., Bobes-Bascaran, M.T., Fonseca-Pedrero, E., . . . Bobes, J. (2011). Sixth version of the Addiction Severity Index: Assessing sensitivity to therapeutic change and retention predictors. *International Journal of Clinical and Health Psychology, 11*(3), 495-508.

Copeland, J., Gilmour, S., Gates, P., & Swift, W. (2005). The Cannabis Problems Questionnaire: Factor structure, reliability, and validity. *Drug and Alcohol Dependence, 80*, 313-319.

de Wit., H. (2009). Impulsivity as a determinant and consequence of drug use: a review of underlying processes. *Addiction Biology, 14*, 22-31. doi: 10.1111/j.1369-1600.2008.00129.x

Dembo, R., Turner, G., Borden, P., Schmeidler, J., & Manning, D. (1994). Screening high-risk youths for potential problems - field application in the use of the Problem-Oriented Screening Instrument for Teenagers (POSIT). *Journal of Child & Adolescent Substance Abuse, 3*(4), 69-93.

Diaz Mesa, E. M., Garcia-Portilla, P., Saiz, P. A., Bobes Bascaran, T.,

Casares, M. J., Fonseca, E., . . . Bobes, J. (2010). Psychometric performance of the 6th version of the Addiction Severity Index in Spanish (ASI-6). *Psicothema, 22*(3), 513-519.

Diaz, R., Castro-Fornieles, J., Serrano, L., Gonzalez, L., Calvo, R., Goti, J., . . . Gual, A. (2008). Clinical and research utility of Spanish Teen-Addiction Severity Index (T-ASI). *Addict Behav, 33*(1), 188-195. doi: 10.1016/j.addbeh.2007.06.002

European Monitoring Centre for Drugs and Drug Addiction. (2015a). Statistical Bulletin 2015. Retrieved from: http://www.emcdda.europa.eu/data/stats2015

European Monitoring Centre for Drugs and Drug Addiction. (2015b). Informe Europeo sobre Drogas. Tendencias y novedades. [European Drug Report. Trends and developments.] Luxembourg: European Monitoring Centre for Drugs and Drug Addiction.

Ewing, J.A. (1984). Detecting alcoholism. The CAGE questionnaire. *JAMA, 252*, 1905-1907.

Fernandez-Artamendi, S., Fernandez-Hermida, J. R., Muniz-Fernandez, J., Secades-Villa, R., & Garcia-Fernandez, G. (2012). Screening of cannabis-related problems among youth: the CPQ-A-S and CAST questionnaires. *Subst Abuse Treat Prev Policy, 7*(1), 13. doi: 10.1186/1747-597X-7-13

Fernández-Artamendi, S., Fernández-Hermida, J.R., & Secades-Villa, Roberto (in press). Intervención psicológica con menores consumidores de drogas [Psychological intervention with child drug users]. In R. Secades-Villa & J. R. Fernández-Hermida (Eds.), *Manual de Adicciones*.

Fernández-Hermida, J.R., Secades-Villa, R., & Fernández-Artamendi, S. (in press). Evaluación clínica de las conductas adictivas [Clinical assessment of addictive behaviours]. En R. Secades-Villa & J. R. Fernández-Hermida (Eds.), *Manual de Adicciones*.

Fernández-Serrano, M.J., Pérez-García, Miguel, Río Valle, J.S., & Verdejo-García, A.J. (2010). Neuropsychological consequences of alcohol and drug abuse on different components of executive functions. *Journal of Psychopharmacology, 24*(9), 1317-1332.

Fernández-Serrano, M.J., Pérez-García, Miguel, & Verdejo-García, A.J. (2011). What are the specific vs. generalized effects of drugs of abuse on neuropsychological performance? *Neuroscience y Biobehavioral Reviews, 35*(3), 377-406. doi: http://dx.doi.org/10.1016/j.neubiorev.2010.04.008

Fernández Ballesteros, R. (Ed.). (2004). *Introducción a la Evaluación Psicológica I [Introduction to Psychological Assessment I]*. Madrid: Pirámide.

Feuerlein, W., Küfner, H., Ringer, C., & Antons, K. (1979). *MALT. Münchner Alkoholismus test. Manual*.

First, M.B., Williams, J.B.W., Karg, R.S., & Spitzer, R.L. (2015). *Structured Clinical Interview for DSM-5 Diagnosis. SCID-5-CV*. Washington D.C.: American Psychiatric Publishing.

Friedman, A.S., & Utada, A. (1989). A method for diagnosing and planning the treatment of adolescent drug abusers: the Adolescent Drug Abuse Diagnosis (ADAD) instrument. *Journal of Drug Education, 19*, 285-312.

Garcia-Altes, A., Olle, J. M., Antonanzas, F., & Colom, J. (2002). The social cost of illegal drug consumption in Spain. *Addiction, 97*(9), 1145-1153. doi: 10.1046/j.1360-0443.2002.00170.x

García-Portilla, P., & Bobes-Bascarán, MT. (2011). *Banco de instrumentos básicos para la Práctica de la Psiquiatría Clínica (6ª edición) [Bank of basic instruments for the Practice of Clinical Psychiatry (6th edition)]*. Barcelona: Psiquiatría Editores, SL.

García-Rodriguez, O., Weidberg, S., Yoon, J. H., García-Fernández, G., & Secades-Villa, R. (2013). Evaluación de la impulsividad mediante una tarea de descuento por demora en adictos a la cocaína [Assessing impulsivity using a delay discounting task in cocaine addicts]. *Revista Iberoamericana de Psicología y Salud, 4*(1), 23-35.

Gonzalez-Saiz, F., de las Cuevas, C., Barrio, G., & Domingo-Salvany, A. (2008). Spanish version of the Severity of Dependence Scale (SDS). *Medicina Clinica, 131*(20), 797-798. doi: 10.1016/s0025-7753(08)75509-x

Gossop, M., Best, D., Marsden, J., & Strang, J. (1997). Test-retest reliability of the severity of dependence scale. *Addiction, 92*(3), 353-353. doi: 10.1111/j.1360-0443.1997.tb03205.x

Gossop, M., Darke, S., Griffiths, P., Hando, J., Powis, B., Hall, W., & Strang, J. (1995). The Severity of Dependence Scale (SDS) - Psychometric properties of the SDS in English and Austrian samples of heroin, cocaine and amphetamine users. *Addiction, 90*(5), 607-614. doi: 10.1111/j.1360-0443.1995.tb02199.x

Gowing, L.R., Ali, R.L., Allsop, S., Marsden, J., Turf, E.E., West, R., Witton, J. (2015). Global statistics on addictive behaviors: 2014 status report. *Addiction, 110*(6).

Gudmundsson, E. (2009). Guidelines for translating and adapting psychological instruments. *Nordic Psychology, 61*(2), 29-45.

Gutiérrez-Zotes, J.A., Ramos Brieva, J.A., & y Sáiz Ruiz, J. (2001). Desarrollo de la versión española del cuestionario de personalidad Zuckerman-Kuhlman (ZKPQ-III) y propiedades psicométricas [Developing the Spanish version of the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ-III) and psychometric properties]. *Psiquis, 22*(6), 239-250.

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerstrom, K. O. (1991). The Fagerstrom Test for Nicotine Dependence - A revision of the Fagerstrom Tolerance Questionnaire. *British Journal of Addiction, 86*(9), 1119-1127.

Heishman, S.J., Singleton, E.G., & Liguori, A. (2001). Marijuana Craving Questionnaire: development and initial validation of a self-report instrument. *Addiction, 96*, 1023-1034.

Jentsch, J.D., & Pennington, Z.T. (2014). Reward, interrupted: Inhibitory control and its relevance to addictions. *Neuropharmacology, 76*, 479-486.

Kaminer, Y., Burkstein, O.G., & Tarter, R.E. (1991). The Teen Addiction Severity Index: Rationale and reliability. *The International Journal of Addictions, 26*, 219-226.

Kazdin, A.E. (Ed.). (2000). *Encyclopedia of Psychology*. Washington, D.C.: American Psychological Association Publication.

Knight, J. R., Goodman, E., Pulerwitz, T., & DuRant, R. H. (2001). Reliability of the Problem Oriented Screening Instrument for

Teenagers (POSIT) in adolescent medical practice. *Journal of Adolescent Health, 29*(2), 125-130.

Kokkevi, A., & Hartgers, C. (1995). EuropASI: European adaptation of a multidimensional assessment instrument for drug and alcohol dependence. *European Addiction Research, 1*(4), 208-210.

Leccese, M., & Waldron, H. B. (1994). Assessing adolescent substance use: a critique of current measurement instruments. *Journal of Substance Abuse Treatment, 11*(6), 553-563.

Leigh, B.C., & Stacy, A.W. (1993). Alcohol outcome expectancies: Scale construction and predictive utility in higher order confirmatory models. *Psychological Assessment, 5*, 216-229.

Lievens, Delfine, Vander Laenen, Freya, & Christiaens, Johan. (2014). Public spending for illegal drug and alcohol treatment in hospitals: an EU cross-country comparison. *Substance Abuse Treatment Prevention and Policy, 9, 29*. doi: 10.1186/1747-597x-9-26

López-Nuñez, C., Fernández-Artamendi, S., Fernández-Hermida, J.R., Campillo-Álvarez, A., & Secades-Villa, R. (2012). Spanish Adaptation and Validation of the Rutgers Alcohol Problems Index (RAPI). *International Journal of Clinical and Health Psychology, 12*(2), 251-264.

Martin, C.S., Pollock, N.K., Bukstein, O.G., & Lynck, K.G. (2000). Inter-rater reliability of the SCID alcohol and substance use disorders section among adolescents. *Drug and Alcohol Dependence, 59*, 173-176.

Martin, G., Copeland, J., Gates, P., & Gilmour, S. (2006). The Severity of Dependence Scale (SDS) in an adolescent population of cannabis users: reliability, validity and diagnostic cut-off. *Drug and Alcohol Dependence, 83*(1), 90-93. doi: 10.1016/j.drugalcdep.2005.10.014

Martínez-Loredo, V., Fernandez-Hermida, J. R., Fernández-Artamendi, S., Carballo-Crespo, J.L., & García-Rodriguez, O. (2015). Spanish adaptation and validation of the Barratt Impulsiveness Scale for early adolescents (BIS-11-A). *International Journal of Clinical and Health Psychology, 15*(3), 274-282.

Martino, S., Grilo, C.M., & Fehon, D.CV. (2000). Development of the Drug Abuse Screening Test for Adolescents (DAST-A). *Addictive Behaviors, 25*, 57-70.

Mayer, J. E., & Filstead, W. J. (1979). Adolescent Alcohol Involvement Scale - Instrument for measuring adolescent use and misuse of alcohol. *Alcoholism-Clinical and Experimental Research, 3*(2), 186-186.

McBride, C. M., Curry, S. J., Stephens, R. S., Wells, E. A., Roffman, R. A., & Hawkins, J. D. (1994). Intrinsic and extrincsic motivation for change in cigarette smokers, marijuana smokers, and cocaine users. *Psychology of Addictive Behaviors, 8*(4), 243-250.

McLellan, A.T., Luborsky, L., Woody, G.E., & O'Brien, C.P. (1980). An improved diagnostic evaluation instrument for substance abuse patients: The Addiction Severity Index. *Journal of Nervous and Mental Disease, 168*, 26-33.

Michalec, E., Zwick, W. R., Monti, P. M., Rohsenow, D. J., Varney, S., Niaura, R. S., & Abrams, D. B. (1992). A Cocaine high-risk situations questionnaire: development and psychometric properties. *Journal of Substance Abuse, 4*(4), 377-391. doi: 10.1016/0899-3289(92)90044-x

National Institute on Drug Abuse. (2014). *Principles of adolescent substance use disorder treatment: A research-based guide*: National Institute on Drug Abuse.

Newcomb, M.D. (1995). Identifying High-Risk Youth: Prevalence and patterns of adolescent drug Abuse. In E. Rahdert & D. Czechowicz (Eds.), *Adolescent drug abuse: Clinical assessment and therapeutic interventions*. (pp. 7-38). Rockville, MD: National Institutes of Health.

Nutt, D. J., King, L. A., Phillips, L. D., & Independent Scientific Committee on Drugs. (2010). Drug harms in the UK: a multicriteria decision analysis. *Lancet, 376*(9752), 1558-1565. doi: 10.1016/s0140-6736(10)61462-6

Observatorio Español de las Drogas y las Toxicomanías [Spanish Monitoring Centre on Drugs and Drug Addiction]. (2014). Estadísticas 2013. Alcohol, tabaco y drogas ilegales en España [Statistics 2013. Alcohol, tobacco and illegal drugs in Spain]. Madrid: Plan Nacional Sobre Drogas.

Organización Mundial de la Salud [World Health Organisation]. (1992). *Clasificación Internacional de Enfermedades, 10ª Edición [International Classification of Diseases, 10th Edition]*. Madrid: Meditor.

Pérez-Gálvez, B., García Fernández, L., de Vicente Manzanaro, M.P., Oliveras Valenzuela, M.A., & Lahoz Lafuente, M. (2010). Spanish validation of the Drug Abuse Screening Test (DAST-20 and DAST-10). *Health and Addictions, 10*, 35-50.

Plan Nacional Sobre Drogas. (2009). Consumo problemático de cannabis en estudiantes españoles de 14-18 años: Validación de escalas [Problematic use of cannabis in Spanish students 14-18 years: Validation of scales]. Madrid: Ministerio de Sanidad y Política Social [Ministry of Health and Social Policy].

Plan Nacional Sobre Drogas. (2013). Encuesta sobre alcohol y drogas en población general en España EDADES 2011-2012 [Survey on alcohol and drugs in the general population in Spain EDADES 2011-2012]. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad [Ministry of Health, Social Services and Equality].

Plan Nacional Sobre Drogas. (2014). Informe de la Encuesta Estatal Sobre Uso de Drogas en Estudiantes de Enseñanzas Secundarias (ESTUDES) 2012-2013 [Report of the National Survey on Drug Use in Secondary Education Students (ESTUDES) 2012-2013]: Delegación del Gobierno para el Plan Nacional Sobre Drogas [Government Delegation for the National Plan on Drugs]. Ministerio de Sanidad y Política Social [Ministry of Health and Social Policy].

Pulido, J., Indave-Ruiz, B.I., Colell-Ortega, E., Ruíz-García, M., Bartroli, M., & Barrio, G. (2014). Estudios poblacionales en España sobre daños relacionados con el consumo de alcohol [Population studies in Spain on damage related to alcohol consumption]. *Revista Española de Salud Pública, 88*, 493-513.

Saldivia, S., Vicente, B., Valdivia, M., & Melipillan, R. (2013). Validación de la entrevista diagnóstica estructurada DISC-IV para la identificación de trastornos psiquiátricos en niños y adolescentes en Chile [Validation of the diagnostic structured interview DISC-IV for identifying psychiatric disorders in children and adolescents in Chile]. *Revista Chilena de Neuropsiquiatría, 51*(1), 70-78.

Saunders, J. B., Aasland, O. G., Babor, T., De La Fuente, J., & Grant,

M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption II. *Addiction, 86*, 791-804.

Selzer, M.L. (1971). The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *American Journal of Psychiatry, 127*(12), 1653-1658. doi: doi:10.1176/ajp.127.12.1653

Sheffer, C.E., Christensen, D.R., Landes, R.D., Carter, L.P., Jackson, L. & Bickel, W. (2014). Delay Discounting rates: A strong prognostic indicator of smoking relapse. *Addictive Behaviors, 39*, 1682-1686. doi: 10.1016/j.addbeh.2014.04.019

Shiffman, S., Waters, A. J., & Hickcox, M. (2004). The Nicotine Dependence Syndrome Scale: A multidimensional measure of nicotine dependence. *Nicotine y Tobacco Research, 6*(2), 327-348. doi: 10.1080/1462220042000202481

Skinner, H. A. (1982). The Drug-Abuse Screening Test. *Addictive Behaviors, 7*(4), 363-371. doi: 10.1016/0306-4603(82)90005-3

Skinner, H. A., & Holt, S. (1987). *Alcohol Clinical Index*. Toronto: Addiction Research Foundation.

Spitzer, R., Williams, J., & Gibbon, B. (1987). *Instructions manual for the Structured Clinical Interview for the DSM-III-R*. New York: New York State Psychiatric Institute.

Spitzer, R., Williams, J. W., Gibbon, M., & First, M. B. (1992). The Structured Clinical Interview for DSM-III-R (SCID): I: History, rationale, and description. *Archives of General Psychiatry, 49*(8), 624-629. doi: 10.1001/archpsyc.1992.01820080032005

Stanger, C., Ryan, S.R., Fu, H.Y., Landes, R.D., Jones, B.A., Bickel, W., & Budney, A.J. (2012). Delay Discounting predicts adolescent substance abuse treatment outcome. *Experimental and Clinical Psychopharmacology, 20*(3), 205-212. doi: 10.1037/a0026543

Tarter, R.E. (1990). Evaluation and treatment of adolescent substance abuse: A decision tree method. *American Journal of Drug and Alcohol Abuse, 16*, 1-46.

Tarter, R.E., Kirisci, L, Mezzich, A., Cornelius, J.R., Pajer, K, Vanyukov, M… Clark, D. (2003). Neurobehavioral desinhibition in childhood predicts early age at onset of substance use disorder. *American Journal of Psychiary, 160*(6), 1078-1085.

Vandrey, R., Budney, A. J., Kamon, J. L., & Stanger, C. (2005). Cannabis withdrawal in adolescent treatment seekers. *Drug and Alcohol Dependence, 78*(2), 205-210. doi: 10.1016/j.drugalcdep.2004.11.001

Verdejo-García, A.J., López-Torrecillas, Francisca, Orozco Giménez, C., & Pérez-García, Miguel. (2004). Clinical implications and methodological challenges in the study of the neuropsychological correlates of cannabis, stimulant, and opioid abuse. *Neuropsychology Review, 14*(1), 1-41.

Verdejo-García, A.J., Toribio, I., Orozco, C., Puente, K.L., & Pérez-García, M. (2005). Neuropsychological functioning in methadone maintenance patients versus abstinent heroin abusers. *Drug and Alcohol Dependence, 78*(3), 283-288. doi: http://dx.doi.org/10.1016/j.drugalcdep.2004.11.006

Washio, Y., Higgins, S.T., Heil, S.H, McKerchar, T.L., Badger, G.J., Skelly, J.M., & Dantona, R.L. (2011). Delay Discounting is associated with treatment response among cocaine-dependent outpatients. *Experimental and Clinical Psychopharmacology, 19*(3), 243-248. doi: 10.1037/a0023617

Weaver, M.F., Hopper, J.A., & Gunderson, E.W. (2015). Designer drugs 2015: assessment and management. *Addiction science y clinical practice, 10*, 8-8. doi: 10.1186/s13722-015-0024-7

White, W. L., Dennis, M., & Tims, F. (2002). Adolescent treatment: Its history and current renaissance. *Counselor, 3*(2), 20-23.

Winters, K.C. (1992). Development of an adolescent substance abuse screening instrument: The Personal Experience Screening Questionnaire. *Addictive Behaviors, 17*, 479-490.

Winters, K.C. (2001). Assessing adolescent substance use problems and other areas of functioning. In P. M. Monti, S. M. Colby & T. A. O'Leary (Eds.), *Adolescents, Alcohol, and Substance Abuse*. New York, NY: The Guildford Press.

Winters, K.C., McLellan, A.T., & Dembo, R. (1999). *Screening and Assessing Adolescents for Substance Use Disorders: Treatment Improvement Protocol (TIP) Series 31*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

WHO ASSIST Working Group. (2002). The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Development, realiability and feasibility. *Addiction, 97*(9), 1183-1194.

World Health Organization. (2008). The global burden of disease: 2004 Update. Switzerland: World Health Organization.

World Health Organization. (2014). Global status report on alcohol and health 2014. Luxemburg: World Health Organization.

World Health Organization. (2015). Alcohol. Fact Sheet. Retrieved from: http://www.who.int/mediacentre/factsheets/fs349/en/

Papeles del Psicólogo, 2016. Vol. 37(1), pp. 62-68
http://www.papelesdelpsicologo.es

THE ASSESSMENT OF ENTREPRENEURIAL PERSONALITY: THE CURRENT SITUATION AND FUTURE DIRECTIONS

Javier Suárez-Álvarez and Ignacio Pedrosa
Universidad de Oviedo

El espíritu emprendedor de las personas resulta fundamental en la sociedad moderna ya que supone una importante fuente de innovación, empleo, productividad y crecimiento. Si bien los primeros modelos teóricos surgen desde aproximaciones económicas y sociológicas, la psicología proporciona modelos que integran diferentes aspectos centrados en el individuo como los cognitivos, actitudinales y de la personalidad permitiendo así un estudio mucho más detallado. El objetivo de este trabajo es presentar las principales aportaciones de la psicología a la evaluación de la personalidad emprendedora. Para ello se llevó a cabo una revisión de los principales modelos e instrumentos de medida que se han desarrollado hasta la fecha. Los resultados confirman que la personalidad emprendedora tiene una estructura multidimensional pudiendo destacar ocho dimensiones fundamentales: motivación de logro, autoeficacia, toma de riesgos, innovación, autonomía, tolerancia al estrés, locus de control interno y optimismo. Desde un punto de vista metodológico, la Teoría de Respuesta a los Ítems y los Tests Adaptativos Informatizados representan los más avanzados y modernos métodos de evaluación. En la actualidad, se dispone de un amplio número de instrumentos de medida para evaluar la personalidad emprendedora. Las líneas futuras de investigación deberán orientarse hacia la construcción de modelos multidimensionales, así como a proporcionar alternativas que permitan reducir la deseabilidad social y otros sesgos inherentes a los autoinformes.

Palabras Clave: *Personalidad emprendedora, Espíritu emprendedor, Teoría de Respuesta a los ítems, Tests adaptativos informatizados, Autoinforme.*

Entrepreneurship is fundamental in modern society because it represents an important source of innovation, employment, productivity, and growth. While the first theoretical models arose from economic and sociological approaches, psychology provides models that integrate different aspects such as cognitions, attitudes and personality, which allow a more detailed study. The purpose of this paper is to show the main contributions of psychology to the assessment of the enterprising personality. For this purpose, the main models and instruments developed to date were reviewed. The results confirm that the enterprising personality has a multidimensional structure and eight personality traits can be highlighted: achievement motivation, risk-taking, autonomy, self-efficacy, stress tolerance, innovativeness, internal locus of control, and optimism. From a methodological point of view, Item Response Theory and Computerised Adaptive Tests represent the most advanced and modern methods for assessing enterprising personality. There are currently several measurement instruments available. Future areas of research should be directed at the construction of multidimensional models as well as providing alternatives that facilitate a reduction in social desirability and other biases inherent in self-reports.

Key words: *Enterprising personality, Entrepreneurship, item response theory, Computerised adaptive tests, self-report.*

Many internationally renowned organisations, such as the Global Entrepreneurship Research Association (GEM, 2015) and the Organisation for Economic Co-operation and Development (OECD, 2014), annually assess entrepreneurial activity in a large number of countries. Currently, governments, organisations, universities and individuals are joining forces to try to understand the process of entrepreneurship, since its promotion is essential for the development of market economies (OECD / The European Commission, 2013). Furthermore, failure as an entrepreneur involves a cost to society in terms of lost opportunities and resources, and it ultimately causes significant consequences for the individual, both economically and psychologically (Zhao, Seibert, & Lumpkin, 2010).

It is important to distinguish between entrepreneurship and business ownership, as the latter is confined to the exclusive domain of the company as one specific example of the many possible manifestations of entrepreneurship. Depending on the objectives to be achieved, it is possible to identify at least three types of entrepreneur in terms of the objectives they aim to achieve: extra-entrepreneur, intra-entrepreneur and personal entrepreneur. The extra-entrepreneur is a person whose goal is to develop new external projects related to business creation (Rauch & Frese, 2007b). The intra-entrepreneur is a person who is responsible for creating innovation within a business, improving projects that are already underway (Lumpkin, 2007). The personal entrepreneur is characterised by a high level of personal control and initiative, and is able to handle difficult situations, for example, stressors, unemployment or career changes (Frese & Fay, 2001). A particular case of this type of entrepreneur would be those who are oriented towards volunteer work in the community, in non-governmental organisations, or the researcher who proposes a theory or technique that has not previously been explored.

The study of all matters relating to entrepreneurship has grown exponentially in recent years, establishing itself as a multidisciplinary field of research (Rauch & Frese, 2007a; Sánchez, 2011). This consolidation is explained largely by the contribution of models and theories proposed from an economic and sociological approach (Chell, 2008). Psychology has also played a central role in legitimising and even popularising the study of entrepreneurship (Baum, Frese, Baron, & Katz, 2007; Hisrich, Lagan-Fox & Grant, 2007). One possible

Correspondence: Javier Suárez-Álvarez. *Departamento de Psicología. Universidad de Oviedo. Plaza Feijoo, s/n. 33003 Oviedo. España. E-mail: suarezajavier@uniovi.es*

explanation is the tendency of psychology not only to construct theories, but also to construct rigorous measurement instruments that enable us to support such theories empirically (Katz, 2007). In this sense, the psychometric methodology involves the fundamental scientific research tool for the systematic and rigorous study of the phenomenon of entrepreneurship. Three approaches have mainly been used to describe entrepreneurship: the economic, the sociological and the psychological approaches.

THE ECONOMIC APPROACH

The origin of the theories of entrepreneurship emerges mainly from an economic perspective and has its roots in Europe, especially in France (Cantillon, 1756; Say, 1803), the UK (Marshall, 1890) and Austria (Kirzner, 1973; Schumpeter, 1934). Economists tend to develop theories about the decisions that are relevant to the use of resources in order to obtain economic results, such as the performance results of companies, industries and countries. These types of theories tend to give more weight to economic variables (e.g., resources, capital, information or business opportunities) than individual aspects. However, from the beginning classical economic theories have tended to incorporate aspects such as innovation or leadership (Marshall, 1890), to assume that the personal characteristics of the entrepreneur can be acquired (Schumpeter, 1934) and, ultimately, to accept the subjective and individual character of entrepreneurship.

THE SOCIOLOGICAL PERSPECTIVE

In recent decades, different approaches have been emerging, from a sociological perspective, that attempt to provide a more complete picture of the issues involved in the process of entrepreneurship (Chell, 2008). According to these approaches, family background and education are two key aspects in the development of entrepreneurship. Developing in an entrepreneurial environment has a positive and facilitating influence on entrepreneurial behaviour (Altinay, Madanoglu, Daniele & Lashley, 2012). Similarly, the possibility of receiving training on how to be an entrepreneur turns adolescence into a particularly interesting stage (Unger, Rauch, Frese & Rosenbusch, 2011). The main reason is that potentially entrepreneurial students who attend specific training courses would increase their skills, knowledge and abilities to be able to take advantage of the opportunities presented to them compared to those who do not receive such training (Volery, Muller, Oser, Naepflin & del Rey, 2013).

One of the most interesting contributions of the sociological perspective is the emphasis that it places on subjectivity. For example, many of the proposed models highlight the importance of how people perceive the viability of their projects and their perception of control over the resources (Ajzen, 1991; Shapero & Sokol, 1982; Veciana, 1999). This subjectivity depends largely on the culture and context where the person is. The cultural norms and laws and regulations of each individual country have an important influence on the perception and behaviour of entrepreneurs (GEM, 2015; OECD, 2014).

According to the latest GEM Spain 2014 (GEM, 2015), approximately six out of ten entrepreneurs are men; however, the difference between men and women entrepreneurs setting up businesses has declined over the past two years. Of all entrepreneurs, 47.6% have received some sort of higher education or graduate degree, and 43.5% have at some point in their lives received specific training in entrepreneurship. The adults with higher incomes were the ones that showed a greater propensity to start new businesses. Of all potential entrepreneurs (i.e., people with an intention to create a company in less than three years), 18.2% are aged between 18 and 24 years. Of this same group of potential entrepreneurs, 56.6% say they have no specific training, a percentage that rises to 63.6% for those who are leaving a job at a company. These circumstances lead us to believe that specific training focused on entrepreneurship would facilitate both the development and the consolidation of new projects.

Spain is characterised by a perception of having a lower number of opportunities than other European countries. About 16% of Spaniards perceive that there are business opportunities, while in countries such as Britain and Germany the figure is over 30% (GEM, 2015). However, it is curious that the perception of entrepreneurial knowledge and skills is above the European average. About 50% of Spaniards consider themselves to have sufficient skills and knowledge to start a business, while in countries such as Germany and France the percentage is less than 40% (GEM, 2015). These results suggest that variables such as self-concept, motivation and expectations may play an important role in both the perception of opportunities and the perception of competence. In this regard, educational research has already gone into great detail about the significant weight that such variables have, for example, in academic achievement (Suárez-Álvarez, Fernández-Alonso & Muñiz, 2014).

According to the latest *Entrepreneurship at a Glance* (OECD, 2014), necessity was an important driver in emerging economies such as China and India, but also in Korea, Estonia, Greece and Spain, which partly reflects the economic crisis. In fact, in 2013, 29.8% of Spanish entrepreneurs who started a company said that they did so after considering that it was their only career option (GEM, 2015). The combination of opportunities, capacities and resources does not necessarily lead to entrepreneurial activity if the costs of the opportunity (e.g., lost earnings or poorer health coverage) and the initial costs outweigh the potential benefits. In fact, as noted by the OECD, "the regulatory framework and taxes become a critical factor that affects the business performance of countries" (OECD, 2014, p. 86). In sum, these findings shed light on some of the problems that entrepreneurs are now finding and emphasise the importance of education in the process of entrepreneurship.

THE CONTRIBUTION OF PSYCHOLOGY

The research carried out to date has shown that entrepreneurial behaviour is influenced by numerous factors including economic, social and personal aspects (Chell, 2008; Rauch & Frese, 2007a). Based on these results it seems reasonable to believe that entrepreneurial behaviour is multidimensional. Therefore, developing models and comprehensive explanations that realistically reflect entrepreneurship requires the consideration of various dimensions together. Figure 1 proposes an integral model of entrepreneurship that captures the essence of the main models developed to date (Rauch & Frese, 2000; Rauch & Frese, 2007a; Sánchez, 2011) and incorporates the latest research findings on entrepreneurial personality. It is a comprehensive model that presents the major aspects involved in entrepreneurial activity. The comprehensive model of entrepreneurship (Figure 1) can serve as a preliminary outline on which to base future research. While the model is plausible to the extent that each of the issues separately has

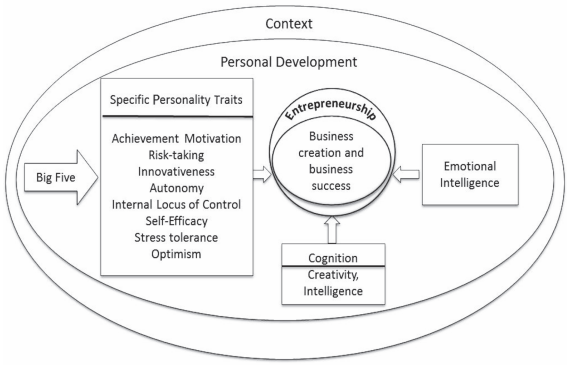
proven to be connected with entrepreneurial activity, more research is required to relate the set of variables as a whole.

The model pays special attention to the dimensions that comprise the area of personal development, which is influenced by the variables that encompass the socio-economic context, such as education, family, culture and the system of rules, laws and regulations of the countries (GEM, 2015; OECD, 2014). Within the area of personal development, the work focused on emotional intelligence deserves special attention (Ahmetoglu, Leutner, & Chamorro-Premuzic, 2011), because of the relationship it has with aspects such as innovation (Suliman & Al-Shaikh, 2007) or achievement motivation and self-efficacy (Muñiz, Suárez-Álvarez, Pedrosa, Fonseca-Pedrero & García-Cueto, 2014). Another key part of this model relates to the cognitive aspects, which include such constructs as cognitive styles (Sánchez, Carballo & Gutiérrez, 2011), creativity (Ward, 2004) and intelligence (Newton & McGrew, 2010). The study of personality has gained particular momentum in recent years and two main approaches can be distinguished: researchers who prefer to use broad personality traits, such as the Big Five (Brandstätter, 2011; Zhao et al, 2010); and those who propose the use of more specific traits that are closer to entrepreneurial activity (Rauch & Frese, 2007a, 2007b; Suárez-Álvarez, Pedrosa, García-Cueto & Muñiz, 2014).

The supporters of using general personality traits argue that these factors (extraversion, emotional stability, responsibility, agreeableness and openness to experience), account for around 13% of the variance of entrepreneurial activity and about 10% of business success (Zhao et al., 2010), and correlate with the activity of business owners and managers (Brandstätter, 2011). Specifically, the dimensions of responsibility and openness to experience are the ones that have a greater relationship with both entrepreneurial behaviour and business performance (Zhao et al., 2010). Also used within this line of research, although much less representatively, are the personality factors assessed by the Eysenck Personality Questionnaire Revised (Furnham, 2002) and the 16 personality factors of Cattell (Chell, 2008).

The specific personality traits that seem to be most related to entrepreneurial personality are achievement motivation, self-efficacy, risk-taking, innovation, autonomy, stress tolerance, internal locus of control and optimism (Baum et al, 2007; Muñiz et al., 2014; Rauch & Frese, 2007a, 2007b; Suárez-Álvarez et al, 2014; Zhao et al, 2010). The central argument supporting the use of models of specific personality traits rather than broad traits, is that the specific traits would be able to explain more specific aspects of the entrepreneurial personality (Laguna, 2013; Lanero, Vázquez & Muñoz-Adánez, 2015; Tyszka, Cieslik, Domurat & Macko, 2011), so the predictions made based on them would be more accurate. This is a plausible hypothesis on which the first data consistent is starting to be collected, moderate relations being found with regards to business creation and success (Rauch & Frese, 2007a, 2007b). In fact, today there are findings that suggest that the more specific traits of the entrepreneurial personality add evidence of predictive validity of business success to the Big Five personality traits (Leutner, Ahmetoglu, Akhtar & Chamorro-Premuzic, 2014). In other words, the inclusion of both measures during the evaluation process would improve decision making and predictive power. For this reason, the model presented in Figure 1 comprises the two models of entrepreneurial personality together.

FIGURE 1
COMPREHENSIVE MODEL OF ENTREPRENEURSHIP, ADAPTED BY RAUCH AND FRESE (2007A)



ASSESSMENT OF THE ENTREPRENEURIAL PERSONALITY

To date, several measuring instruments have been developed to assess the various personality traits involved in entrepreneurial behaviour, such as achievement motivation (Suárez-Álvarez, Campillo-Álvarez, Fonseca-Pedrero, García-Cueto & Muñiz, 2013), locus of control (Suárez-Álvarez, Pedrosa, García-Cueto & Muñiz, in press) or self-efficacy (Moriano, Palací & Morales, 2012) to name just a few of them. Additionally, there are numerous tools that have been developed based on the construct of entrepreneurial orientation (Covin & Wales, 2012) and that assess dimensions such as risk taking, proactivity, innovation, autonomy and competitive aggressiveness. However, the number of instruments is significantly reduced when the aim is the joint evaluation of the entrepreneurial personality traits in one single instrument with methodological consistency, and the number of instruments developed in Spain is even more scarce (Muñiz et al., 2014; Sánchez, 2010).

In recent years, significant contributions have been made to the assessment of the specific traits of the entrepreneurial personality. In Table 1, we present the main instruments for assessing entrepreneurship that have been developed to date. Some of these scales have been translated and adapted to different languages (Almeida et al, 2014; Caird, 2006, Liñán & Chen, 2006) and are aimed at evaluating different groups such as adolescents (Muñiz et al, 2014), university students (Caird, 2006) and workers (Almeida et al., 2014). Another aspect to note is the tendency to develop methods using self-report instruments, usually measured by Likert scales.

In Table 2, we present an overall assessment indicative of the quality of the measuring instruments. This is determined according to the criteria established by the European Federation of Psychologists Associations (EFPA) for the evaluation of tests (Evers et al., 2013) and the Standards for Educational and Psychological Assessment (American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 2014). The information shown in Table 2 corresponds mainly to the information provided by the authors in the original document in which the development of the instrument is shown. This information is completed with scientific papers indexed in international databases. This excludes the possible existence of documents that are not indexed in these databases which provide

information on aspects that are not covered in this table. First, it is striking that while some authors mention content validity, few provide data based on expert judgment and quantitative indicators (Pedrosa, Suárez-Álvarez & García-Cueto, 2013). Clearly, another great omission is the study of DIF, which identifies whether there are items that systematically harm certain groups of people such as, for example, men or women (Sandilands, Oliveri, Zumbo & Ercikan, 2013).

MEASUREMENT INSTRUMENTS IN SPAIN

At present there are at least four measurement instruments to assess the entrepreneurial personality in Spain: EIQ (Liñan & Chen, 2006); COE (Sánchez, 2010); META (Almeida, Ahmetoglu & Chamorro-Premuzic, 2014); and BEPE (Muñiz et al., 2014). It is important to note that EIQ, COE and BEPE were originally developed in Spain while META was originally developed in the UK. Although the latter can be answered in Spanish on its website (http://www.metaprofiling.com), it should be noted that there is no information available on the psychometric properties of the translation and adaptation of the instrument to the Spanish context to date. Therefore it is not possible to assess the suitability of the instrument for use in Spain by international standards (Muñiz, Elosua & Hambleton, 2013). On the other hand, the main limitation of the instruments developed in Spain is the lack of criterion validity evidence (Table 2). While the use of these measurement instruments for research may be suitable for certain purposes, their use would still not be adequate for making important decisions that affect people based on their score on entrepreneurship. To do this, it would be necessary to accumulate more evidence of validity in relation to external variables and test their predictive ability. In sum, although significant progress has been made in the assessment of the entrepreneurial personality in Spain, there is still a long way to go.

DISCUSSION AND CONCLUSIONS

The figure of the entrepreneur is central to the economy of any country, as it constitutes an important source of innovation, employment, productivity and growth. The interest in this figure has evolved over recent decades maintaining, as a common denominator, the person as central to the entrepreneurial process (Baum et al., 2007). Economic and sociological perspectives have contributed substantially to the theoretical development of the entrepreneurial process, while psychology has taken the lead in recent years, being noteworthy for its

TABLE 1
MAIN MEASUREMENT INSTRUMENTS FOR ASSESSING ENTREPRENEURSHIP

Name	Reference	Dimensions
Skills Confidence Inventory [SCI]	Betz, Borgen & Harmon (2005)	Realistic, investigative, artistic, social, enterprising and conventional
General Enterprising Tendency [GET2]	Caird (2006)	Need for achievement, autonomy, determination, creativity and risk taking
Entrepreneurial Aptitude Test [TAI]	Favretto, Pasini & Sartori (2003)	Goal orientation, leadership, adaptation, achievement motivation, personal development, innovation, flexibility and autonomy
Entrepreneurial Intention Questionnaire [EIQ]	Liñán & Chen (2006)	Featuring professional, social value, entrepreneurship and entrepreneurial intention
Cuestionario de orientación emprendedora [COE, questionnaire of entrepreneurial orientation]	Sánchez (2010)	Locus of control, self-efficacy, risk appetite and proactivity
Measure of Entrepreneurial Talents and Abilities [META]	Almeida, Ahmetoglu & Chamorro-Premuzic (2014)	Creativity, opportunism, proactivity and vision
Batería de Evaluación de la Personalidad Emprendedora [BEPE, Battery for the assessment of the entreprising personality]	Muñiz, Suárez-Álvarez, Pedrosa, Fonseca-Pedrero & García-Cueto (2014)	Achievement motivation, risk taking, innovation, autonomy, self-efficacy, stress tolerance, internal locus of control and optimism.

TABLE 2
PSYCHOMETRIC ASSESSMENT OF MEASUREMENT INSTRUMENTS FOR THE ASSESSMENT OF ENTREPRENEURSHIP

Test	Reliability	Validity evidence: content	Validity evidence: construct	Validity evidence: criteria	DIF	Available in Spanish
SCI	✔	✔	✔	✔	-	-
GET2	✔	-	✔	✔	-	-
TAI	✔	-	✔	✔	-	-
EIQ	✔	-	✔	-	-	✔
COE	✔	-	✔	-	-	✔
META	✔	-	✔	✔	-	✔
BEPE	✔	✔	✔	-	✔	✔

Note: DIF = Differential Item Functioning.SCI= *Skills Confidence Inventory;* GET2= *General Enterprising Tendency v2;* TAI= *Entrepreneurial Aptitude Test;* EIQ= *Entrepreneurial Intention Questionnaire;* COE= Cuestionario de Orientación Emprendedora [Questionnaire of Entrepreneurial Orientation]; META= *Measure of Entrepreneurial Talents and Abilities;* BEPE= Batería de Evaluación de la Personalidad Emprendedora [Assessment Battery of Entrepreneurial Personality]

contribution to the evaluation of the entrepreneurial personality. While the existing instruments represent an important advance in terms of measurement, there is still a long way to go. For example, in spite of the boom of Item Response Theory (IRT) in recent years, it is remarkable that only one of the instruments was developed based on this methodological framework (Muñiz et al., 2014). The implementation of IRT in this area would enable us to increase the number of computerised adaptive tests, taking advantage of the many benefits associated with them in terms of effectiveness and efficiency (van der Linden & Glas, 2010). Moreover, there is a notable lack of information provided in the instruments that have already been developed in relation to the analysis of items, evidence of predictive validity, test-retest reliability and especially both differential item functioning (DIF) and bias analysis, deficiencies which are also commonly found in other measuring instruments (Hernández, Tomás, Ferreres & Lloret, 2015).

Internationally, META is probably the measurement instrument that has shown the most validity evidence in recent years, which makes it a suitable tool for evaluating entrepreneurship in adult workers (Ahmetoglu et al, 2011; Almeida et al, 2014; Leutner et al., 2014). Moreover, it has been translated and adapted into ten languages, including Spanish. In the case of Spain, the Battery for Entrepreneurial Personality Assessment (BEPE, Muñiz et al, 2014; Suárez-Álvarez et al, 2014) is noteworthy. This measurement instrument stands out because it offers the joint assessment of the specific characteristics of the entrepreneurial personality as well as being oriented towards adolescents, which enables the early detection of potential entrepreneurs. It also facilitates the assessment of entrepreneurship using 87 items, and has demonstrated adequate psychometric properties including content validity evidence by experts (Suárez-Álvarez et al., 2014) and DIF according to sex (Muñiz et al., 2014). However, it would be necessary to gather more evidence of validity to support the predictive ability of BEPE. Additionally, the recent development of a computerised adaptive version is worth mentioning (BEPE-A; Pedrosa, Suárez-Álvarez García-Cueto & Muñiz, 2015). This instrument, based on the methodological framework of IRT, allows the progressive selection of questions depending on the answers that the participant has given to the preceding items, resulting in a test adapted to the individual (De Ayala, 2009). Using this methodology, the results have shown the ability to assess entrepreneurship accurately with an average of ten items. Adding to this short, effective and rigorous assessment, the fact that it is available in a computerised version opens the possibility of on-line assessment with the benefits that this entails in terms of geographical scope, ease of implementation, and the savings in human, material and financial resources.

Another aspect to highlight is the frequent use of self-report methods. This methodology involves assuming the risk that the person will misrepresent their own answer to fit a certain profile (i.e., social desirability). This type of response bias would directly affect the validity of the decisions that are made based on the scores obtained in the measurment instruments, particularly in personality tests and with significant consequences for people. To solve this problem, various alternatives have been proposed including forced-choice items, where people must choose between two items with similar social desirability (Brown & Maydeu-Olivares, 2012). Good examples of this type of measurement instrument in the context of personality could be WorkFORCE (Naemi, Seybert, Robbins & Kyllonen, 2014) which assesses fit to the work profile and TAPAS (Stark et al., 2014) for selecting personnel in the military context, both developed by the Educational Testing Service. Moreover, as an alternative to self-tests there are situational tests (Olea, Abad & Barrada, 2010) and implicit association tests (Greenwald, Poehlman, Uhlmann & Banaji, 2009). Another interesting alternative is the one currently being conducted by the Psychometrics Centre at the University of Cambridge. Their recent findings include that personality assessment based on computers (i.e., indicators obtained through social networks such as Facebook or Twitter) is more accurate and valid than that obtained by humans (Youyou, Kosinski & Stillwell, 2015).

Future directions in the evaluation of the entrepreneurial personality should be oriented towards improving measurement at different levels (Table 3). First, the use of models to estimate IRT would improve the psychometric properties of the measurement instruments in terms of accuracy. Second, thanks to the use of IRT, computerised adaptive tests could be developed, increasing the effectiveness and efficiency of evaluations compared to the classic format. Third, developing measurement instruments using forced-choice items from IRT would reduce the effects of social desirability in the responses. Fourth, it would be interesting to supplement the data obtained from self-reports with other sources such as implicit association tests or situational tests.

TABLE 3
CURRENT SITUATION AND FUTURE DIRECTIONS IN THE ASSESSMENT OF ENTREPRENEURSHIP

	Now	Future
Psychometric Models	Classical Test Theory	Item Response Theory
Measurement Instruments	General Tests	Computerised Adaptive Tests
Scale	Self-report	Implicit Association Tests, Situational Tests, Forced Choice Items..

ACKNOWLEDGEMENTS

The authors would like to thank Professors José Muñiz and Eduardo García-Cueto for their comments of support and help in previous versions of this manuscript. This research was partly funded by the Ministry of Economy and Competitiveness and the Ministry of Education, Culture and Sports (PSI2011-28638, BES2012-053488 and AP2010-1999).

REFERENCES

Ahmetoglu, G., Leutner, F., & Chamorro-Premuzic, T. (2011). EQ-nomics: Understanding the relationship between individual differences in trait emotional intelligence and entrepreneurship. *Personality and Individual Differences, 51*(8), 1028-1033. doi: 10.1016/j.paid.2011.08.016.

Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211. doi:10.1016/0749-5978(91)90020-t

Almeida, P.I.L., Ahmetoglu, G. & Chamorro-Premuzic, T. (2014). Who wants to be an entrepreneur? The relationship between

vocational interests and individual differences in entrepreneurship. *Journal of Career Assessment, 22*(1), 102-112. doi: 10.1177/1069072713492923

Altinay, L., Madanoglu, M., Daniele, R., & Lashley, C. (2012). The influence of family tradition and psychological traits on entrepreneurial intention. *International Journal of Hospitality Management, 31*(2), 489-499. doi: 10.1016/j.ijhm.2011.07.007.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Baum, J.R., Frese, M., Baron, R.A. & Katz, J.A. (2007). Entrepreneurship as an Area of Psychology Study: An Introduction. In J.R. Baum, M. Frese & R.A. Baron (Eds.), *The Psychology of Entrepreneurship* (pp. 1-18). London: Lawrence Erlbaum.

Betz, N.E., Borgen, F.H. & Harmon, L.W. (2005). *Manual for the Skills Confidence Inventory: Revised Edition*. Mountain View, CA: Consulting Psychologists Press.

Brandstatter, H. (2011). Personality aspects of entrepreneurship: A look at five meta-analyses. *Personality and Individual Differences, 51*(3), 222-230. doi: 10.1016/ j.paid.2010.07.007

Brown, A. & Maydeu-Olivares, A. (2012). How IRT can solve problems of ipsative data in forced-choice questionnaire. *Psychological Methods, 18*(1), 36-52.

Caird, S. (2006). General Measure of Enterprising Tendency Version 2 (GET2). In T. Mazzarol (Ed.), *Entrepreneurship and Innovation* (pp. 247-266). Australia: Tilde University Press.

Cantillon, R. (1756). *Essai sur la Nature du Commerce en General [An Essay on the General Nature of Business]*. London: Fletcher Gyles.

Chell, E. (2008). *The entrepreneurial personality: A social construction*. New York: Routledge.

Covin, J.G. & Wales, W.J. (2012). The measurement of entrepreneurial orientation. *Entrepreneurship Theory and Practice, 36*(4), 677-702. doi: 10.1111/j.1540-6520.2010.00432.x

De Ayala, R.J. (2009). *The theory and practice of item response theory*. New York: Guilford Press.

Evers, A., Muñiz, J., Hagemeister, C., Hstmælingen, A., Lindley, P., Sjöberg, A., & Bartram, D. (2013). Assessing the quality of tests: Revision of the EFPA review model. *Psicothema, 25*(3), 283-291.

Favretto, G., Pasini, M. & Sartori, R. (2003). Attitudine imprenditoriale e misura psicométrica: il TAI [The entrepreneur attitude and psychometric measurement: the TAI]. *Revista di Psicologia del Lavoro e dell´Organizzazione, 9,* 271-282.

Frese, M., & Fay, D. (2001). Personal initiative (PI): An active performance concept for work in the 21st century. In B.M. Staw & R.M. Sutton (Eds.), *Research in Organizational Behavior* (Vol. 23, pp. 133-187). Amsterdam: Elsevier Science.

Furnham, A. (2002). Personality at work: individual differences in the workplace. London: Routledge.

Global Entrepreneurship Monitor (2015). *Informe GEM España 2014 [GEM Report Spain 2014]*. Retrieved on 17 June 2015 from http://www.gem-spain.com/wp-content/uploads/2015/04/GEM_es_2014.pdf

Greenwald, A.G., Poehlman, T.A., Uhlmann, E.I. & Banaji, M.R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*, 17-41.

Hernández, A., Tomás, I., Ferreres, A. & Lloret, S. (2015). Tercera evaluación de tests editados en España [Third evaluation of tests published in Spain]. *Papeles del Psicólogo, 31*(1), 1-8.

Hisrich, R., Langan-Fox, J., & Grant, S. (2007). Entrepreneurship research and practice: A call to action for psychology. *American Psychologist, 62*(6), 575-589.

Katz, J.A. (2007). Education and training in Entrepreneurship. In J.R. Baum, M. Frese & R.A. Baron (Eds.), *The Psychology of Entrepreneurship* (pp. 209-235). London: Lawrence Erlbaum.

Kirzner, I. (1973). *Competition and entrepreneurship*. Chicago: Chicago University Press.

Laguna, M. (2013). Self-efficacy, self-esteem, and entrepreneurship among the unemployed. *Journal of Applied Social Psychology, 43*(2), 253-262. doi: 10.1111/j.1559-1816.2012.00994.x

Lanero, A., Vázquez, J.L. & Muñoz-Adánez, A. (2015). Un modelo social cognitivo de intenciones emprendedoras en estudiantes universitarios [A cognitive social model of entrepreneurial intentions in university students]. *Anales de Psicología, 31*(1), 243-259. doi: 10.6018/analesps.31.1.161461

Leutner, F., Ahmetoglu, G., Akhtar, R. & Chamorro-Premuzic, T. (2014). The relationship between the entrepreneurial personality and the Big Five personality traits. *Personality and Individual Differences, 63*, 58-63. doi: 10.1016/j.paid.2014.01.042

Liñán, F. & Chen, Y.W. (2006). *Testing the entrepreneurial intention model on a two-country sample*. Bellaterra: Universitat Autònoma de Barcelona.

Lumpkin, G.T. (2007). Intrapreneurship and innovation. In J.R. Baum, M. Frese & R. Baron (Eds.), *The Psychology of Entrepreneurship* (pp. 237-264). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Marshall, A. (1890). *Principles of Economics: An Introductory Volume*. London: MacMillan.

Moriano, J.A., Palací, J.F. & Morales, J.F. (2006). Adaptación y validación en España de la escala de autoeficacia emprendedora [The adaptation and validation in Spain of the entrepreneurial self-efficacy scale]. *Revista de Psicología Social, 21*(1), 51-64.

Muñiz, J., Elosua, P. & Hambleton, R. (2013). Directrices para la traducción y adaptación de los tests: segunda edición [Guidelines for the translation and adaptation of tests: Second Edition]. *Psicothema, 25*(2), 151-157. doi:10.7334/psicothema2013.24

Muñiz, J., Suárez-Álvarez, J., Pedrosa, I., Fonseca-Pedrero, E. & García-Cueto, E. (2014). Enterprising personality profile in youth: Components and assessment. *Psicothema, 26*(4), 545-553. doi: 10.7334/psicothema2014.182

Naemi, B., Seybert, J., Robbins, S. & Kyllonen, P. (2014). *Examining the WorkFORCE™ Assessment for Job Fit and Core Capabilities of the FACETS™ Engine (Research Report)*. Princeton, NJ: Educational Testing Service.

Newton, J.H. & McGrew, K.S. (2010). Introduction to the special issue: current research in Catell-Horn-Carrol-based assessment. *Psychology in the Schools, 47*(7), 621-634. doi: 10.1002/pits.20495

OCDE (2014). *Entrepreneurship at a Glance 2014*. Retrieved on 17 June 2015 from http://www.OCDE-ilibrary.org/docserver/download/3014031e.pdf?expires=1435 052455&id=id&accname=guest&checksum=16E3E11BDE0005B81D5DC49095A37D7E

OCDE/The European Commission (2013). *The missing entrepreneurs: Policies for inclusive entrepreneurship in Europe*. Brussels: OECD

Publishing. doi: 10.1787/9789264188167-en.

Olea, J., Abad, F.J. & Barrada, J.R. (2010). Tests informatizados y otros nuevos tipos de tests [Computerised tests and other new types of test]. *Papeles del Psicólogo, 31*(1), 94-107.

Pedrosa, I., Suárez-Álvarez, J., García-Cueto, E. & Muñiz, J. (2015). *Development of a Computerized Adaptive Test for Enterprising Personality Assessment in youth*. Article sent for publication.

Pedrosa, I., Suárez-Álvarez, J., García-Cueto, E. (2013). Evidencias sobre la Validez de Contenido: Avances Teóricos y Métodos para su Estimación [Evidence for Content Validity: Theoretical Developments and Estimating Methods]. *Acción Psicológica, 10*(2), 3-20.

Rauch, A., & Frese, M. (2000). Psychological approaches to entrepreneurial success. A general model and an overview of findings. In C.L. Cooper & I.T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology* (pp. 101-142). Chichester: Wiley.

Rauch, A. & Frese, M. (2007a). Born to be an entrepreneur? Revisiting the personality approach to entrepreneurship. In J.R. Baum, M. Frese & R.J. Baron (Eds.), *The psychology of entrepreneurship* (pp. 41-65). Mahwah, NJ: Erlbaum.

Rauch, A. & Frese, M. (2007b). Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' personality traits, business creation, and success. *European Journal of Work and Organizational Psychology, 16*(4), 353-385. doi: 10.1080/13594320701595438

Sánchez, J.C. (2010). Evaluación de la personalidad emprendedora: Validez factorial del cuestionario de orientación emprendedora (COE) [Evaluation of the entrepreneurial personality: Factorial validity of the entrepreneurial orientation questionnaire]. *Revista Latinoamericana de Psicología, 42*(1), 41-52.

Sánchez, J.C. (2011). Entrepreneurship as a legitimate field of knowledge. *Psicothema, 23*(3), 427-432.

Sánchez, J.C., Carballo, T. & Gutiérrez, A. (2011). The entrepreneur from a cognitive approach. *Psicothema, 23*(3), 433-438.

Sandilands, D., Oliveri, M.E., Zumbo, B., & Ercikan, K. (2013). Investigating sources of differential item functioning in international large-scale assessment using a confirmatory approach. *International Jounral of Testing, 13*(2), 152-174.doi: 10.1080/15305058.2012.690140

Say, J. B. (1803). *Traité d´économie politique [Treaty of Political Economy]*. Paris: Calmann-Lévy Éditeur.

Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, MA: Harvard University Press.

Shapero, A. & Sokol, L. (1982). The Social Dimensions of Entrepreneurship. In C.A. Kent, D.L. Sexton & K.H. Vesper (Eds.), *Encyclopedia of Entrepreneurship* (pp. 72-90). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Stark, S., Chernyshenko, O.S., Drasgow, F., Nye, C.D., White, L.A., Heffner, H., & Farmer, W.L. (2014). From ABLE to TAPAS: A new generation of personality tests to support military selection and classification decisions. *Military Psychology, 26*(3), 153-164. doi: http://dx.doi.org/10.1037/mil0000044

Suárez-Álvarez J., Campillo-Álvarez A., Fonseca-Pedrero E., García-Cueto E., & Muñiz J. (2013). Professional training in the workplace: The role of achievement motivation and locus of control. *The Spanish Journal of Psychology, 16*, E35.http://dx.doi.org/10.1017/sjp.2013.19

Suárez-Álvarez, J., Fernández-Alonso, R., & Muñiz, J. (2014). Self-concept, motivation, expectations and socioeconomic level as predictors of academic performance in mathematics. *Learning and Individual Differences, 30,* 118-123. doi: doi:10.1016/j.lindif.2013.10.019

Suárez-Álvarez, J., Pedrosa, I., García-Cueto, E. & Muñiz, J. (2014). Screening Enterprising Personality in Youth: An empirical model. *Spanish Journal of Psychology, 17*(E60). doi: 10.1017/sjp.2014.61

Suárez-Álvarez, J., Pedrosa, I., García-Cueto, E. & Muñiz, J. (in press). New validity evidence to support Locus of Control bi-dimensionality. *Anales de Psicología.*

Suliman, A.M., & Al-Shaikh, F.N. (2007). Emotional intelligence at work: Links to conflict and innovation. *Jordan Employee Relations, 29*(2), 208-220.

Tyszka, T., Cieslik, J., Domurat, A. & Macko, A. (2011). Motivation, self-efficacy, and risk attitudes among entrepreneurs during transition to a market economy. *The Journal of Socio-Economics, 40*, 124-131.

Unger, J.M., Rauch, A., Frese, M. & Rosenbusch, N. (2011). Human capital and entrepreneurial success: A meta-analytical review. *Journal of Business Venturing, 26*(3), 341-358. doi: 10.1016/j.jbusvent.2009.09.004

van der Linden, W.J & Glas, C.A.W. (2010). *Elements of Adaptive Testing*. New York, NY: Springer.

Veciana, J.M. (1999). Creación de empresas como programa de investigación científica [The creation of companies as a scientific research program]. *Revista Europea de Dirección y Economía de la Empresa, 8*(3), 11-36.

Volery, T., Muller, S., Oser, F., Naepflin, C. & del Rey, N. (2013). The Impact of Entrepreneurship Education on Human Capital at Upper-Secondary Level. *Journal of Small Business Management, 51*(3), 429-446. doi: 10.1111/jsbm.12020

Ward, T.B. (2004). Cognition, creativity, and entrepreneurship. *Journal of Business Venturing, 19*, 173-188.

Youyou, W., Kosinski, M. & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences, 112*(4), 1036-1040. doi:10.1073/pnas.1418680112

Zhao, H., Seibert, S.E. & Lumpkin, G.T. (2010). The Relationship of Personality to Entrepreneurial Intentions and Performance: A Meta-Analytic Review. *Journal of Management, 36*(2), 381-404. doi: 10.1177/0149206309335187

Papeles del Psicólogo, 2016. Vol. 37(1), pp. 69-73
http://www.papelesdelpsicologo.es

ASSESSMENT OF QUALITY OF LIFE: PRESENT AND FUTURE METHODOLOGICAL CHALLENGES

Isabel Benítez
Universidad de Granada. Tilburg University (Países Bajos)

El impacto de la Calidad de Vida en áreas como la salud, el rendimiento escolar o la participación social ha estimulado el desarrollo de distintas aproximaciones que han tratado de abordar tanto la definición como la evaluación de este constructo. Sin embargo, a pesar de los beneficios que supone el enfoque multidisciplinar, esta diversificación ha impedido alcanzar una definición única del constructo y, por tanto, un instrumento o procedimiento de evaluación consensuado. El objetivo de este estudio es plantear los retos metodológicos que afectan al estudio de la Calidad de Vida en la actualidad. Se presenta una breve descripción de la evolución del constructo en los distintos ámbitos, los avances más novedosos y los planteamientos que guiarán la investigación futura en marco nacional e internacional.
Palabras clave: *Calidad de vida, Evaluación, Avances metodológicos,, Psicología transcultural.*

The growing importance of quality of life in diverse domains, such as health, school performance and social participation, has led to the development of new conceptualisations and assessments of the construct. This diversity of perspectives brings about many benefits, but it also creates an obstacle for the formulation of a single unifying definition of the construct and, therefore, an agreed instrument or assessment framework. The aim of this study is to discuss the current methodological challenges in the measurement of quality of life. Firstly, we provide a brief description of the construct as defined in various areas, then we examine the new methodological developments and different applications. We also present an overview of the different possibilities for future developments in defining and measuring quality of life in national and international studies.
Key words: *Quality of Life, Assessment, Methodological challenges, Cross-cultural Psychology.*

AN OVERVIEW OF THE CONCEPT OF QUALITY OF LIFE

According to the *Diccionario de la Real Academia Española* [*Dictionary of the Royal Spanish Academy*] (DRAE), quality of life refers to "the set of conditions that help to make life enjoyable and valuable." This simple definition that would fit the popular idea of quality of life, begins to raise difficulties when we wish to specify what "to make life enjoyable and valuable" means or when we wish to determine the "set of conditions" that favours this state. On both counts, personal experiences, ambitions or expectations (among other things) introduce a subjective factor that makes the concept difficult to extend to multiple people, especially if these people come from different socio-demographic or cultural groups. This situation becomes clear if we think of our inner circle, where we can easily identify people we know whose priorities are far from our own, which clearly leads to a different assessment of what gives us quality of life. The same problem occurs on a large scale, when studies aim to assess quality of life on a national level, and of course in international studies aimed at comparing participants from different countries.

In common parlance, the term quality of life is used to refer to different aspects such as satisfaction with specific conditions, commodities at a socio-economic level, facilities for meeting the needs of daily life, or even happiness. How often we have heard someone say "This is quality of life", and more importantly, in how many different contexts and with how many different nuances? Precisely this familiarity with the concept is one of the reasons why quality of life is, as Campbell, Converse, and Rodgers (1916) noted, something that many people talk about but nobody knows how to define clearly. Many years later, Barofsky (2012) also indicated that the "everydayness" of the term is a constraint to both the definition and the measurement of this construct.

Correspondence: Isabel Benítez. *Universidad de Granada. Facultad de Psicología. Campus de Cartuja s/n. 18071 Granada. España. E-mail: ibenitez.baena@gmail.com*

This situation poses a challenge in science where, ideally, "personalised" definitions should be left behind in order to establish common definitions and agreed criteria to guide the activity of the research community. Therefore, it is in this context where concern for the systematic study of the concept of quality of life as well as its assessment emerged in the 60s (Gómez & Sabeh, 2001). Although there has always been interest in the construct, in this period a change of perspective occurred in which the idea of proposing solutions subsequent to the emergence of the problem was replaced by a concept of social change that seeks to promote an improvement in society (Casas, 2004). In other words, the idea of quality of life became an activity aimed at promoting positive behaviours that improve people's situations.

From then until now, the concept of quality of life has been used in various fields such as psychology, health, education, economics or politics, which has led to the study of the construct at different levels of generalisation. While psychology has focused on the individual aspects of people (Aroila, 2003), economics and politics have addressed issues concerning society or the community, considering quality of life as the indispensable motor of innovation for social evolution (Yúdice, 2002). On the other hand, in healthcare and education both fronts have been covered, focusing on groups of people with specific circumstances, such as a specific pathology (e.g., Lara, Ponce, & de la Fuente, 1995), or special educational needs (e.g., Gómez-Vela, Verdugo, & González-Gil, 2007). This diversity of approaches has meant that the meaning of quality of life is complex and has definitions that adjust to the focus of interest in each case. However, in an attempt to reduce this diversity, two main branches emerged that divide the research on quality of life in health sciences and social sciences. From the common objective of knowing the most important aspects for peoples' lives and their influence on the different life areas of human beings, the two perspectives present an approach to the concept that incorporates different nuances.

On the one hand, in the healthcare field, the concept emerges of quality of life related to health (HRQOL, or Health-Related Quality of Life) which was initially defined in 1948 by the World Health Organization (World Health Organization, WHO) as a state of complete physical, mental and social well-being that goes beyond the mere absence of disease (WHO, 1998). Some of the most relevant studies in this approach have shown the powerful influence of quality of life in such significant aspects as patient adherence to prescribed treatments (Carballo et al., 2004) or the evolution of disease (Jones et al, 2006; Lemonnier et al, 2014)..

On the other hand, in the field of social science, efforts have been focused on trying to unify the concept of quality of life which, as Veenhoven (2000) indicates, has been used interchangeably with other constructs such as well-being or happiness. In this context, two of the basic traditions described by Schwartzmann (2003) are included: research into the concept of happiness in psychology and the study of social indicators in sociology. The main concern in the social sciences is people's environment, highlighting the more private component of quality of life; and therefore the studies address aspects such as its influence on social participation (Nakamura et al, 2014; Wendel-Vos, Schuit, Tijhuis, & Kromhout, 2004) or the development of personal and professional relationships (Pinquart & Sorenser, 2000). Quality of life is then defined as the "experience that people have of their own ways and conditions of life" (Casas, 2004, p. 309); referring to objective factors, such as living conditions in themselves, and subjective elements that reflect people's own perception that they have of the situation. In line with the controversial role of subjectivity mentioned above, the research is characterised by proposing assessment focused on obtaining indicators of the presence or absence of subjective well-being in people. In other words, the elements considered in the traditional assessment focused on objective conditions are maintained but interest is moved to the discovery of the private and subjective aspects. In this line, Schalock and Verdugo (2002) describe the indicators of quality of life based on three personal dimensions that reflect the well-being of the person: specific perceptions, behaviours and conditions. Meanwhile Casas (2011) distinguishes two social indicators of subjective well-being: overall life satisfaction and satisfaction with specific or peripheral aspects.

These definitions reflect the efforts made to clarify the concept of quality of life and make it easier for researchers to outline new studies by differentiating and limiting the aspects of interest in each area of study. However, at the same time they present other difficulties such as the measurement of the construct. The same diversity described above is evident in the tools available for evaluating quality of life, as shown by previous reviews of the existing instruments for the evaluation of quality of life (Blanco & Chacón, 1985, Bowling, 1991).

Following the approach of social sciences, possibly the closest to the readers, we find instruments that refer to the division of quality of life into general and specific aspects. Among the general aspects, the main focuses of attention have been subjective well-being (*The Satisfaction with Life Scale*; Diener, Emmons, Larsen, & Griffin, 1985), happiness *(Happiness Measures*; Fordyce, 1988) and overall satisfaction (*Life Satisfaction Scale*; Huebner, 1994). The assessment of specific aspects has focused on the investigation of the main life areas of individuals (Cummins, 2003; Cummins, Eckersley, Van Pallant, Vugt & Misajon, 2003). For example, Zabriskie and McCormick (2003) used an adapted version of the *Satisfaction With Life Scale* to assess satisfaction with family life (*Satisfaction with Family Life Scale*); Bowling and Hammond (2008) review the properties of the *Michigan Organizational Assessment Questionnaire* designed to measure job satisfaction (*Michigan Organizational Assessment Questionnaire Job Satisfaction Subscale*); and Heyland et al. (2002) investigated satisfaction with the services received in healthcare settings.

This situation reflects the current reality in the research on quality of life, which has led to decisions being made to approach the study of this construct, from different fronts both nationally and internationally, as described in the next section.

THE STUDY OF QUALITY OF LIFE

Despite the clear influence of the methodological issues in the research on quality of life, the ultimate goal of the assessment is to draw conclusions regarding this variable, and it is on this point that the national and international studies focus. In Spain (as in other countries), much of the research focuses on the assessment of specific groups or content-specific assessment, while at the international level the aim is to establish "universal" indicators of quality of life for comparative purposes. Described below are some of the current lines of work in the two contexts.

Quality of life in international and transcultural studies

On the international scene, various organisations have proposed approaches intended to assess different countries or groups in a standardised way. The healthcare aspect is represented by the WHO, which in recent years has tried to introduce the individual perception of patients in relation to their quality of life, as part of the assessment of their functionality (WHO, 1994). In its classification models, the WHO proposes indicators that describe health conditions globally, such as the International Classification of Functioning, Disability and Health. This model includes the concept of functionality, which incorporates personal and environmental factors and their interaction with the disease suffered by the individual.

Studies of the social aspect are associated with survey research whereby items are administered to different groups which are subsequently compared in relation to the amount of the variable. In this context, the main limitation comes from the difficulty of establishing common indicators to the different groups evaluated and ensuring equivalence in the definition of these indicators. The European Statistical System Committee (ESSC) attempted to address this difficulty by developing, in November 2011, a list of dimensions to measure the quality of life in the European Union. These dimensions are divided into specific indicators so that the assessment of these indicators provides, according to this approach, information on people's quality of life. For example, the dimension "overall experience with life" covers three themes: satisfaction with life, emotions and goals. These themes are subdivided into indicators that are formulated in terms of items, such that the application of those items would measure the construct quality of life. Other dimensions are "leisure and social interactions", "environment" and "physical and financial security," which are defined the same way in terms of themes and indicators as described above.

This perspective has been accepted by many researchers since it offers a broad framework in which studies of various types have a place. In fact, numerous international studies with comparative objectives have followed the guidelines proposed by the ESSC. Some examples are the *European Values Study* (EVS), the *European Social Survey* (ESS), the

European Quality of Life Survey (EQLS), and the *World Values Survey* (WVS). In all of these studies, the dimensions and indicators proposed for generating quality of life items are used. However, this scheme does not meet the needs of researchers concerned with the subjective sphere.

Another approach to quality of life from the social perspective is that proposed by the Organisation for Economic Co-operation and Development (OECD), which has presented one of the most universal panoramas to date by creating the *Better Life Index*. Based on the responses of participants from different countries on five continents, this index assesses the most relevant aspects for the citizens, but this also leaves out the more private aspects of the assessment. Therefore, despite international efforts to reach a satisfactory definition for the different areas of study, the inclusion of the subjective aspects that capture personal perceptions of quality of life has not yet been achieved.

Beyond the definition of the construct, in the international framework another of the most important challenges relates to achieving equivalence in measurement. Ensuring the equivalence of the responses provided by different groups is, as indicated by Van de Vijver and Matsumoto (2011), the only way to make valid comparisons between the groups assessed. Both the level of equivalence and the presence of bias have been analysed previously in the context of quality of life (Meng, King-Kallimanis, Gum & Wamsley, 2013; Scott et al, 2009a.) The construct equivalence is particularly important in this area, as although subjective well-being can be conditioned by objective conditions, an individual assessment is likely to be more determined by the specific circumstances of each person. That is, two people in the same circumstances in two different countries could assess their quality of life differently in relation to continua such as wealth-poverty or health-disease. This fact is the main challenge in the international arena: ensuring that the quality of life indicators that are established are independent of the context and/or focus attention on subjective aspects that must also be interpreted through the groups. In other words, the assessment of subjective elements, both in themselves and in a comparative scenario, currently constitute the biggest challenge on an international level.

Quality of life in specific contexts

In Spain, several research teams are working directly or indirectly in the study of quality of life. Two main objectives can be identified: the study of the quality of life construct itself, which is approached from the assessment of groups selected based on demographics (sex or age) or groups of people with specific circumstances (patients, caregivers, etc.); and the creating of instruments that capture the aspects of interest.

Studies unifying both concerns currently represent the most complex and challenging option. For example, one of the most important lines in the study of quality of life is currently focused on the evaluation of subjective well-being in children and adolescents. The main objective is to understand the determinants of subjective well-being in this group. However, the study of the concept involves complex methodological challenges. For example, the assessment of children and adolescents involves conducting longitudinal studies or having instruments adapted to the characteristics of the participants in different age groups. In both cases we are faced with situations in which participants experience a developmental change in the course of the study, and therefore, it is necessary that the instruments used capture the same content in all administrations, so that it is possible to draw conclusions about the changes associated with subjective well-being beyond the changes resulting from growth. This means that researchers must generate tools that capture equivalent indicators in groups that differ in their demographics and, in all likelihood, in their cognitive abilities to cope with the task set. Recent research seeks to respond to this situation by including graphic materials that have proved their usefulness in obtaining information from younger participants (Nic Gabhainn & Sixsmith, 2006).

CHALLENGES IN THE STUDY OF QUALITY OF LIFE

In view of the above, one might ask what the most important challenges are in the study of quality of life. This approach requires us to return to the most recent studies, extract their limitations and propose innovative approaches in relation to the critical points described above.

First, both the definition and assessment of quality of life have been widely addressed in studies that have attempted to gather different formulations and propose comprehensive definitions (Blanco & Chacón, 1985, Bowling, 1991; Casas, 2004; Gómez & Sabeh, 2001). However, the need to respond to concerns from different fields of study has hindered the achievement of proposals accepted by the scientific community as a whole. To overcome this limitation involves possibly returning to the beginning with the conceptualisation and extracting the common concerns of psychologists, sociologists, health workers, educators and other professionals interested in the study of quality of life. Despite the ambitious project, psychometrics can provide a theoretical and methodological framework to guide the process of defining the construct (Crocker & Algina, 1986), establishing the necessary steps to collect and integrate the different perspectives in a detailed definition of the dimensions and indicators that enables the generation of items that measure people's quality of life.

Another challenge for research in this field is to create versions that permit the assessment of groups that speak different languages, and to establish equivalence in the measurements made using these instruments. As mentioned in previous sections, this task has mainly been approached by organisations interested in international comparison. However, ensuring equivalence, or in other words the absence of bias between groups, is still a challenge. Bias refers to the presence of elements in the measuring instruments that do not have the same meaning across groups (Poortinga, 1989). The studies of bias in the context of quality of life have so far pursued different objectives, such as, for example, examining the adequacy of the translated versions of assessment instruments (Scott et al., 2009a), or obtaining validity evidence of the usefulness of a tool to assess different groups (Rendas-Baum, Yang, Varon, Bloudek, DeGryse & Kosinski, 2014). However, as Scott et al. (2009b) suggest, there is still no agreement on the nature and impact of bias in assessments of the quality of life.

In this regard, recent studies have tried to determine both the impact of and the elements generating bias in assessments of quality of life. To do this, mixed designs that combine quantitative and qualitative methodologies represent the most promising option in recent years, as they integrate findings of a different nature with the aim of achieving a more global and sophisticated view of the phenomenon studied (Tashakkori & Teddlie, 1998). Benítez, Van de Vijver and Padilla (in press) used statistical techniques to detect bias at item level and cognitive interviews in order to explain its causes. These authors describe three main sources of bias: linguistic (words and expressions that do not have the same meaning in the different versions), contextual (differences in interpreting the nuances connected with the agreements established in

each country or culture) and substantive (differential interpretations due to the specific circumstances of the groups or countries evaluated). Also Benítez, He, Van de Vijver, and Padilla (under review) used a mixed design to interpret the presence of bias related to cultural trends expressed during the process of responding to the items. Specifically, this study describes the causes of differences in the frequencies of choice of specific alternatives between the groups.

In addition to the mixed research, the use of qualitative procedures in itself represents an important contribution to the study of bias in evaluating the quality of life. To date, its implementation has pursued two fundamental objectives: to identify and understand the origin of the differences between the groups and to provide a comprehensive view of the bias integrating the various levels (item, method and construct). Among the studies of the first group is the work by Smits et al. (2005) in which differential interpretations of the symptoms related to mental health (considered negative indicators of quality of life) among participants from Turkey and Morocco are described. In the second group, Benítez, Padilla and Van de Vijver (2015) illustrate a comprehensive evaluation of bias using cognitive interviews. The authors provide qualitative evidence of the presence of differences between groups unrelated to the construct as well as specific elements of the groups demonstrating the non-equivalent composition of quality of life construct through the groups evaluated.

CONCLUSIONS AND DISCUSSION

The aim of this study was to present an overview of the research into quality of life as well as the past, present and future challenges in this field. Both the theoretical aspects and the revised empirical studies have demonstrated the importance that quality of life has in society today, and how interest in the impact of quality of life has spurred the development of rigorous scientific studies that have contributed to significant progress in recent years.

Following the review and reflection made in this paper, several conclusions can be drawn. The first clear result, derived from the contents described, is the need to continue research into quality of life leveraging the efforts made so far. That is, despite the diversification and the multidisciplinary nature surrounding the construct, there are important points of departure that must be a reference for future research. For example, the consensus definitions in the different areas represent the most advanced theoretical bases to date, and as such, should be considered in future studies pending further investigations that incorporate the aspects that have been most critical, such as those related to the more subjective aspects.

Likewise, the assessment tools available gather and capture the most current theoretical approaches, so their application may be relevant in studies that replicate the conditions for which the instrument was originally created. However, in the case of comparative studies it is necessary to ensure equivalence in the construct measured in the different groups involved, as well as the lack of bias at different levels (Van de Vijver & Matsumoto, 2011). To do this, it is proposed to implement mixed designs that include a statistical assessment of the bias and a qualitative review of the interpretations made by the participants, so we have information on the aspects being captured differentially through the groups.

Currently, the literature on quality of life is extensive, although it is essential to promote new studies that advocate a "universal" definition of the construct. According to the limitations observed in the previous research, this definition would consist of a model of dimensions, indicators and relationships that would enable the generation of a set of items that capture the quality of life construct in a standardised and non-biased way. Apart from these purely methodological challenges, future research should also address the substantive requirements. This would include studies aimed at promoting quality of life in patients with specific health conditions, subjective well-being in children and adolescents, or other investigations not mentioned above focused on quality of life related to ecological behaviours, or with personal characteristics such as attachment to the place of residence or belonging to majority or minority groups (Benitez, He & Adams, 2015).

Despite the idealism of the reflections raised, the progress observed in recent years shows the interest and dedication of the various professionals making headway in the research into quality of life. Therefore, this paper aims to be a starting point to guide and lead the steps planned from the various fronts toward a common goal.

REFERENCES

Aroila, R. (2003). Calidad de vida: Una definición integradora [Quality of life: An inclusive definition]. *Revista Latinoamericana de Psicología, 35*(2), 161-164.

Barofsky, I. (2012). Can quality or quality-of-life be defined? *Quality of Life Research, 21*(4), 625-631.

Benítez, I., He, J., & Adams, B. (2015). *The role of environmental satisfaction in wellbeing: A comparison between mainstreamers and minority groups*. XII Congreso de Psicología Ambiental [Congress of Environmental Psychology] (pp. 288). Granada: University of Granada.

Benítez, I., He, J., Van de Vijver, A.J.R., & Padilla, J.L. (Under review). Linking Response Styles to Response Processes: A Cross-Cultural Mixed Methods Approach. *International Journal of Psychology*.

Benítez, I., Padilla, J.L., & Van de Vijver, A.J.R. (2015). *The qualitative assessment of bias: Contributions of Cognitive Interviewing methodology to the bias definition*. 13th European Conference on Psychological Assessment (pp. 27). Zurich: University of Zurich.

Benítez, I., Van de Vijver, A.J.R., & Padilla, J.L. (In press). A Global Strategy for Investigating Item Bias in Cross-Cultural Studies: A Mixed Methods Approach Integrating Differential Item Functioning and Cognitive Interviews. *Journal of Mixed Methods Research*.

Blanco, A., & Chacón, F. (1985). La evaluación de la calidad de vida [The assessment of quality of life]. In F.F. Morales, A.B. Abarca, C.H. Casal, & J.M.F Dols, *Psicología Social Aplicada [Applied Social Psychology]* (pp. 183-210). Bilbao. Desclée de Brouwer.

Bowling, A. (1991). *Measuring health: a review of quality of life measurement scales*. Milton Keynes: Open University Press.

Bowling, N. A., & Hammond, G. D. (2008). A meta-analytic examination of the construct validity of the Michigan Organizational Assessment Questionnaire Job Satisfaction Subscale. *Journal of Vocational Behavior, 73*(1), 63-77.

Campbell, A., Converse, P.E., & Rodgers, W.L. (1916). *The quality of American life*. New York, NY: Russell Sage Foundation.

Carballo, E., Cadarso-Suárez, C., Carrera, I., Fraga, J., de la Fuente, J., Ocampo, A., Ojea, R., & Prieto, A. (2004). Assessing relationships between health-related quality of life and adherence to antiretroviral therapy. *Quality of Life Research, 13,* 587-599.

Casas, F. (2004). De afrontar problemas sociales a promover la calidad de vida [From tackling social problems to promoting quality of

life]. *Revista de Historia de la Psicología, 25*(4), 305-322.

Casas, F. (2011). Subjective social indicators and child and adolescent well-being. *Child Indicators Research, 4*(4), *555-575*.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando, FL: Harcourt Brace Jovanovich.

Cummins, R.A. (2003). Normative life satisfaction: Measurement issues and a homeostatic model. *Social Indicators Research, 64,* 225-256.

Cummins, R.A., Eckersley, R., Van Pallant, J., Vugt, J., & Misajon, R. (2003). Developing a national index of subjective well-being: The Australian Unity Well-being Index. *Social Indicators Research, 64,* 159-190.

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71-75.

Fordyce, M. W. (1988). A review of research on the happiness measures: A sixty second index of happiness and mental health. *Social Indicators Research, 20*(4), 355-381.

Gómez, M., & Sabeh, E. (2001). *Calidad de vida: Evolución del concepto y su influencia en la investigación y la práctica [Quality of life: The evolution of the concept and its influence on research and practice].* Salamanca: Instituto Universitario de Integración en la Comunidad [University Institute of Integration in the Community], Faculty of Psychology, University of Salamanca.

Gómez-Vela, M., Verdugo, M. Á., & González-Gil, F. (2007). Calidad de vida y autoconcepto en adolescentes con necesidades educativas especiales y sin ellas [Quality of life and self-concept in adolescents with and without special education needs]. *Infancia y Aprendizaje, 30*(4), 523-536.

Heyland, D. K., Rocker, G. M., Dodek, P. M., Kutsogiannis, D. J., Konopad, E., Cook, D. J., Peters, S., Tranmer, J.E., & O'Callaghan, C.J. (2002). Family satisfaction with care in the intensive care unit: Results of a multiple center study. *Critical Care Medicine, 30*(7), 1413-1418.

Huebner, E. S. (1994). Preliminary development and validation of a multidimensional life satisfaction scale for children. *Psychological Assessment, 6*(2), 149.

Jones, P.B., Barnes, T.R.E., Davies, L., Dunn, G., Lloyd, H., Hayhurst, K.P., Murray, R.M., Markwick, A., & Lewis, S.W. (2006). Randomized Controlled Trial of the Effect on Quality of Life of Second- vs First-Generation Antipsychotic Drugs in Schizophrenia Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1). *Archives of General Psychiatry, 63,* 1079-1087.

Lara, M.C., Ponce, S., & de la Fuente, J. R. (1995). Conceptualización y medición de la calidad de vida de pacientes con cáncer [Conceptualisation and measurement of quality of life in cancer patients]. *Revista de Investigación Clínica, 47*(4), 315-327.

Lemonnier, I., Guillemin, F., Arveux, P., Clément-Duchêne, C., Velten, M., Woronoff-Lemsi, M.C., Jolly, D., & Baumann, C. (2014). Quality of life after the initial treatments of non-small cell lung cancer: a persistent predictor for patients' survival. *Health and Quality of Life Outcomes, 12,* 73-83.

Meng, H., King-Kallimanis, B.L., Gum, A., & Wamsley, B. (2013). Measurement bias of the SF-36 Health Survey in older adults with chronic conditions. *Quality of Life Research, 22,* 2359-2369.

Nakamura, P.M., Teixeira, I.P., Smirmaul, B.P.C., Sebastião, E., Papini, C.B., Gobbi, S., & Kokubun, E. (2014). Health related quality of life is differently associated with leisure-time physical activity intensities according to gender: a cross-sectional approach. *Health and Quality of Life Outcomes, 12,* 98-108.

Nic Gabhainn, S., & Sixsmith, J. (2006). Children photographing well being: facilitating participation in research. *Children & Society, 20*(4), 249-259.

Pinquart, M., & Sorensen, S. (2000). Influences of Socioeconomic Status, Social Network, and Competence on Subjective Well-Being in Later Life: A Meta-Analysis. *Psychology and Aging, 15*(2), 187-224.

Poortinga, Y.H. (1989) Equivalence of cross cultural data: an overview of basic issues. *International Journal of Psychology, 24, 737–756.*

Real Academia Española [Royal Spanish Academy]. (2015). *Diccionario de la lengua española [Dictionary of the Spanish Language]* (22.a ed.). Madrid, Spain: Autor.

Rendas-Baum, R., Yang, M., Varon, S. F., Bloudek, L. M., DeGryse, R. E., & Kosinski, M. (2014). Validation of the Headache Impact Test (HIT-6) in patients with chronic migraine. *Health and Quality of Life Outcomes, 12*(1), 1-10.

Schalock, R.L., & Verdugo, M.A. (2002). *The concept of quality of life in human services: A handbook for human service practitioners.* Washington, DC: American Association on Mental Retardation.

Schwartzmann, L. (2003). Calidad de vida relacionada con la salud: aspectos conceptuales [Quality of life related to health: Conceptual aspects]. *Ciencia y Enfermería, 9*(2), 9-21.

Scott, N.W., Fayers, P.M., Aaronson, N.K., Bottomley, A., de Graeff, A., Groenvold, M., Gundy, C., Koller, M., Petersen, M.A., & Sprangers, M.A.G. (2009a). Differential item functioning (DIF) in the EORTC QLQ-C30: A comparison of baseline, on-treatment and off-treatment data The EORTC Quality of Life Group and the Quality of Life Cross-Cultural Meta-Analysis Group. *Quality of Life Research, 18,* 381–388.

Scott, N.W., Fayers, P.M., Aaronson, N.K., Bottomley, A., de Graeff, A., Groenvold, M., Gundy, C., Koller, M., Petersen, M.A., & Sprangers, M.A.G. (2009b). The practical impact of differential item functioning analyses in a health-related quality of life instrument. *Quality of Life Research, 18,* 1125-1130.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: combining qualitative and quantitative approaches.* Thousand Oaks, CA: Sage.

Van de Vijver, F. J. R., & Matsumoto, D. (2011). Introduction to the methodological issues associated with Cross-Cultural Research. In D. Matsumoto & F. J. R van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 1-16). New York, NY: Cambridge University Press.

Veenhoven, R. (2000). The four qualities of life. *Journal of Happiness Studies, 1*(1), 1-39.

Wendel-Vos, G.C.W., Schuit, A.J., Tijhuis, M.A.R., & Kromhout, D. (2004). Leisure time physical activity and health-related quality of life: Cross-sectional and longitudinal associations. *Quality of Life Research, 13, 667-677.*

World Health Organization. (1994). *Quality of life Assessment- An Annotated bibliography.* Geneva: World Health Organization.

World Health Organization. (1998). *The world health report 1998: life in the 21st century A vision for all.* World Health Organization.

Yúdice, G. (2002). *El recurso de la cultura [The resource of culture].* Barcelona: Gedisa.

Zabriskie, R. B., & McCormick, B. P. (2003). Parent and child perspectives of family leisure involvement and satisfaction with family life. *Journal of Leisure Research, 35*(2), 163.

Papeles del Psicólogo, 2016. Vol. 37(1), pp. 74-79
http://www.papelesdelpsicologo.es

R AS THE ENVIRONMENT FOR DATA ANALYSIS IN PSYCHOLOGICAL EVALUATION

Ana María Ruiz-Ruano and Jorge L. Puga
UCAM Universidad Católica de Murcia

R es un entorno libre para el análisis estadístico de datos y para la generación de gráficos que está cobrando un papel clave en un amplio espectro de áreas de conocimiento. Pese a que el usuario ha de interaccionar con el programa, principalmente, por medio de una consola de comandos; se están invirtiendo esfuerzos en proporcionar entornos gráficos amigables que aproximen al usuario novel a esta especie de lengua franca del análisis estadístico contemporáneo. Además, la flexibilidad y versatilidad del programa, en conjunción con una comunidad de desarrolladores y usuarios global, facilitan que R sea la herramienta elegida en muchos contextos básicos o aplicados. En este trabajo se presenta la utilidad potencial que supone R para la evaluación psicológica en general a la vez que se acentúa el rol que asume la estadística propiamente dicha como instrumento de progreso de la psicología al amparo del método científico.
Palabras clave: *R, Evaluación psicológica, Análisis estadístico, Psicometría.*

R is a free computing environment for statistical data analysis and graph creation. It is becoming a key tool in a wide range of knowledge domains. The interaction with the software is mainly based on a command line interface but efforts are currently being made to develop friendlier graphical user interfaces that will help novice users to become familiar with this programming language. R is a flexible and powerful system thanks to the development community that is working together to improve its capabilities. As a result, it is the chosen statistical software in many applied and basic contexts. This paper highlights the potential usefulness of R for psychological assessment and related areas. Additionally, the relevance of statistical data analysis is emphasised as an instrument that will boost the progress of psychology under the umbrella of the scientific method.
Key words: *R, Psychological assessment, Statistical analysis, Psychometrics.*

The statistical analysis of data has become fundamental in contemporary culture. It is therefore not surprising that some degree of statistical literacy is required in society as a whole in order to aspire to new heights of human development (e.g., Bond, 2009; Seldmeier & Gigerenzer, 2001). For science, which seeks to understand nature, and for technology that yearns to control nature, statistics is fading as the key tool that guides and leads the progress of human communities. However, although today we have huge computational and statistical algorithms and we can implement them in computers powerful enough to handle a considerably large volume of data, it appears that we are not able to take advantage of this situation. It is worth remembering the slogan popularized by Pirelli (the Italian tyre company) in the mid-1990s, because it could clearly be applied to this situation: "power is nothing without control". Today we have very powerful computers, computer software for data analysis that is easy to use, and it is also relatively easy to collect data to carry out studies. However, in spite of all these facilities (power), it is crucial to have a good plan of what you want to do and a good research design (control), taking into account the data analysis to be performed *a posteriori*, to test the initial hypothesis (e.g., Bakan, 1966; Cohen, 1990; Dar, Serlin & Omer, 1994; Munafò et al, 2014; Rosnow & Rosenthal, 1989). In other words, even though we may have all the tools and all the data at our disposal, they will not do us any good if our study is lacking in planning and design. And, contrary to what one might think, being flexible in data analysis that is performed after an abnormal or non-existent planning phase does not help the investigation to progress satisfactorily, but rather leads to fruitless or contrived results with no scientific meaning (Ioannidis, 2005).

In recent years, there have been numerous reports of malpractice in research (e.g., Button et al, 2013; Cohen, 1990; Gigerenzer, 2004; Munafò et al, 2014; Trafimow & Marks, 2015; Vaux, 2012; Weissgerber, Milic, Winham & Garovic, 2015) which present a panorama that is scary, to say the least, with regards to the analysis of statistical data. There are even authors like Ioannidis (2005) who dare to affirm that many of the research papers presented are false and that, in fact, most of the discoveries that are made are merely evidence of these biases existing in the various sub-fields of research. As for malpractice, for example, Munafò et al. (2014) criticise the quality systems prevailing in science, the low reproducibility of the studies, the existence of a bias in scientific publications sponsored by the editorial policies of the scientific journals, the falsifying of the data, the low statistical power of the results presented, the development of illicit practices in research and the *p* value used widely as a critical element of statistical inference. Meanwhile, Button et al. (2013) focus on denouncing the low statistical power presented in the results of neuroscience research and the small sample size used in most studies. The importance of taking into consideration the ethical implications of animal sacrifice is also highlighted (see also e.g., Cressey, 2015, Nature Publishing, 2015) as well as the financial investment that is being made only to fail to obtain, in the majority of cases, useful results for society.

One of the issues that continues to receive most attention in terms of its misuse is the null hypothesis significance testing procedure or NHSTP (e.g., Bakan, 1966; Cohen, 1994; Gigerenzer, 1998; Haller & Krauss,

Correspondence: Jorge López Puga. *Facultad de Ciencias de la Salud. Departamento de Ciencias de la Salud. UCAM Universidad Católica de Murcia. Ctra. de los Jerónimos s/n. 30107 Guadalupe - Murcia. España. E-mail: jpuga@ucam.edu*

2002; Halsey, Curran-Everett, Vowler & Drummond, 2015; Nieuwenhuis, Forstmann & Wagenmakers, 2011; Rosnow & Rosenthal, 1989; Trafimow, 2014; Trafimow & Marks, 2015) and the abusive use of the *p* value as a single element that guides statistical inference in scientific studies. These malpractices were highlighted long ago by Bakan (1966), who referred to it as a flagrant attack against common logic and affirmed that "when we reach a point where statistical procedures are substitutes rather than aids to thinking, and we reach absurd conclusions, then we must return to the basics of common sense" (p.436). Similarly, Cohen (1994) and Gigerenzer (1998) warned of the importance of making good use of hypothesis testing and not turning it into a ritual that has been institutionalised in a scientific culture that has systematically repeated the same mistakes throughout history (Gigerenzer, 2004).

One of the most drastic and controversial steps that have been taken recently regarding the use of testing the null hypotheses was from the 2014 editorial of the journal *Basic and Applied Social Psychology* (Trafimow, 2014), in which it was proposed that authors should use a different statistical methodology than the null hypothesis procedure that used the *p* value as the sole value to guide the statistical inference, and which was explicitly prohibited in the 2015 editorial (Trafimow & Marks, 2015). The authors are asked to exploit the data at the descriptive level, and to use another type of data analysis that might lead to higher quality scientific results. The editors point out that, contrary to what many researchers may think, $p < .05$ is very easy to obtain and sometimes even serves as an excuse to publish poor quality research studies. There is evidence to prove the misuse of the *p* value in a statistical test. For example, the study carried out by Haller and Krauss (2002) who, after observing the misinterpretations made of the *p* value in psychology, concluded that both teachers and students had deficiencies when interpreting the significance of the research results. However, Leek and Peng (2015) suggest that the aforementioned ban will not guarantee an increase in the quality of the *science* that is produced, as well as the fact that the *p* value is "the tip of the iceberg" in the world of statistics and data analysis.

Over the years, psychology as a science has provided a number of measurement tools to carry out work in different areas of study that are closely linked to statistical data analysis. In neuropsychology, clinical, educational and organisational psychology, for example, it is very common to use tests. In fact, as noted by Hernández, Tomás, Ferreres & Lloret (2015), tests are considered basic instruments in psychological assessment to guide the psychologist's decision making. This is why, in order to be used correctly, tests must meet a series of psychometric guarantees, which should be available to the user applying the test. Hernández et al. (2015) note that some of the information that is not provided and that would be very useful is information related to the evidence of content validity, predictive validity, reliability as stability in the measurement, analysis of bias or differential item functioning. They also point out that it is the responsibility of the professional, even if their field of work may not be psychometrics, to train and learn about psychometric advances. However, away from the instruments used in psychological intervention or research, psychology as a science needs the scientific method in daily practice. Statistical data analysis plays an important role in all of these situations, which are not foreign to the professional psychologist. Therefore, like Dar et al. (1994), we could argue that the key to performing good data analysis, regardless of the area of psychological knowledge we are considering, lies in starting from a good base theory and designing impeccably the research you wish to carry out in order to prevent the weaknesses that are attributed to the procedure of testing null hypotheses.

The main objective of this paper is to present the R software as an environment for performing statistical analysis that may be very useful for psychology in general. As we will see below, this environment can enhance the learning of the statistical technique we are using. The next section is about the environment R, which is presented from a historical and philosophical perspective, as well as the perspective of the user interaction with R. Next, we briefly review some of the uses that R offers for psychological evaluation, presenting the different packages and their applicability to this discipline. Finally, we end with a reflection on possible solutions and improvements for attempting to implement useful and quality research for society in general with the support of statistical analysis, in which R plays a crucial role today.

R ENVIRONMENT

R is a working environment for carrying out statistical analysis and creating graphs (R Development Core Team, 2011). The graphical interface of the program is a command console, which means that in order to interact with it you have to write lines of code and run them. We can trace the origin of R back to the 1970s, when the term S was coined to refer to a high-level programming language designed to perform statistical computations. The shift to R occurred in the 1990s, coinciding with the massive expansion of S-Plus, the commercial version of S. Ross Ihaka and Robert Gentleman (University of Auckland, New Zealand) wrote an educational version of the first distribution of R, which was released to the general scientific community in 1996 (Ihaka & Gentleman, 1996). R was structured based on the programming language S and Scheme, a dialect of Lisp. From the beginning, the aim was for it to be a multiplatform language that could be used under different operating systems. In 1997 the R Core Team was created. This was the body responsible for developing the software base and providing continuous support for the development and distribution of R (Fox, 2009). Since then, while continuing to grow and develop, R has established itself as the statistical reference tool par excellence. In fact, Tippmann (2015) notes that R is the data analysis software most used in 2015 on a par with, or even surpassing, other commercial software programs.

The fact that it is considered a working environment means that R is a programming language that comes equipped with a set of tools for calculating and generating statistical graphics (Ihaka & Gentleman, 1996). This working environment was made to be –and still is– free, in the sense that it is not necessary to pay to use, copy or distribute the program (Carleos & Corral, 2013). Furthermore, it is protected against possible privatisation / commercialisation under a GNU license. The GNU GPL (General Public License) philosophy is the basis of freedom and openness of code, and its purpose is to protect the rights and freedoms of end users (http://www.fsf.org). The basic pillars on which R rests are allowing to use, share, study and modify the software code that is at the base of the program. This means that any user of R can develop a new application of it, modifying what already exists, sharing it and using it freely.

R consists of a *base* structure that contains a number of basic functions for making statistical calculations and graphs. Additionally, a range of

accessories called packages have been created by the scientific community to perform statistical analyses of more specific data. This software development activity may be regarded as having been done in an altruistic way and caters more to the intrinsic factors of the person –such as the satisfaction produced by this collaboration, the intrinsic rewards of the job or the fact that they are contributing to a public asset– than to external factors such as scientific recognition or money (Fox, 2009). An R package is simply "a set of functions that maintain some kind of relationship between them" (Elosua, 2011, pp.24-25). For example, there are packages for the statistical analysis of issues as varied as the content of electronic texts, the spatial distribution of urban areas or the neural activation recorded by a neuroimaging technique. All of these are freely available on the Internet at no cost. The CRAN (Comprehensive R Archive Network, http://www.r-project.org) is the network location where all of the downloads are managed. For greater detail on downloading and installing R and its packages see, for example, Elosua (2009) or López (2012, 2013).

The main advantage of R compared with other statistical analysis software is the freedom and the fact that it is free to use. Other advantages of using this working environment include the fact that it promotes the learning of basic statistics (López, 2012; Tippmann, 2015) and that it can be run on different operating systems (Fox, 2009), such as Windows, Mac OS or Linux. Furthermore, Huber et al. (2015) note that the use of R as a statistical program facilitates research and innovation, because it allows the development of rapid prototypes; it is flexible and functional; it allows reproducibility; changes or modifications are made quickly; it provides graphic facilities and it permits interaction with other programming languages such as C and C ++ or JavaScript for web applications. With regard to the interpretation of the outputs generated after performing statistical analysis, certain advantages can also be observed. Since the user of the software has to become more involved in the statistical tool he is using from a technical point of view, he ends up being more adept at interpreting the results generated by the data analysis he applies. One drawback we can point out is the difficulty working with a command environment compared to other programs or interfaces where all you have to do is click (e.g., López, 2013; Tippmann, 2015). However, as discussed below, the graphical user interfaces that have been developed are increasing in number and in quality in order to make the interaction with R more user friendly.

The graphic interface R, as we mentioned above, is based on command lines. However, to facilitate interaction with this program, graphical interfaces have been developed that are more intuitive for the users. These interfaces can be classified into two types, the first are Windows type menus (e.g., R Commander or RKWard) and the second are code editors (e.g., Tinn-R, Emacs or RStudio). Interfaces with Windows-type menus are preferred by users that are familiar with commercial software (e.g., Elosua, 2009; Fox, 2005; López, 2013). The R Commander package developed by Fox (2005) was presented as the most natural transition between commercial statistical analysis software and R (Elosua, 2009). It was developed with the objective of providing a user-friendly platform for basic courses in statistics, aiming to reduce the possibility of committing careless mistakes (e.g., calculating the mean of a nominal variable) and, finally, to make visible the relationship established between what is selected in the menus and the R code that would have had to be used in the command console. Fox (2005) points out that this type of graphical interface has positive and negative aspects for the users. One positive point noted is that it is not necessary to remember commands or arguments of the functions, which reduces the likelihood of errors in the script syntax, and also it is a comfortable environment for the work of novice, infrequent or causal users. The negative aspects highlighted were the fact that it is tedious and difficult to reproduce certain statistical analyses, because it involves navigating through multiple dialogue boxes, and also the fact of having to incorporate numerous analyses on the interface, which can end up becoming a labyrinth that is difficult to use (Fox, 2005). One of the most attractive proposals that have been made in recent times to facilitate the user interaction with R base code can be found in the program JASP (Love et al., 2015). JASP (https://jasp-stats.org) aims to enrich the user experience with a quick, agile and easy to use graphical interface. Meanwhile, interfaces such as RStudio (www.rstudio.com), which is an integrated or interactive development environment, –meaning it is an application that provides facilities and tools to users that work in computer programming– are preferred by users with more advanced programming skills in R or those who wish to try to be productive with code (López, 2013).

Finally, we would like to note that turning to the sources of support for working with R is almost inevitable regardless of the level of skill we have in using it. R comes with a series of PDF manuals that are available to the user from the console itself. It also provides various other sources of support (in html, FAQs, etc.). Another way of obtaining help, in this case regarding specific packages, is by going directly to the directory where the package is hosted on the web or to the local installation directory. Using traditional web browsers can be useful, the Quick-R (www.statmethods.net) or the CRAN webpage (López, 2013). There are manuals in both English (e.g., Field, Miles & Field, 2012) and Spanish (e.g., Arriaza et al, 2008; Elosua, 2011; Elosua & Etxeberria, 2012; López, 2012, 2013) to facilitate the interaction with this software. Therefore, in addition to being supported by a community of developers and users of the software, the researchers and professionals who prefer to use R as a tool for the statistical analysis of data will have a large pool of resources to make their interaction with the software more satisfactory. Table 1 presents a short list of resources, freely available on the web, which can be useful to make the interaction with R more user-friendly.

USES IN PSYCHOLOGICAL ASSESSMENT

In this section we present very briefly some of the R packages that take centre stage in the field of psychological assessment and particularly we discuss some of those most related to psychometric technology. It would be impractical to try to cover them all and that is not the intention of this article. At the moment of writing these words, there are 6,695 packages available for download and it is likely that by the time the article is being read this amount will have increased by a substantial number given the quasi-exponential growth experienced in recent years (Elosua & Etxeberria, 2012). Although only some of the packages are directly related to psychology, it is also true that there are many packages that may be used by our discipline occasionally, even though they have been developed in areas as diverse as topographic pattern recognition. In any case, CRAN contains a listing of the available packages together with documentation that specifies how they can be used efficiently.

In 2007 Patrick Mair and Reinhold Hatzinger inaugurated the *Psychometric Task View* with the aim of bringing together all the initiatives and progress that had been made so far in the field of psychometrics (Mair & Hatzinger, 2007a). The same year, a special issue appeared in the *Journal of Statistical Software* which formally presented some of the most relevant packages for psychometric data analysis with R (de Leeuw & Mair, 2007). The packages (127 at present) were grouped into five major groups (plus one miscellaneous) oriented to psychometric modelling under item response theory, correspondence analysis, developing structural equation models, multidimensional scaling, and classical test theory.

Most of the packages compiled by the *Task View* are concentrated in the group dedicated to statistical modelling under item response theory (IRT). Thus it could be deemed to fit in with the attempt being made in recent years to advance the advantages offered by this theory of psychological measurement compared to Classical Test Theory (Muñiz, 1997; 2010). If we had to highlight any of the packages included in this group, it would be **ltm** (Rizopoulos, 2006) since as well as including the functionality to develop dichotomous and polytomous IRT models (one, two or three parameters), it also contains functions that have been highlighted in different application contexts other than IRT (Falissard, 2012). We should also highlight the packages **eRm** (Mair & Hatzinger, 2007b) and **plRasch** (Anderson, Li, & Vermunt, 2007).

Additionally, we should emphasise four essential packages for working from the perspective of Classical Test Theory (Mair & Hatzinger, 2007a): **psychometric**, **psy**, **psych** and **MiscPsycho**. These four packages bring together basic and advanced functions dedicated to item analysis, and the study of the validity and reliability of tests, as well as useful functions for the development of scales under classical philosophy.

Another large group of packages is the one intended for factor analysis, principal component analysis and the development of structural equation modelling. Factor analysis can be considered a statistical analysis technique that, despite having emerged within psychology, shows its usefulness in knowledge areas as diverse as biology or economics (Ferrando & Anguiano-Carrasco, 2010). Classical factor analysis and principal component analysis are implemented in the **stats** package that is contained in the base distribution of R. Although the versatility of the factanal() and princomp() functions of the **stats** package enables you to run factor analysis and basic principal component analysis, other packages have been developed to complement and optimise the user experience of R in this regard. Structural Equation Models are emerging today as a kind of evolution of factor analysis and associated techniques although they involve even greater complexity from the point of view of statistical estimation (Ruiz, Pardo & San Martin, 2010). In R there are some packages that have been aimed at developing this type of model with similar functionality to the commercial packages available to date. For example, the **sem** package (Fox, 2006) and the **lavaan** package (Rosseel, 2012) are two tools that are hugely popular for the development of structural equation models that allow the estimation of factor models of various kinds while also facilitating their graphical representation. Packages such as **polycor**, among others, for estimating models involving polychoric correlations, **systemfit**, for nonlinear structural alternatives, and **pls**, for estimating the minimum partial quadratic allow great flexibility and versatility in the range of statistical techniques applicable to data generated in the field of study of psychology that can be treated under the structural perspective.

Finally, we should note that different procedures aimed at estimating and testing models of multidimensional scaling can be found in packages such as **MASS**, **MLDS**, **vegan**, **labdsv**, **ecodist** and **ade4**.

CONCLUSIONS

As we have tried to demonstrate, R is a considerably versatile and flexible tool that enables us to carry out statistical data analysis that is difficult to achieve using any other software available today. In addition, the large number of packages that are associated with R makes it easy for it to be used as the only software, which avoids the need to switch from one program to another when specific computing needs arise. Thus, psychology could benefit from this potential in all of its knowledge areas. However, it is also true that interacting efficiently with the program requires a certain amount of skill, which cannot be acquired quickly. In any case, more and more attempts are being made to develop graphical user interfaces that are easier for general users, so that they do not panic when faced with the command console. Let us not forget, as we have pointed out from the beginning, that *power without control is useless* and that statistical data analysis only provides a spectrum of techniques to be applied which have to be backed up by solid scientific theories.

Cohen, (1990) noted that statistical inference, in conjunction with the informed judgment of scientists is a very useful tool. However, he also

TABLE 1
SOME FREE RESOURCES FROM THE WEB THAT CAN BE USEFUL TO MAKE THE INTERACTION WITH R MORE USER-FRIENDLY

Resource Description	Link
Brief manual covering basic and common statistical analysis in Psychology with R Commander (Spanish)	http://hdl.handle.net/10835/1658
One of the best compilations of tutorials on R, which addresses both the basic and more advanced aspects of statistical analysis for Psychology (English)	http://ww2.coastal.edu/kingw/statistics/R-tutorials/
This page contains a set of links to introductory video tutorials on R (Spanish)	http://www.jpuga.es/Docencia/mooc-r.html
Interactive course on operating the R console (English)	http://tryr.codeschool.com/
Basic manual on psychometrics with R Commander (Spanish)	https://web-argitalpena.adm.ehu.es/pasa_pdf.asp?File=UWLGPS5661

believed that statistics are not the most important tool of science but rather the result of what comes before (theory, planning, scenario, etc.). This is why we need to re-emphasise the importance of research planning as an essential part of the scientific method, as many other authors also affirm (e.g., Cohen, 1990; Munafò et al, 2014), so we can obtain results that really are useful for both the scientific community and society in general.

In this sense, a cultural change in the scientific community is necessary (e.g., Morrison, 2012; Munafò et al, 2014; Weissgerber et al, 2015) to allow us to obtain quality work, faithful to professional ethics adapted to the current times, along with a change in the editorial policies of the scientific journals (Weissgerber et al., 2015). This cultural change will improve in part if researchers receive training (Weissgerber et al., 2015), as Leek and Peng (2015) point out when they say that "education is the beginning." As for the changes in editorial policies, at least in psychology, we have the somewhat extreme example of the *Basic and Applied Social Psychology* journal which prohibited authors from sending articles that use the *p* value (Trafimow, 2014; Trafimow & Marks, 2015).

Similarly, and to provide greater quality to the scientific results that are presented, some authors recommend that the exploratory data should be emphasised, i.e., that greater use should be made of descriptive statistics (e.g., Cohen, 1990, 1994; Leek & Peng, 2015; Trafimow & Marks, 2015; Weissgerber et al, 2015). It would also be interesting to provide more graphical results (Cohen, 1994), information on the effect size (e.g., Cohen, 1994), confidence intervals or using Bayesian statistics (e.g., Bakan, 1966; Cohen, 1992, 1994; Haller & Krauss, 2002; Puga, Krzywinski & Altman, 2015a, 2015b; Trafimow & Marks, 2015). Another important point suggested by some authors is to carry out further replications of studies that find results that may be considered weak from a methodological point of view (e.g., Halsey et al. 2015; Huber et al, 2015; Munafò et al, 2014; Rosnow & Rosenthal, 1989). In addition, Ioannidis (2005) suggests that obtaining meaningful results should not be pursued and the previous information regarding the issue being researched should be taken into account. The latter underscores the interest in using Bayesian statistics, the foundations of which are based on pre-existing information.

The credibility of the science of psychology is in our hands and the statistical analysis of data is an essential ally in which to trust in order to optimise the evolution of scientific progress. R is presented today as the lingua franca of statistical analysis and graph generation in a wide range of subject areas. It is about time it was determined whether R will ultimately become something that could be considered as a world heritage (López, 2012) or whether its peak will simply be the result of a passing trend.

ACKNOWLEDGEMENTS

This work was carried out under the Ministry of Economy and Competitiveness (project PSI2014-53427-P).

REFERENCES

Arriaza, A. J., Fernández, F., López, M. A., Muñoz, M., Pérez, S., & Sánchez, A. (2008). *Estadística básica con R y R-Commander.* [*Basic statistics with R and R-Commander.*] Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.

Anderson, C. J., Li, Z., & Vermunt, J. K. (2007). Estimation of models in a Rasch family for polytomous items and multiple latent variables. *Journal of Statistical Software, 20*, 1-36.

Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin, 66*, 423-437. doi: 10.1037/h0020412

Bond, M. (2009, October). Decision-making: risk school. *Nature, 461*, 1189–1192. doi: 10.1038/4611189a

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Neuroscience, 14*, 365-176. doi: 10.1038/nrn3475

Carleos, C. & Corral, N. (2013). Paquetes estadísticos con licencia libre (I). [Statistical packages with free licences (I)] *Revista Electrónica de Metodología Aplicada, 18*, 12-37.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45*, 1304-1312. doi: 10.1037/0003-066X.45.12.1304

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101. doi: 10.1111/1467-8721.ep10768783

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49*, 997-1003. doi: 10.1037/0003-066X.50.12.1103

Cressey, D. (2015, Abril). UK funders demand strong statistics for animal studies. *Nature, 520*, 271-272. doi:10.1038/520271a

Dar, R., Serlin, R. C., & Omer, H. (1994). Misuse of statistical tests in three decades of psychotherapy research. *Journal of Consulting and Clinical Psychology, 62*, 75-82. doi: 10.1037/0022-006X.62.1.75

de Leeuw, J. & Mair, P. (2007). An introduction to the special volume on "psychometrics in R". *Journal of Statistical Software, 20*, 1-5.

Nature Editorial (2015, April). Numbers matter. *Nature, 520*, 263-264. doi:10.1038/520263b

Elosua, P. (2009). ¿Existe vida más allá de SPSS? Descubre R. [Is there life after SPSS? Introducing R] *Psicothema, 21*, 652-655.

Elosua, P. (2011). *Introducción al entorno R.* [*An introduction to the R environment*] Bilbao: Euskal Herriko Univertsitateko Argitalpen Zerbitzua / Editorial Service of the University of the Basque Country.

Elosua, P. & Etxeberria, J. (2012). *R Commander. Gestión y análisis de datos [R Commander. Data management and analysis].* Madrid: La Muralla.

Falissard, B. (2012). *Analysis of questionnaire data with R.* Boca Raton, FL: Taylor & Francis Group.

Ferrando, P. J. & Anguiano-Carrasco, C. (2010). Factor analysis as a research technique in psychology. *Papeles del Psicólogo, 31*, 18-33.

Field, A., Miles, J., *R Commander. Gestión y análisis de datos* Field, Z. (2012). *Discovering statistics using R.* London: SAGE.

Fox, J. (2005). The R commander. A basic-statistics graphical user interface to R. *Journal of Statistical Software, 14*, 1-42.

Fox, J. (2006). Structural equation modeling with the **sem** package in R. *Structural Equation Modeling, 13*, 465-486.

Fox, J. (2009). Aspects of the social organization and trajectory of the R project. *The R Journal, 1/2*, 5-13.

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics, 33*, 587-606. doi: 10.1016/j.socec.2004.09.033

Gigerenzer, G. (1998). We need statistical thinking, not statistical rituals. *Behavioral and Brain Sciences, 21*, 199-200. doi: 10.1017/S0140525X98281167

Haller, H. & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research Online, 7*, 1-20.

Hasley, L. G., Curran-Everett, D., Vowler, S. I., & Drummond, G. B. (2015). The fickle *P* value generates irreproducible results. *Nature Methods*, *12*, 179-185. doi: 10.1038/nmeth.3288

Hernández, A., Tomás, I., Ferreres, A., & Lloret, S. (2015). Tercera evaluación de tests editados en España [Third evaluation of tests published in Spain]. *Papeles del Psicólogo*, *36*, 1-8.

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., … & Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, *12*, 115-121. doi: 10.1038/nmeth.3252

Ihaka, R. & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, *5*, 299-314. doi: 10.1080/10618600.1996.10474713

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine*, *2*, 696-701. doi: 10.1371/journal.pmed.0020124

Leek, J. T. & Peng, R. D. (2015, Abril). P values are just the tip of the iceberg, *Nature*, *520*, 612. doi: 10.1038/520612a

López, J. (2012). *Introducción al análisis de datos con R y R Commander en psicología y educación [Introduction to data analysis with R and R Commander in psychology and education.]*. Almería: Editorial of the University of Almería. Available from http://hdl.handle.net/10835/1658.

López, J. (2013). *Primeros pasos analizando datos estadísticos con R. Un texto para acompañarte [First steps to analysing statistical data with R. A text to accompany you.]*. Murcia: Universidad Católica San Antonio.

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., Ly, A., Gronau, Q. F., Smira, M., Epskamp, S., Matzke, D., Wild, A., Rouder, J. N., Morey, R. D. & Wagenmakers, E.-J. (2015). JASP (Version 0.7)[IT program].

Mair, P. & Hatzinger, R. (2007a). Psychometrics task view. *R News*, *7*, 38-40.

Mair, P. & Hatzinger, R. (2007b). Extended rasch modeling: The eRm package for the application of IRT models in R. *Journal of Statistical Software*, *20*, 1-20.

Morrison, R. P. (2012). Reforming science: methodological and cultural reforms. *Infection and Immunity*, *80*, 891-896. doi: 10.1128/IAI.06183-11

Munafò, M., Noble, S., Browne, W. J., Brunner, D., Button, K., Ferreira, J., … & Blumenstein, R. (2014). Scientific rigor and the art of motorcycle maintenance. *Nature Biotechnology*, *32*, 871-873. doi:10.1038/nbt.3004

Muñiz, J. (1997). *Introducción a la Teoría de Respuesta a los Ítems [Introduction to Item Response Theory]*. Madrid: Pirámide.

Muñiz, J. (2010). La teoría de los tests: Teoría Clásica y Teoría de Respuesta a los Ítems. [Test Theory: Classical Theory and Item Response Theory] *Papeles del Psicólogo*, *31*, 57-66.

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nature Neuroscience*, *14*, 1105-1107. doi: 10.1038/nn.2886

Puga, J. L., Krzywinski, M., & Altman, N. (2015a). Points of Significance: Bayes' theorem. *Nature Methods*, *12*, 277–278. doi:10.1038/nmeth.3335

Puga, J. L., Krzywinski, M., & Altman, N. (2015b). Points of Significance: Bayesian statistics. *Nature Methods*, *12*, 377–378. doi: 10.1038/nmeth.3368

R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, *17*, 1-25.

Rosnow, R. L. & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, *44*, 1276-1284. doi: 10.1037/0003-066X.44.10.1276

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1-36.

Ruiz, M. A., Pardo, A., & San Martín, R. (2010). Structural equation models. *Papeles del Psicólogo*, *31*, 34-45.

Seldmeier, P. & Gigerenzer, G. (2001). Teaching bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General, 130*, 380–400. doi: 10.1037/0096-3445.130.3.380

Tippmann, S. (2015, January). Programming tools: Adventures with R. *Nature*, 109-110. doi:10.1038/517109a

Trafimow, D. (2014). Editorial. *Basic and Applied Social Psychology*, *36*, 1-2. doi: 10.1080/01973533.2014.865505

Trafimow, D. & Marks, M. (2015). Editorial. *Basic and Applied Social Psychology*, *37*, 1-2. doi: 10.1080/01973533.2015.1012991

Vaux, D. (2012, December). Know when your numbers are significant. *Nature*, *492*, 180-181. doi:10.1038/492180a

Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond bar and line graphs: Time for a new data presentation paradigm. *PLOS Biology*, *13*, 1-10. doi: 10.1371/journal.pbio.100212