Assessment instruments on children's internalizing behaviour: A review of their psychometric properties, functioning and applications

Maria A. Efstratopoulou
School of Social Sciences,
Department of Special Education and Inclusion,
Bishop Grosseteste University, U.K.
maria.efstratopoulou@bishopg.ac.uk

Abstract

The study aims to summarize assessment instruments for teachers and practitioners to record externalized and internalized problems in children. The author sampled articles on children's psychopathology over the last twenty-five years, selected checklists and rating scales with multiple citations, and reviewed their properties. Most instruments use teachers and parent-report forms. Only a few instruments assess a number of different disorders. The psychometric properties of these instruments are presented in tables. Discussion and recommendation for teachers and researchers on which assessment instrument is the best to use when rating children's behaviour, indicated that the rater must define in advance the goals of assessment, taking into consideration the factors that the instrument measures and its psychometric properties. The use of more than one instrument is recommended.
Keywords: rating scales; checklists; disorders; validity; reliability.

1 Introduction

Unskilled, aggressive hyperkinetic and impulsive children are quickly rejected and ostracized from peer groups, and become frequent targets of bulling in school settings by their peers (Snyder, 2004). Research on children's psychopathology indicates that a large number of school-aged children with symptoms of depression, anxiety and mental health problems and are often left without necessary early diagnosis, psychiatric assessment and help (Efstratopoulou, Janssen, & Simons, 2012). In addition, many children facing attentional, emotional, and/or behavioral problems are placed in public typical elementary schools without a first screening. These children are 'at risk' for school failure, emotional difficulties and significant negative adult outcomes compared to their typical developing peers (Eisenberg, Fabes, Guthrie, & Reiser, 2000).

Detection efforts are particularly critical during the early years. Students with symptoms of emotional and/or behavioral disorders (EBD) and children with social in interaction with peers and teachers, experience negative outcomes within and beyond the school setting (Landrum, Tankersley, & Kauffman, 2003; Volkmar, Lord, Bailey, Schultz, & Klin, 2004). Students with EBD who do not receive necessary support often experience a host of negative outcomes, including peer and teacher rejection, academic underachievement, school dropout, substance abuse, depression, unemployment, and involvement in the juvenile justice system (Bullis & Yovanoff, 2006; Morris, Shah, & Morris, 2002; Wagner & Davis, 2006; Wagner, Newman, Cameto, Levine, & Garza, 2006; Zigmond, 2006). Given the costs associated with emotional and behavioural disorders and mental health problems, to students themselves, their families, and society as a whole, it is not surprising that a great body of research focusing on reducing the incidence of these

disorders through systematic screening and intervention efforts (Kauffman & Landrum, 2009; Lane, 2007).

Despite the usefulness of rating instruments for describing children's deviant behaviours, the relatively modest agreements among rating sources raise questions about the validity of information and the importance of setting on children's behavior. In general, concordance has been found to be higher when informants have similar relationships (for example between class teachers and physical educators) than between teachers and parents (Efstratopoulou, Janssen & Simons, 2012a).

Parents can observe their child in a wide range of situations; nonetheless, information from the parents is not always reliable and tends to follow a pattern of idealized expectations and cultural stereotypes that can affect the reliability of their reports (Mash & Johnston, 1983). For example, some parents when rate their child's behaviour are very sensitive or have a low threshold for certain behaviours, whereas other parents may underreport deviant or troublesome children's behaviours. The accuracy of parents as raters may vary greatly depending on factors as parent's education, the amount of stress associated with the child's behaviours, and hidden agenda's that parents may have when rating a child. Apart from parents, teachers and especially teachers in primary education interact with children during many hours a day. Hence, several assessment instruments have been developed to gather information about children's wellbeing using teacher's ratings (Achenbach, 1991).

Children experiencing anxiety problems (APA; American Psychological Association, 2013) have a consistent pattern, of uncontrollable and excessive anxiety or worry, with the concerns covering a broad range of events or activities, irritability, restlessness, fatigue, difficulty in concentrating, muscle tension, sleep disturbances and the disorder commonly begins at around age ten, is persistent, and frequently co-occurs with depression (Beidel, 1989). Depression is another relatively common disorder that often first appears in childhood or adolescence. Children with major depression experience depressed mood (or irritability) loss of interest in their usual activities plus other symptoms such as sleep or appetite disturbance, loss of energy, or trouble concentrating (APA, 2000). These symptoms must be present nearly every day for two weeks or more. Standardized questionnaires are also used to measure depression and determine whether a child's level of symptoms are in the non-depressed range or indicate mild, moderate, or severe levels of depression.

The purpose of this study was to review the most popular assessment instruments in children's psychopathology, and examine their properties. Based on specific selection criteria a list of instruments that may be suitable for assessing children's deviant behavior within school context, will be considered with the purpose of making recommendations for researchers and educators on how to select the most appropriate to use for educational or research purposes.

2   Method

2.1   Search Strategy

Literature searches, using tailored search terms appropriate to each database were conducted using Web of Science, Medline, PubMed, PsychoINFO and SportDISCUS. Typical search term included: emotional and behavioral problems, measurement, checklists, rating scales, motor behavior, education, children and physical activity. Rating scales and checklists cited in child and adolescent psychiatric literature over the last 25 years were selected and reviewed. Obviously, not all available instruments could be included in this study. Inclusion criteria for assessment instruments to be considered for further evaluation required that: a) that have long track records in educational research,

b) assess elementary school-aged children's behavior, c) have specific educators' and parents' versions, d) have established psychometric properties and ongoing citations in the literature. Instruments assessing only physical fitness and motor functioning abilities or provide only self-report versions were excluded. The search strategy resulted in 47 instruments measuring externalizing and/or internalizing disorders in children which were initially screened against the inclusion/exclusion criteria for potential relevance. After this initial screening, 11 measures were assessed as potentially relevant and presented in details in Tables and text. Table 1 includes the description and psychometric properties of eleven assessment instruments for children internalizing behavior (e.g., Depression, Anxiety) and developmental disorders (e.g., Autism). The information provided focusing on instruments functioning and implementation, subscales, psychometric properties, number of items, and response format.

Table 1. Characteristics and psychometric properties of Anxiety, Depression and Developmental Disorders instruments for children

Scale	Factors	Reliability	Validity	Items, Scoring, Versions
Multidimensional Anxiety Scale for Children (MASC; March, 1997)	1. Physical symptoms 2. Social anxiety 3. Harm avoidance 4. Separation anxiety	IC: .60-.90 (clin) TR: .65-.93 (clin)	CONV:.63 (clin) DISC: .74 (clin)	39 items 4 point scales Teacher/Parent version Self-report version
Screen for Child Anxiety Related Emotional Disorders (SCARED; Birmaher, 1999)	1. Somatic/ panic 2. Generalized Anxiety 3. Separation Anxiety 4. Social phobia 5. School phobia	IC: .78-.90 (clin) TR:.70-.90 (clin)	CONV: good (clin) DISC: good (clin)	41 items 3 point scales Teacher/Parent version Brief (5-item version) Self-report: SCARED-R (66 items)
Anxiety Scale for Children Re**vised (SASC-R;** La Greca, 1999)	1. Fear of Negative Evaluation 2. Social Avoidance 3. Generalized Social Avoidance	IC: .69-.86 (com) TR: .70 (clin; com)	CONV: adequate (com) DISC: good (clin)	18 items SAS-A: Self report (22 items)
Fear Survey Schedule for Children- Revised (FSSC-R; Ollendick, 2002)	1. Fear of Failure and Criticism 2. Fear of the Unknown 3. Fear of Injury & Small Animals	IC:.90 (com) TR:.80 (com)	CONV: good (com) DISC: good (com)	80 items 5 point scale Teacher/ Parent version
Children's Depression Rating Scale-Revised (CDRS-R; Poznanski & Mokros, 1999)	1. Mood 2. Suicidality 3. Sleep problems 4. Social withdrawal 5. Somatization	IC: .80- .96 (clin) TR: .81 (com, clin)	CONV: .75- .92 DISC: good (clin)	17 items 5-7 points scale Teacher/ Parent version Self- report version Brief 5-items version
Child Behaviour Checklist (**C**BCL; Achenbach,1991) **Teacher Report Form** (TRF; Achenbach, 1991)	1. Withdrawal 2. Somatic Problems 3. Anxiety/Depression 4. Social Problems 5. Thought Problems 6. Attention Problems 7. Delinquent Behaviour 8. Aggressive Behaviour	IC: .80-.85 (CBCL) IC: .88-.90 (TRF) TR:.68-.90 (CBCL) TR: .73-.96 (TRF)	CONV: good (com; school; clin) DISC: high (school, clin)	120 items 3-point scale Parent Version (CBCL) Teacher version (TRF) Self-report version (YRF)
Motor Behaviour Checklist (MBC; Efstratopoulou, Jansen, & Simons, 2012)	1. Rules breaking 2. Low energy 3. Stereotyped behaviours 4. Hyperactivity /Impulsivity 5. Lack of Attention 6. Lack of Social interaction (10) 7. Lack of Self- regulation (12)	IC: .82-95 TR: .81-87	CONV: good (clin, school	59 items 7-point scale PE /teacher version

					62 items
Baby and Infant Screen for Children with Autism Traits-Part 1 (Matson et al., 2010)	1. Aggressive/disruptive Behavior 2. Stereotypic behaviour 3. Self-injurious behaviour	IC: .60- .90 (clin) TR: .65-.93 (clin)	CONV: .63 (clin) DISC: .74 (clin)		3-point scale Clinician/Parent versions BISCUIT-Part 3 (17-items)
Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, & Barton, 1999)	1. Stereotyped movements 2. Social reciprocity 3. Imaginative play	IC: .94-. 76 (clin) TR: .70- .90 (clin)	CONV: good (clin) DISC: good (clin)		23 items Yes-No Critical items (score) Clinician/Parent versions
Childhood Autism Rating Scale (CARS; Schopler, Reichler, Renner, 1988)	1. Relating to others 2. Body use 3. Adaptiveness to change 4. Listening response 5. Verbal communication	IC: .88- .96 (com, clin) TR: .81- .88 (com)	CONV: good (clin) DISC: moderate (clin)		15 items 4 -point scale Clinician/Parent versions

*Note: IC=Internal Consistency; TR=Test-Retest; CONV=Convergent; DISC: Discriminant; clin= Clinical sample; com= Community sample; school=School sample

3   Anxiety and Depression Scales

3.1   Multidimensional Anxiety Scale for Children (MASC)

The Multidimensional Anxiety Scale for Children (MASC) (March, 1997; March, Parker, Sullivan, Stallings, & Conners, 1997) is an empirically derived scale that assesses a spectrum of anxiety symptoms rather than a single anxiety construct. Its four major factors are invariant across age and gender, and three factors can be subdivided: physical symptoms (tense/restless and somatic/autonomic), social anxiety (humiliation/rejection and public performance fears), harm avoidance (perfectionism and anxious coping), and separation anxiety. Two of these major factors match the DSM-IV diagnoses of social phobia and separation anxiety disorder (SAD), while the total score matches generalized anxiety disorder (GAD). Both internal and test–retest reliability of the MASC range from moderate to excellent across the factors and is unaffected by age or gender but is lower for African-American than for white youths (March & Sullivan, 1999). Inter-informant agreement varies with the informant dyads. Parent–child concordance demonstrated the expected low agreement, being better for easily observable symptoms, and better for mother–child dyads than for father–child dyads. Validity appears moderate to good. A nice feature is that the subscales and the Anxiety Disorders Index discriminated youths according to the presence or absence of anxiety disorders with 88% accuracy. Youths with obsessive compulsive disorder (OCD) demonstrated significant decreases in their compulsive behaviours but not in their MASC-anxiety (Geller et al., 2001), further supporting discriminant validity compared with other anxiety scales.

3.2   Screen for Child Anxiety Related Emotional Disorders (SCARED)

The Screen for Child Anxiety Related Emotional Disorders (SCARED) (Birmaher, 1999; Birmaher, Brent, Chiappetta, Bridge, Monga, & Baugher, 1999) was developed with heterogeneous youths presenting to a mood and anxiety disorders clinic. Its five factors conform with DSM-IV disorders: somatic/panic, GAD, SAD, social phobia, and school phobia. Internal and test–retest reliabilities vary from moderate to excellent for the total scale and subscales. The overall low inter informant agreement, or concordance, for parent and child was better with adolescents than with children. Validity appears promising. Girls score higher than boys on all scales, and the SCARED correlates well with other internalizing scales, including with the similarly MASC (Muris, Bjorg, Moulaert, & Merckelbach, 1998a), and the older RCMAS (Muris, Merckelbach, Brakel, Mayer, & Dongen, 1999; Muris, Merckelbach, Brakel, Mayer, & Dongen, 1998b). All scales discriminate anxious from other youths, especially those with disruptive disorders and to

Journal of Physical Activity, Nutrition and Rehabilitation, 2016
www.panr.com.cy
info.panr@unic.ac.cy

some degree from depressed youths. The social phobia and school phobia perform less well than the other scales. The construct of anxiety is clear because it is based on DSM-IV, and its discriminative abilities are impressive. If its suggested ability to discriminate anxiety from depression is replicated, it would be especially helpful in examining internalizing disorders. The SCARED has functioned as well in Dutch community studies as in the original clinical sample (Muris et al., 1998a; Muris, 1999). The 3-point scoring may not be optimally sensitive to treatment effects, although initial treatment studies have been variable (Research Unit on Pediatric Psychopharmacology Anxiety Study Group, 2001).

3.3 Social Anxiety Scale for Children-Revised (SASC-R)

The Social Anxiety Scale for Children-Revised (SASC-R) (La Greca, 1999; La Greca & Stone, 1993) is based on the hypothesis that social anxiety fosters the development of maladaptive social behaviours that lead to anxiety disorders. The SASC-R examines social anxiety in relation to peer functioning, a major source of affective experience. Its three subscales are moderately inter correlated: Fear of Negative Evaluation, Social Avoidance and Distress - Specific to New Peers or Situations (SAD-New), and Generalized Social Avoidance and Distress (SAD-G). Internal consistency for all scales and subscales is moderate to very good and test-retest reliability is good (Ginsberg, La Greca, & Silverman, 1998; La Greca et al., 1993). Subscale scores are preferred because they assess distinct aspects of functioning. In school and clinical samples, higher fear of negative evaluation, SAD-New, and SAD-G scores correlate with poorer social acceptance, self-worth, and conduct (Ginsburg et al., 1998). Because social anxiety correlated more strongly with social acceptance than with conduct, the authors claim support for divergent validity. Support for discriminate validity was offered by the association of the SASC-R subscales with youths'peer status (popular, rejected, neglected, and average). The SASC-R makes new contributions to understanding children's anxiety. Its construct is clear and distinguishes two conceptually and clinically relevant forms of social anxiety: SAD-G and SAD-New. This approach to anxiety is novel and relevant to children's self-perception. However, data are needed regarding its use with normative samples, clinical samples, the clinical relevance of screening social anxiety in school children, and overall utility.

3.4 Fear Survey Schedule for Children-Revised (FSSC-R)

The Fear Survey Schedule for Children-Revised (FSSC-R) (Ollendick, 1983, 2002; Ollendick, Matson, & Helsel, 1985; Ollendick, King, & Frary, 1989) is a revision of the original FSSC developed in the 1960s. Another revision, the FSSC-II, performs comparably with the FSSC-R but is not as popular (Burnham & Gullone, 1997; Gullone & King, 1992). The Fear Survey Schedule for Children-Revised (FSSC-R) is a widely used self-report questionnaire that purports to measure the number of fears and the overall level of fearfulness in children. A number of studies have shown that the ten most common childhood fears can be found on the Danger and Death subscale of the FSSC-R, with upwards of 50% of children endorsing such fears. However, some researchers have questioned the validity of these findings, suggesting that these items do not reflect actual childhood fears that children have or experience on a daily or regular basis. Assessment of the construct tapped by the FSSC-R indicates that children may endorse their affective response to the image of the stimulus rather than their actual fear responses (McCathie & Spence, 1991). The FSSC-R assesses both the number and intensity of fears. Its five

subscales are invariant across gender and age: Fear of Failure and Criticism, Fear of the Unknown, Fear of Injury and Small Animals, Fear of Danger and Death, and Medical Fears. All items are highly intercorrelated, questioning the utility of the subscales, and either the total scale or subscales may be used depending upon intended application (Ollendick et al., 1989). Internal and test–retest reliabilities are well established. Girls endorse the same fears as boys, but endorse greater intensity of these fears; younger children also report greater fearfulness (Burnham & Gullone, 1997; Gullone & King, 1993; Spence & McCathie, 1993). The FSSC-R has discriminated phobic children from controls and among various phobias (Weems, Silverman, Saavedra, Pina, & Lumpkin, 1999). Fears of failure and criticism have best discriminatory power (King, Gullone, & Ollendick, 1992). However, boys with anxiety disorders have not endorsed fears different from boys with disruptive disorders, nor even from normal boys (Perrin & Last, 1992). Both the FSSC-R (Weems et al., 1999) and the FSSC-II (Bouldin & Pratt, 1998) have been modified for parent-report about their children. Many studies have established the psychometric properties and utility of the FSSC-R and have shown that the intensity, frequency, and pattern of fears have remained remarkably stable over time (Spence & McCathie, 1993). The FSSC-Hawaii (FSSC-HI) (Muris, & Ollendic, 2002) is a modified version that includes a number of contemporary fear stimuli and situations (e.g., "drugs", "being raped", "AIDS"). The psychometric properties of the FSSC-HI were examined in a large sample of Belgium adolescents (n=551) aged 12-19 years. Results showed that a five- and seven-factor model both provided satisfactory fits for the structure of the FSSC-HI. Furthermore, the internal consistency of the scale was good and this appeared to be true for the five-factor as well as the seven-factor solution. Support was found for the convergent validity of the FSSC-HI. That is, FSSC-HI scores correlated in a meaningful way with scores on alternative measures of childhood anxiety. Finally, a considerable number of the "new" fear items were found to rank high in the top 10 of most common fears.

3.5   Children's Depression Inventory (CDI)

The Children's Depression Inventory (CDI) (Kovacs, 1985, 1992) is a five-factor solution inventory for both children and adolescents: i) dysphoric mood, ii) acting out, iii) loss of personal and social interest, iv) self-deprecation, and v) vegetative symptoms (Craighead, Curry, & Ilardi, 1998; Weiss, Weisz, Politano, Carey, Nelson, & Finch, 1991). A major developmental difference is the extent to which externalizing behavior is part of the depression construct tapped by the CDI. Children's negative cognitions are related to their misbehaviours and are endorsed as depression. For adolescents, these two issues appear independent. Factors (i), (iii), and (iv) appear most related to depression, while factor (ii) relates to acting out, and factor (v) to anxiety (Hodges & Craighead, 1990; Hodges, Siegel, Mullins, & Griffin, 1983). The role of the factors is unclear and total scores are generally preferred (Craighead et al., 1995). Internal consistency is adequate (Crowley, Thompson, & Worchel, 1994; Joiner, Schmidt, & Schmidt, 1996). Test–retest reliability is highly variable and somewhat lower for boys than girls and for community than for psychiatric youths, as might be expected by the instability of depressive feelings in the general population (Joiner, Schmidt, & Schmidt, 1996; Kovacs, 1985, 1992; Nelson & Politano, 1990). Similar to other such measures, there is poor child–adult concordance, which improves as children mature (Renouf & Kovacs, 1994). Validity has been supported by correlations with multiple other scales, including those measuring related constructs such as self-esteem, cognitive distortions, locus-of-control, attributional style, and underachievement and the predictive validity for future functioning (Mattison, Handford, Kales, & Goodman, 1990). The CDI has been used in many developmental, psychosocial, and ethnic studies. It has shown that during early adolescence, pubertal status is a better predictor of depressive symptoms than chronological age in white, but not African-

Journal of Physical Activity, Nutrition and Rehabilitation, 2016
www.panr.com.cy
info.panr@unic.ac.cy

American or Hispanic, girls (Hayward, Gotlib, Schraedley, & Litt, 1999). The CDI has also been used to reexamine the conceptual aspects of juvenile depression with the suggestion that environmentally impoverished youths experience a complex response to their environments and special needs rather than a biological disease (Barreto & McManus, 1997; Menke, 1998).

3.6   Children's Depression Rating Scale-Revised

The Children's Depression Rating Scale-Revised (CDRS-R) (Poznanski &Mokros, 1999) is a clinician-rated scale developed specifically for children. It was based on the hypothesis that depression in childhood would present more homogeneously than adolescent depression that would combine features of childhood and adult depression However, the CDRS-R is widely used with teenagers. The CDRS-R has three unique features: it integrates information from multiple sources, incorporates behavior during the interview, and several items are not specific to depression. Thus, CDRS-R criteria differ from DSM criteria for depression. The clinician completes the scale independently with the parent and child, producing three scores: parent's, child's, and a combined score. The CDRS-R is unidimensional with six clusters. Over multiple samples, the internal reliability has been described as adequate, and test– retest reliability is good. Even more encouraging is the good to excellent concordance, or interrater reliability, among examiners, consistent with the good concordance demonstrated with structured interviews. Concurrent validity has been supported by correlations with multiple types of ratings (King et al., 1997; Shain, Naylor, & Alessi, 1990). As with other scales, parent–child agreement varies (Mokros, Poznanski, Grossman, & Freeman, 1987). Clinicians rated school children higher than their parents rated them on mood, suicidality, and sleep problems and lower for social withdrawal and somatization. However, in a mood disorders clinic parent's scores were higher than children's scores. Overall, parent–child concordance and clinical status appear to affect clinician's ratings. The CDRS-R has been widely used in research. It has been sensitive to medication effects, (Emslie, Rush, Weinberg, Gullion, Rintelmann, & Hughes, 1997), in phenomenological studies, has revealed features of depression in diverse populations, such as school refusers (Borchardt, Giesler, Bernstein, & Crosby, 1994), youths exposed to violence (Freeman, Mokros, & Poznanski, 1993), and international samples (Canals, Marti-Henneberg, Fernandez-Ballart, & Domenech, 1995; Sharan, Mehta, & Chaudhry, 1999) and has elucidated the course of depression in adolescents with chronic versus acute depressions (Shain, King, Naylor, & Alessi, 1991).

3.7   Child Behavior Checklist (CBCL) and Teacher Report Form (TRF)

The Child Behavior Checklist (CBCL) (Achenbach, 1991; ASEBA; Achenbach, 2001) and Teacher Report Form (TRF) (Achenbach, 1991, 2001) are among the most widely used measures of children's emotional and behavioral problems in both clinical and research settings. The items measure three broad-band scales: Internalizing, Externalizing, and Total Problems, and eight syndrome scales: Withdrawal, Somatic Problems, Anxiety/Depression, Social Problems, Thought Problems, Attention Problems, Delinquent Behavior, and Aggressive Behavior (Achenbach, 1991). The items on both CBCL and TRF, were rated as Not True (0), Somewhat or Sometimes True (1), or Very True or Often True (2), and summed to yield (a) eight syndrome scale scores, (b) six DSM-oriented scale and (c) broad-band scale scores (including internalizing and externalizing total scores). With well-established normative data and standardized clinical cutoffs, the instruments have demonstrated strong psychometric properties (Achenbach, 1991; Chen, Faraone,

Biederman, & Tsuang, 1994; Drotar, Stein, & Perrin, 1995) and good test-retest and inter-rater reliability (Achenbach, 1991). Manuals of the CBCL and the TRF present several kinds of evidence for the validity showing that the item scores, the syndrome scores, and the clinical cut points all significantly discriminate between referred for services and non-referred students (Achenbach, 1991).

3.8 Motor Behavior Checklist for children (MBC)

The Motor Behavior Checklist for children (MBC) (Efstratopoulou, Janssen, & Simons, 2012) is a scale designed to be completed by the primary school teachers and/or physical educators who know the child well enough to rate his/her motor related behavior. Responders are asked to observe the child during physical education classes and free play situations and to rate each behavior on a 5-point Likert scale ranging from "never" (0) to "almost always" (4). The final version of the MBC for children consisted of 59 motor related behavior items included in two broadband factors (Externalizing and Internalizing) and seven problems scales: Rules breaking (7 items), Low energy (4 items), Stereotyped behaviours (2 items), Hyperactivity/Impulsivity (14 items), Lack of attention (10 items), Lack of social interaction (10 items), and Lack of self-regulation (12 items).

The development of the MBC for children involved three different phases. During the first phase, primary school teachers (PE) were asked to report the full spectrum of deviant motor related behavior they observed during teaching hours and to describe the most frequent and troublesome behavior. During the second phase, diagnostic criteria that describe observable motor-related behavior that can occur in school settings were selected from the DSM-IV (American Psychiatric Association, 2000) and the ICD-10 (World Health Organization, 1992), by a team of experts in adapted physical activity and psychomotor therapy were selected. In a third phase a cluster analysis was used in order to analyze the main categories of deviant behaviour discerned by the PE teachers and to investigate the overlap between educators believes in children's emotional and/or behavioral disorders and the official diagnostic criteria in children's psychopathology. Confirmatory Factor Analyses (CFAs) were used to examine the construct validity of the MBC. A normative database of primary students (N=841), rated by their physical educators in school settings, were used to identify the factor structure of the MBC list and to investigate the internal consistency, the reproducibility and the inter-rater agreement.

A series of Confirmatory Factor Analyses (CFAs) revealed a second order model with two (Externalizing and Internalizing) broadband domains and seven problems scales. Alpha values for all the subscales were excellent (Efstratopoulou, Janssen, & Simons, 2012a) suggesting that the list was homogenous in content. More specifically, for the factor Rules breaking (7 items), alpha coefficient was .95, for factor Low energy (4 items), alpha=.82, for factor Stereotyped behavior (2 items), alpha=.85, for factor Hyperactivity/Impulsivity (14 items), alpha=.95, for factor Lack of attention (10 items), alpha=.95, for factor Lack of social interaction (10 items), alpha=.94 and finally for factor Lack of self-regulation (12 items), the alpha coefficient was .91. In addition, for the externalizing scale (31 items), alpha coefficient was .93, and for the internalizing scale (28 items) the alpha coefficient was .91. MBC for children is a new practical valid and reliable instrument which gives the educators the ability to asses an array of problematic behaviours in their students, providing separate scores on different problem scales and stretching the attention to the warning signs of the most problematic domains.

4 Developmental Disorders

Educational research indicated that autism may not be an excessively rare disorder (Volkmar, Lord, Bailey, Schultz, & Klin, 2004), but it could represent the extreme of a

Journal of Physical Activity, Nutrition and Rehabilitation, 2016
www.panr.com.cy
info.panr@unic.ac.cy

quantitative distribution of autistic traits that are present in the general population (e.g. Spiker, Lotspeich, Dimiceli, Myers, & Risch, 2002; Constantino & Todd, 2003). Problem behaviours observed with autism include physical aggression, self-injury, property destruction, stereotyped behaviours, and tantrums are highly disruptive to classroom, community, and home environments and without intervention, they are more likely to increase than improve (Horner, 2002). Children with ASD, indicate stereotyped and repetitive motor mannerisms, impairments of facial expression, postures, and gestures, and are often characterized as clumsy and as having problems in motor coordination (Berkeley, Zittel, Pitney & Nichols, 2001; Piek, & Dyck, 2004). Autistic traits are widely distributed in the general population, and there are many unselected children by the lack of appropriate screening instruments (Skuse, Mandy, & Scourfield, 2005).

Due to the effectiveness of early intervention on the outcome of individuals with ASD, here is a race to identify children with ASD at younger ages (Matson, Boisjoli, Hess, & Wilkins, 2010). For this reason, a top priority in the field of autism is the development of precise early diagnostic tools that are designed to assess symptoms of ASD in young children. The Baby and Infant Screen for Children with autism Traits-Part 1 (BISCUIT-Part 1) (Matson et al., 2010), the Modified Checklist for Autism in Toddlers (M-CHAT) (Robins, Fein, & Barton, 1999) and the Childhood Autism Rating Scale (CARS) (Schopler, Reichler, &Renner, 1988) are among the most popular screening instruments designed to screen for ASD in young children. The instruments consider examiner's observations and parent's responses concerning children's social development, attention and ability to use imaginative play skills in order to determine whether the child in question appears to be at risk for a PDD like autism.

4.1 The Baby and Infant Screen for Children with autism Traits (BISCUIT)

The Baby and Infant Screen for Children with autism Traits-Part 1 (BISCUIT-Part 1) (Matson et al., 2010), is an informant-based measure designed to aid in the diagnosis of ASD in toddlers. The measure contains 62 items scored on a 3-point Likert-type scale. Informants are asked to rate the child on each item, comparing them to a typically developing child as: a) not different; no impairment (score=0), b) somewhat different; mild impairment (score=1) or c) very different; severe impairment (score=2). The BISCUIT Part 1 meets the standard for good internal consistency with an internal reliability coefficient of .97 (Matson, Wilkins, et al., 2009b). Scores from each item are tallied and children receiving scores above 17 are considered in the 'at-risk' range for an ASD, warranting further evaluation. Validity studies have revealed excellent sensitivity and specificity (Matson, Wilkins, Sevin, Knight, Boisjoli, & Sharp, 2009a). Reliability analyses have been conducted with the BISCUIT–Part 1 and excellent internal consistency has been established (Matson, Boisjoli, Hess, & Wilkins, 2010).

4.2 Checklist for Autism in Toddlers (CHAT, M-CHAT)

The Checklist for Autism in Toddlers (CHAT) is a screening test that a pediatrician may use to determine if she should send a child on to a specialist psychologist or psychiatrist for further assessment. The original version, the M-CHAT (Robins, Fein, & Barton, 1999) was developed by neuropsychologists Diana Robins and Deborah Fein and clinical psychologist Marianne Barton. The CHAT is designed to be used with children who are at least eighteen months old. It is filled out by an examiner who answers questions based on personal observation of subject children, and on parental or guardian reports. The test addresses children's social development and also tests their ability to simultaneously

focus on an object that another person is also paying attention to (joint attention). It considers children's ability to use imaginative play skills and their ability to point to objects on command. It measures eye contact and social reciprocity. The examiner considers his/her own observations and the parents' responses to nine questions concerning their child's behavior to determine whether the child in question appears to be at risk for a PDD like autism. The push for earlier and earlier autism diagnoses has resulted in the Modified Checklist for Autism in Toddlers (M-CHAT) (Robins, Fein, & Barton, 1999) which is a shortened version of the CHAT designed to be filled out by parents. The questionnaire focuses on stereotyped movements, social reciprocity and imaginative play. Parents are directed to seek professional attention if their child's answers suggest symptoms of autism may be present. The Modified Checklist for Autism in Toddlers-Revised (M-CHAT-R) is a scientifically validated tool for screening children between 16 and 30 months of age that assesses risk for autism spectrum disorder (ASD). The revision, which improves specificity, was released in December 2013. The American Academy of Pediatrics (AAP) recommends that all children receive autism screening at 18 and 24 months of age, and the M-CHAT-R is one of the AAP's recommended tools.

4.3   Childhood Autism Rating Scale (CARS)

The Childhood Autism Rating Scale (CARS) (Schopler, Reichler, &Renner, 1988) is the strongest, best-documented, and most widely used clinical rating scale for behaviours associated with autism. It has been used in studies all around the world and translated into many languages (Nordin, Gillberg, & Nyden, 1998; Sponheim, 1996). It consists of 15 items on which children and adults are rated, generally after observation, on a 4-point scale. The scale requires minimal training. Training is available on videotape or in brief workshops. Points are added and a standard cut-off of 30 has been suggested and validated with various samples (Garfin et al., 1988). The five domains include: relating to others, body use, adaptiveness to change, listening response and verbal communication. The test examiner answers these questions after observing the child subjects' behavior, reviewing reports concerning the child's behavior, and interviewing the parents, and then computes a CARS score which is compared to normative data that describe how normal children and children known to have PDD diagnoses score on the test. Most of the information about the CARS is from studies of autistic children who function in the mild to moderate range of mental handicap.

CARS has been shown to discriminate autistic children from children without autism and some mental handicap (Schopler et al., 1988; Teal & Wiebe, 1986). Convergence between the CARS and the Autism Diagnostic Interview (ADI) was good for autistic children, but less good for young, non-autistic mentally handicapped children (Lord, 1995). Thus, the evidence that the CARS list accurately identifies children with autism is stronger than the evidence that it discriminates between children with autism and mental-age matched children with other disorders. Besides direct observation by a clinician, for which the CARS was designed, it has also been used in chart review, scored directly by parents and teachers, and used as part of a parent interview (Schopler et al., 1988). CARS does not diagnose autism, but it does help identify who acts like an autistic person. A child's CARS score thus helps examiners to know whether that child's behavior is most similar to typical children's behavior, mildly autistic children's behavior, or severely autistic children's behavior.

5   Conclusions and Recommendations

Children's mental health problems have emerged from a long history of misunderstanding and neglect to become the central concern of researchers, educators and practitioners.

Journal of Physical Activity, Nutrition and Rehabilitation, 2016
www.panr.com.cy
info.panr@unic.ac.cy

The last few decades witnessed an explosion of knowledge about the nature of disorders that affect children, their frequency of occurrence, their developmental course, and the effectiveness of treatments.

The purpose of assessment in fields such as education, child welfare, and mental health has almost always been to inform decision making. Choices made on behalf of individuals or groups have generally been related to eligibility or programming concerns. Consequently, assessment related questions have been both varied and complex. This is particularly true in the assessment of individuals with challenging behaviours.

In the domain of educational and mental health programming, questions range from qualification for special education services to diagnosis of psychiatric disorders. Once established, there are a host of ongoing questions from what type of behavioral, social, academic, and vocational supports might best serve an individual child to what kind of treatment program might best meet a child's mental health needs and the program effectiveness over the short and long term. As a result, it is critical that the decisions made on behalf of children with emotional and behavioral disorders be based on accurate assessment data. A significant amount of time and energy must be enlisted in collecting data that can be used to make accurate assessment-related decisions. Thus, important questions in the process of assessing children's educational and behavioral status in schools become: What information should be collected? What methods should be used? From whom should the information be collected? This review indicates that selection of the optimal checklist or rating scale for a particular application may be difficult. In addition, the study highlights the evidence for assessment instrument for children with emotional, behavioral and developmental disorders but also reveals limitations and clear directions for needed research.

Many studies seem to choose a scale based on its popularity. Given that the psychometric properties of all of the instruments reviewed here are sound, our consideration of the advantages and limitations of each method emphasizes the utility of the approach. Because the determination of emotional and behavioral disorders is problematic, it is imperative that steps be taken to maximize the technical adequacy of the formal instruments used in the assessment process. That is, the tools incorporated in the assessment should be highly reliable and valid so that useful data are gathered for decision-making purposes. One of the biggest challenges for researchers, educators, practitioners and parents is accurate assessment in order to distinguish between normal developmental changes and the emergence of a disorder (atypical changes) determining whether early signs of a disorder will emerge as a full-blown disorder, or resolve into healthy functioning. So, before selecting an instrument for educational or research purposes take these into considerations:

  a) Considering first rating scales as a group, characteristics that likely contribute to their ubiquitous use include that they are easy to administer and score, take little rater or clinician time, and are cost-efficient, allowing the clinician to obtain information from multiple raters across settings. Most also have parallel parent forms and short screening forms. This is especially important for scales assessing externalizing and internalizing disorders because they tend to overlap in their constructs and symptom profiles. To obtain a complete and reliable assessment of the child's adaptive and problem behavior, it is important to have multiple informants and measures that may be used longitudinally.
  b) Because of the modest correlation between children's disorders and impaired functioning, assessment of behavioral symptoms must include evaluation of the

child's functioning in the key domains of peer and sibling and parent and teacher relationships, academic progress, and the classroom and family. There are numerous measures available that have been validated for each of these domains. Brief assessments of those domains may be sufficient for both research and clinical purposes (i.e., those included on measures of global impairment), but additional research is needed. Objective assessment of functioning can be efficiently accomplished with idiographic measures of daily behaviours (both problematic and adaptive). Studies of the incremental validity of such idiographic measures beyond global ratings of impairment have not yet been conducted and are needed.

c) Concerning internalizing scales, their biggest challenge, and their greatest promise, is their ability to discriminate anxiety disorders from depressive disorders. In any event, no one scale is likely to provide all of the information desired. In general, consistent with general recommendations for obtaining a robust measure of a child's psychopathology, more than one scale should be used to evaluate a specific internalizing construct (Myers & Winters, 2002).

The implication for assessment is clear: Raters from a single source or setting do not provide a comprehensive picture of the current levels of functioning for a child with externalizing or/and internalizing behaviours. Ratings from both parents and teachers are indicated for comprehensive assessment (Power, Costigan, Leff, Eiraldi, & Landau, 2001). The diagnostic phase of assessment should be completed with minimal time and expense so that resources can be focused on other aspects of assessment, particularly treatment planning. We argue that the main focus of assessment should be on target behavior selection, contextual factors, functional analyses, treatment planning, and outcome monitoring.

This study indicates that the selection of the optimal checklist or rating scale for a particular application may be difficult. Many studies seem to choose a scale based on its popularity and performance with children despite limited data regarding its functioning with youths. Thus, when the scale does not demonstrate sensitivity to treatment effects, it is unclear whether there are truly no benefits from treatment, or whether the scale was a poor outcome measure. Over the past decade, many older versions of assessment scales have been revised, reexamined psychometrically, applied to different populations, or simply used widely. Within these guidelines, assessment instruments for children can facilitate research and augment educational practice, having also into consideration that a truly universal and inclusive scale is nearly impossible to create.

References

Achenbach, T. M. (1991). *Manual for the teacher's report form and 1991 Profile.* Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2002). Ten-year comparisons of problems and competencies for national samples of youth self, parent, and teacher reports. *Journal of Emotional and Behavioral Disorders, 10*, 194-203.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA school-age forms and profiles.* Burlington: University of Vermont, Research Center for Children, Youth, and Families.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Barreto, S., & McManus, M. (1997). Casting the net for depression. among ethnic minority children from the high-risk urban communities. *Clinical Psychology Review, 17*, 823–845.

Journal of Physical Activity, Nutrition and Rehabilitation, 2016
www.panr.com.cy
info.panr@unic.ac.cy

Beidel, D. C. (1989). Assessing Anxious Emotions: A Review of psychophysiological assessment in Children. *Clinical Psychology Review, 9,* 717-736.

Berkeley S. L., Zittel L. L., Pitney L.V., & Nichols S. E. (2001). Locomotor and object control skills of children diagnosed with autism. *Adapted Physical Activity Quaterly, 18*, 405-416.

Birmaher, B. (1999). *Screen for child anxiety related emotional disorders (SCARED).* Division of Child Psychiatry, Western Psychiatric Institute and Clinic.

Birmaher, B., Brent, D. A., Chiappetta, L., Bridge, J., Monga, S., & Baugher, M. (1999). Psychometric properties of the Screen for child anxiety related emotional disorders scale (SCARED): a replication study. *Journal of American Academy of Child and Adolescent Psychiatry, 38,* 1230–1236.

Borchardt, C. M., Giesler, J., Bernstein, G. A., & Crosby, R. D. (1994). A comparison of inpatient and outpatient school refusers*. Child Psychiatry and Human Development, 24*, 255–264.

Bouldin, P., & Pratt, C. (1998). Utilizing parent report to investigate young children's fears: modification of the Fear Survey Schedule for Children-II: a research note. *Journal of Child Psychology and Psychiatry, 39,* 271–277.

Bullis, M. & Yovanof, P. (2006). Community employment experiences of formerly incarcerated youth. *Journal of Emotional and Behavioral Disorders, 14*, 71-85.

Burnham, J. J., Gullone, E. (1997). The Fear Survey Schedule for Children-II: a psychometric investigation with American data. *Behaviour Research and Therapy, 35,* 165–173.

Canals, J., Marti-Henneberg, C., Fernandez-Ballart, J., & Domenech, E. (1995). A longitudinal study of depression in an urban Spanish pubertal population. *European Child and Adolescent Psychiatry, 4*, 102–111.

Chen, W. J., Faraone, S. V., Biederman, J., & Tsuang, M. T. (1994). Diagnostic accuracy of the child behavior checklist scales for attention-deficit hyperactivity disorder: a receiver-operating characteristic analysis. *Journal of Consulting and Clinical Psychology, 62*, 1017–1025.

Constantino, J. N. & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Archives of General Psychiatry, 60*, 524-530.

Craighead, W. E., Curry, J. F., Ilardi, S. S. (1995). Relationship of Children's Depression inventory factors to major depression among adolescents. *Psychological Assessment, 7,*171–176.

Crowley, S. L., Thompson, B., & Worchel, F. (1994). The Children's depression inventory: a comparison of generalizability and classical test theory analyses. *Educational and Psychological Measurement, 54,* 705–713.

Drotar, D., Stein, R. E., & Perrin, E. C. (1995). Methodological issues in using the child behavior checklist and its related instruments in clinical child psychology research. *Journal of Clinical Child Psychology*, *24*, 184-192.

Efstratopoulou, M., Janssen, R., & Simons, J. (2012). Differentiating children with Attention-Deficit/Hyperactivity Disorder, Conduct Disorder, Learning Disabilities and Autistic Spectrum Disorders by means of their motor behavior characteristic *Research in Developmental Disabilities, 33,* 196–204.

Efstratopoulou, M., Janssen, R., & Simons, J. (2012a). Agreement among physical educators, teachers and parents on children's behaviours: A Multitrait-Multimethod design approach. *Research in Developmental Disabilities, 33*, 1343–1351.

Eisenberg, N., Fabes, R. A., Guthrie, I. K., & Reiser, M. (2000). Dispositional emotionality 18and regulation: their role in predicting quality of social functioning. *Journal of Personality and Social Psychology*, *78*, 136.

Emslie, G. J., Rush, A., Weinberg, W. A., Gullion, C. M., Rintelmann, J., & Hughes, C. W. (1997). Recurrence of major depressive disorder in hospitalized children andadolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*, 785-792.

Freeman, L. N., Mokros, H., & Poznanski, E. O. (1993). Violent events reported by normal urban school-aged children: characteristics and depression correlates. *Journal of the American Academy of Child and Adolescent Psychiatry, 32, 419–423.*

Geller, D. A., Hoog, S. L., Heiligenstein, J. H., Ricardi, R. K., Tamura, R., Kluszynski, S., Jacobson, J. G., & Fluoxetine Pediatric OCD study team (2001). Fluoxetine treatment for obsessive-compulsive disorder in children and adolescents: a placebo controlled clinical trial. *American Academy of Child and Adolescent Psychiatry, 40*, 773–779.

Ginsburg, G. S., La Greca, A. M., & Silverman, W. K. (1998). Social anxiety in children with anxiety disorders: relation with social and emotional functioning. *Journal of Abnormal Child Psychology, 26,* 175–185.

Gullone, E., King, N. J. (1992). Psychometric evaluation of a revised fear survey schedule for children and adolescents. *Journal of Child Psychology and Psychiatry 33*, 987–998.

Gullone, E., King, N. J. (1993). The fears of youth in the 1990s: contemporary normative data. *Journal of Genetic Psychology, 154*, 137–153.

Hayward, C., Gotlib, I. H., Schraedley, P. K., & Litt, I. F. (1999). Ethnic differences in the association between pubertal status and symptoms of depression in adolescent girls. *Journal of Adolescence Health, 25*, 143–149.

Hodges, K., & Craighead, W. E. (1990). Relationship of Children's Depression Inventory factors to diagnosed depression. *Psychological Assessment, 2*, 489–492.

Hodges, K., Siegel, L. J., Mullins, L., & Griffin, N. (1983). Factor analysis of the children's depression inventory. *Psychological Reports, 53*, 759–763.

Horner, R., Carr, E., Strain, P., Todd, A., & Reed, H. (2002). Problem behavior interventions for young children. *Journal of Autism and Developmental Disorders, 32,* 423–446.

Joiner, T. E., Schmidt, K. L., & Schmidt, N. B. (1996). Low-end specificity of childhood measures of emotional distress: differential effects for depression and anxiety. *Journal of Personality Assessment, 67,* 258–271.

Kauffman, J. M., & Landrum, T. J. (2009). Politics, civil rights, and disproportional identification of students with emotional and behavioral disorders. *Exceptionality*, *17*, 177-188.

King, N. J., Gullone, E., & Ollendick, T. H. (1992). Manifest anxiety and fearfulness in children and adolescents. *Journal of Genetic Psychology, 153*, 63–73.

King, C. A., Katz, S. H., Ghaziuddin, N., Brand, E., Hill, E., & McGovern, L. (1997). Diagnosis and assessment of depression and suicidality using the NIMH diagnostic interview schedule for children (DISC-2.3). *Journal of Abnormal Child Psychology, 25*, 173–181.

Kovacs, M. (1985). The Children's depression inventory (CDI). *Psychopharmacology Bulletin, 21,* 995–998.

Kovacs, M. (1992). *Children's depression inventory manual.* Multi-Health Systems.

La Greca, A. M. (1999). *Social anxiety scale for children-revised (SASC-R), manual.* Department of Psychology, University of Miami, Coral Gables.

La Greca, A. M., & Stone, W. L. (1993). Social anxiety scale for children-revised: factor structure and concurrent validity. *Journal of Clinical Child Psychology, 22*, 17–27.

Journal of Physical Activity, Nutrition and Rehabilitation, 2016
www.panr.com.cy
info.panr@unic.ac.cy

Lane, D. (2007). "*Engineering feedback: a student-developed approach to the assessment of peer evaluation in civil engineering.*" Chicago: National Communication Association.

Landrum, T. J., Tankersley, M., James, M., & Kauffman, J. M. (2003). What is Special about special education for students with emotional or behavioral disorders? *The Journal of Special Education, 37,* 148-156.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, 36,* 1365–1382.

March, J. S. (1998). *Manual for the multidimensional anxiety scale for children (MASC)*. Toronto: Multi-Health Systems.

March, J. S., Parker, J. D., Sullivan, K., Stallings, P., & Conners, C. K. (1997). The multidimensional anxiety scale for children (MASC): factor structure, reliability, and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 554–565.

Mash, E. J., & Johnston, C. (1983). Parental perceptions of child behavior problems, parenting self-esteem, and mothers' reported stress in younger and older hyperactive and normal children. *Journal of consulting and clinical psychology*, *51*, 86.

Matson, J. L., Wilkins, J., Sharp, B., Knight, C., Sevin, J. A., & Boisjoli, J. A. (2009a). Sensitivity and specificity of the baby and infant screen for children with autism traits (BISCUIT): validity and cut-off scores for autism and PDD-NOS in toddlers. *Research in Autism Spectrum Disorders, 3,* 336-344.

Matson, J. L., Wilkins, J., Sevin, J. A., Knight, C., Boisjoli, J. A., & Sharp, B. (2009b). Reliability and item content of the baby and infant screen for children with autism traits (BISCUIT): parts 1, 2, and 3. *Research in Autism Spectrum Disorders, 3* , 336–344.

Matson, J. L., Boisjoli, J. A., Hess, J. A., & Wilkins, J. (2010). Factor structure and diagnostic fidelity of the baby and infant screen for children with autism traits-Part 1 (BISCUIT23, part 1). *Developmental Neurorehabilitation, 13,* 72–79.

Mattison, R. E., Handford, H. A., Kales, H. C., & Goodman, A. L. (1990). Four-year predictive value of the children's depression inventory. *Psychological Assessment, 2,* 169–174.

McCathie, H., & Spence, S. H. (1991). What is the revised fear survey schedule for children measuring? *Behaviour Research and Therapy, 29*, 495–502.

Menke, E. M. (1998). The mental health of homeless school-age children. *Journal of Child and Adolescence Psychiatry Nursing, 11*, 87–98.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., & Reed, G. M. (2001). Psychological testing and psychological assessment: A review of the evidence and issues. *American Psychologist, 56,* 128–165.

Mokros, H. B., Poznanski, E., Grossman, J. A., & Freeman, L. N. (1987). A comparison of child and parent ratings of depression for normal and clinically referred children. *Journal of Child Psychology and Psychiatry, 28*, 613–627.

Morris, R. J., Shah, K. & Morris, Y. P. (2002). Internalizing behavior disorders. *Interventions for children with or at risk for emotional and behavioral disorders*, *15*, 223-241.

Muris, P., Bjorg, G., Moulaert, V., & Merckelbach, H. (1998a). Correlations between two multidimensional anxiety scales for children. *Perceptual and Motor Skills, 87*, 269–270.

Muris, P., Merckelbach, H., Brakel, A., Mayer, B., Dongen, L. (1998b). The Screen for Child Anxiety Related Emotional Disorders (SCARED): relationship with anxiety and depression in normal children. Personality and Individual Differences, 24, 451–456.

Muris, P., Merckelbach, H., Brakel, A., Mayer, B., & Dongen, L. (1999). The Revised Screen for Child Anxiety Related Emotional Disorders (SCARED-R): factor structure in normal children. *Personality and Individual Differences, 26*, 99–112.

Myers, K., & Winters, N. C. (2002). Ten-year review of rating scales. II: Scales for internalizing disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(6), 634-659.

Nelson, W. M., & Politano, P. M. (1990). Children's Depression Inventory: stability over repeated administrations in psychiatric inpatient children. *Journal of Clinical Child Psychology, 19,* 254–256.

Nordin, V., Gillberg, C., & Nyden, A. (1998). The Swedish version of the Childhood Autism Rating Scale in a clinical setting. *Journal of Autism and Developmental Disorders, 28,* 69–75.

Ollendick, T. H. (1983), Reliability and validity of the Revised Fear Survey Schedule for Children (FSSC-R). *Behaviour Research and Therapy 21*, 685–692.

Ollendick, T. H., Matson, J. L., & Helsel, W. J. (1985). Fears in children and adolescents: normative data. *Behaviour Research and Therapy 23*, 465–467.

Ollendick, T. H., King, N. J., & Frary, R. B. (1989). Fears in children and adolescents: reliability and generalizability across gender, age, and nationality. *Behaviour Research and Therapy, 27,* 19–26.

Ollendick, T. H. (2002). *The Fear Survey Schedule for Children-Revised.* Department of Psychology, Child Study Center, Virginia Polytechnic Institute and State University, Blacksburg.

Perrin, S., Last, C. G. (1992). Do childhood anxiety measures measure anxiety*? Journal of Abnormal Child Psychology, 20,* 567–578.

Piek, J.P., & Dyck, M.J. (2004). Sensory-motor deficits in children with developmental coordination disorder, attention deficit hyperactivity disorder and autistic disorder. *Human Movement Sciences, 23*, 475–88.

Power, T. J., Costigan, T. E., Leff, S. S., Eiraldi, R. B., & Landau, S. (2001). Assessing ADHD across settings: Contributions of behavioral assessment to categorical decision making. *Journal of Clinical Child Psychology, 30*, 399–412.

Poznanski, E. O., & Mokros, H. B. (1999). *Children's Depression Rating Scale-Revised* (CDRS-R). Los Angeles: Western Psychological Services.

Renouf, A. G., & Kovacs, M. (1994). Concordance between mothers' reports and children's self-reports of depressive symptoms: a longitudinal study. *Journal of American Academy of Child and Adolescence Psychiatry, 33,* 208–216.

Research Unit on Pediatric Psychopharmacology Anxiety Study Group (2001). Fluvoxamine for the treatment of anxiety disorders in children and adolescents. *The New England Journal of Medicine, 344,* 1279–1285.

Robins, D. L., Fein, D., & Barton, M. L. (1999). *The modified checklist for autism in toddlers* (M-CHAT). Storrs, CT: Selfpublished.

Schopler, E., Reichler, R., & Renner, B. (1988). *The Childhood Autism Rating Scale* (CARS). Los Angeles: Western Psychological Services.

Skuse, H. D., Mandy, L. P. W., & Scourfield, J. (2005). Measuring autistic traits: heritability, reliability and validity of the Social and Communication Disorders Checklist. *British Journal of Psychiatry, 187,* 568-572.

Snyder, J. (2004). Child impulsiveness-inattention, early peer experiences, and the development of early onset conduct problems. *Journal of Abnormal Child Psychology, 32*, 579–594.

Sponheim, E. (1996). Changing criteria of autistic disorders: A comparison of the ICD-10 research criteria and DSM-IV with DSM-III-R, CARS, and ABC. *Journal of Autism and Developmental Disorders, 26*, 513–525.

Journal of Physical Activity, Nutrition and Rehabilitation, 2016
www.panr.com.cy
info.panr@unic.ac.cy

Spence, S. H., & McCathie, H. (1993). The stability of fears in children: a two-year prospective study: a research note. *Journal of Child Psychology and Psychiatry, 34*, 579–585.

Spiker, D., Lotspeich, L. J., Dimiceli, S., Myers, R., & Risch, N. (2002). Behavioral phenotypic variation in autism multiplex families: evidence for a continuous severity gradient. *American Journal of Medical Genetics, 114*, 129-136.

Shain, B. N., King, C. A., Naylor, M., & Alessi, N. (1991). Chronic depression and hospital course in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 428–433.

Sharan, P., Mehta, M., & Chaudhry, V. P. (1999). Psychiatric morbidity in children suffering from acute lymphoblastic leukemia*. Pediatric Hematology and Oncology, 16*, 49–54.

Teal, M. B., & Wiebe, M. J. (1986). A validity analysis of selected instruments used to assess autism. *Journal of Autism and Developmental Disorders, 16*, 485–494.

Volkmar, F. R., Lord, C., Bailey, A., Schultz, R. T. & Klin, A. (2004). Autism and pervasive developmental disorders. *Journal of Child Psychology and Psychiatry, 45*, 135-170.

Wagner, M., & Davis, M. (2006). How are we preparing students with emotional disturbances for the transition to young adulthood? Findings from the National Longitudinal Transition Study—2. *Journal of Emotional and Behavioral Disorders*, *14*, 86-98.

Wagner, M., Newman, L., Cameto, R., Levine, P., & Garza, N. (2006). *An Overview of Findings from Wave 2 of the National Longitudinal Transition Study-2* (NLTS2). National Center for Special Education Research.

Weems, C. F., Silverman, W. K., Saavedra, L. M., Pina, A. A., & Lumpkin, P. W. (1999). The discrimination of children's phobias using the revised fear survey schedule for children. *Journal of Child Psychology and Psychiatry*, *40*, 941–952.

Weiss, B., Weisz, J., Politano, M., Carey, M., Nelson, W., & Finch, A. (1991). Developmental differences in the factor structure of the Children's Depression Inventory. Psychological Assessment, *3*, 38–45.

World Health Organization. (1992). *The IDC-10 classification of mental and behavioral disorders: Diagnostic criteria for research*. Geneva: Author.

Zigmond, N. (2006). Twenty-four months after high school paths taken by youth diagnosed with severe emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*, *14*, 99-107.