Running head. Validation of the CESD-ID-*R*

Validation of a Revised Version of the Center for Epidemiologic Depression Scale for Youth with Intellectual Disabilities (CESD-ID-*R*)

Elizabeth Olivier, Département de psychopédagogie et d'andragogie, Université de Montréal, Canada
Corina Lacombe, School of Psychology, University of Ottawa, Canada
Alexandre J.S. Morin, Substantive-Methodological Synergy Research Laboratory, Department of Psychology, Concordia University, Canada
Simon A. Houle, Substantive-Methodological Synergy Research Laboratory, Department of Psychology, Concordia University, Canada
Cynthia Gagnon, Département de psychoéducation et de psychologie, Université du Québec en Outaouais, Saint-Jérome, Canada
Danielle Tracey, School of Education, Western Sydney University, Sydney, Australia
Rhonda G. Craven, Institute for Positive Psychology and Education, Australian Catholic University, Sydney, Australia
Christophe Maïano, Cyberpsychology Laboratory and Département de psychoéducation et de psychologie, Université du Québec en Outaouais, Saint-Jérome, Canada

Acknowledgements

This study was supported by grants from the Australian Research Council (DP140101559) and from the Social Science and Humanities Research Council of Canada (430-2012-0091, 435-2014-0909). In the preparation of this paper, the first author was also supported by funding from the Canadian Institute for Military and Veteran Health Research (CIMVHR), the third author by a grant from the Social Science and Humanities Research Council of Canada (435-2018-0368), and the fourth author by a doctoral scholarship from the Social Science and Humanities Research Council of Canada.

Corresponding author:
Alexandre J.S. Morin,
Department of Psychology, Concordia University, Canada,
7141 Sherbrooke W, Montreal, QC, Canada, H4B 1R6
alexandre.morin@concordia.ca

This is the prepublication version of the following manuscript:

Olivier, E., Lacombe, C., Morin, A.J.S., Houle, S.A., Gagnon, C., Tracey, D., Craven, R.G., & Maïano, C. (In Press, 2021). Validation of a revised version of the Center for Epidemiologic Depression Scale for youth with intellectual disabilities (CESD-ID-*R*). *Journal of Autism and Developmental Disorders*. Early View doi: 10.1007/s10803-021-05334-9 [ISI 2020: 4.291]

Abstract

Aims. This study proposes a revision (*R*) of the Center for Epidemiologic Studies Depression Scale for youth with ID (CESD-ID) in English and French.
Methods. 346 youth (36.02% girls) with mild (51.26%) and moderate (48.78%) ID (11-22 years; M=15.69), enrolled in secondary schools in Canada (French-speaking; N= 115), and Australia (English-speaking; N= 231), as well as their parents and teachers, participated in this study.
Results. Results support the reliability, factor validity, equivalence (sex, ID level, comorbidities, and country), and convergent validity (with youth-, parent-, and teacher-rated measures of depression, anxiety, and loneliness/social isolation) of the CESD-ID-*R*.
Conclusions. The CESD-ID-*R* allows youth with ID to provide a reliable and valid assessment of their depressive mood and happiness suitable for epidemiological studies.
Keywords: CESD-ID; depression; intellectual disability; special education needs; measurement; confirmatory factor analysis; scale validation.

Up to 14% of young people suffer from depression (Costello et al., 2011; ISQ, 2018), placing depression as one of the top 10 causes of disability-adjusted life years for adolescents and young adults (GBD 2019 Diseases and Injuries collaborators, 2020; Weinberger et al., 2018). Although similarly high rates are observed among typically developing (TD) youth and youth with an intellectual disability (ID) (Maïano et al., 2018; Polancyk et al., 2015), little research has specifically focused on depression among youth with ID, a population vulnerable to many biopsychosocial difficulties (Berg et al., 2015). This research shortage can be partly attributed to the challenges associated with measuring internal states among youth with ID, as well as to the relative unavailability of short and easy to use self-report measures of depression validated in English and other languages for this population. Therefore, several studies rely on informants' reports, usually a parent, teacher, or clinician (Rojahn et al., 2011), despite depressive symptoms being hard to discern externally (Smith, 2007; Turk et al., 2012).

This insufficient understanding of depression as experienced by youth with ID is problematic in light of the repeated calls to empower people with ID by providing them with a means to enable their agency and express their own voice in research focusing on their internal reality (Hartley & MacLean, 2006). As long as the bulk of research on depression among youth with ID remains focused on informant reports, it is hard to verify whether the results truly reflect the reality or simply informants' expectations about the nature, drivers, and outcomes of depression among this population. In an effort to provide researchers with tools for the measurement of depressive symptoms among youth with ID, this study proposes a revised version of the Center for Epidemiologic Studies Depression Scale for youth with ID (CESD-ID), the CESD-ID-*R*. As the CESD-ID was initially developed in French, we also propose equivalent English and French versions.

Depression Among Youth With and Without ID

Depression encompasses a range of manifestations, including depressive mood, loss of interest and pleasure, feelings of guilt and worthlessness, suicidal ideation, difficulties concentrating, and somatic manifestations (weight loss or gain, loss of energy and fatigue, psychomotor retardation or agitation, sleep difficulties) (APA, 2020). As a result, depression interferes with all aspects of youth's academic, familial, and social life (APA, 2020).

Among the TD population, the risk for depression is not distributed equally. Rates of depression are similar between boys and girls during childhood, but become more prevalent among girls at the approach of adolescence, a difference that lasts into adulthood (Costello et al., 2011). Youth with comorbid conditions also present a higher risk of depression than those without (Garber & Rao, 2014). However, it remains unclear whether these differences are found in youth with ID. Some suggest that males and females with ID present a similar risk of depression (Austin et al., 2018; Chester et al., 2013; Whitney et al., 2019), whereas others indicate that this risk is higher among females (Hsieh et al., 2020). Similarly, there is no agreement on whether comorbid conditions increase (Whitney et al., 2019), or not (Hsieh et al., 2020), the risk of depression. Finally, among youth with ID, the risk of depression seems to differ as a function of ID levels (Whitney et al., 2019).

Besides, youth with ID who manifest depressive symptoms tend to have a less positive psychosocial adjustment, including higher levels of anxiety (Glenn et al., 2003; Masi et al., 2002), a co-occurrence also observed when using informant reports (Masi et al., 2002; Rojahn et al., 2011). In turn, these emotional difficulties impact the social interactions of youth with ID, by increasing their tendencies to avoid social interactions (Rojahn et al., 2011), leading to loneliness, social isolation, and even peer victimization (Klein et al., 2018; Olivier et al., 2020).

The Need for Validated Epidemiologic Measures of Depression

Over the years, various measures have been proposed for the assessment of depression among individuals with ID. The most commonly used of those measures, which have been validated for individuals with ID, are presented in Table 1. Readers seeking a more comprehensive coverage of available measures are referred to Hermans and Evenhuis (2010) and Perez-Achiaga et al. (2009). As can be seen from this table, few of these measures are suitable to capture youth's own perceptions of their depressive symptoms in the context of large-scale epidemiological studies of adolescents with ID. Such studies require linguistically simple and relatively short, yet comprehensive measures, of depressive symptoms validated for use as self-reported instruments among youth with and without ID to ensure comparability between these two populations. For instance, the Glasgow Depression Scale for people with Intellectual Disabilities (GDS-ID; Cuthill et al., 2003) is well suited for use among English-speaking adult populations with ID, but has never been validated for use among adolescents,

or among French-speaking populations. Furthermore, the CESD is among the five most frequently used measured of depression among TD populations (Santor et al., 2006), whereas the GDS-ID has never been validated for use among TD populations, suggesting that the CESD might provide a better alternative, in the long run, for studies seeking to compare depressive symptoms among these two populations. Among the scales reported in Table 1, only the CESD / CESD-ID, an instrument specifically developed to address these limitations (Maïano et al., 2011), meets these criteria. Unfortunately, only a French version of this instrument currently exists.

More precisely, Maïano et al. (2011) proposed a French adaptation of the CESD for youth with ID (CESD-ID), a measure whose original version was specifically designed for epidemiologic studies (Morin et al., 2011; Radloff, 1977). The original CESD is a 20-item self-report questionnaire covering four subscales: depressive affect, positive affect, somatic and retarded activity, and interpersonal relationships. After adapting this measure for youth with ID, Maïano et al. (2011) conducted factor analyses using this complete set of 20 items. Matching their intention to devise a shorter version of this instrument, they retained a subset of 14 items for the CESD-ID, following the removal of six items which were associated with problematically high modification indices or covariance residuals. Their results further showed that these 14 items replicated the four-factor structure of the original CESD among a sample of youth with ID, with adequate scale-score reliability, and evidence of measurement invariance as a function of youth sex and age. Maïano et al. (2011) also demonstrated the superiority of a graphically assisted response scale (see Appendix B) relative to a purely verbal response scale.

Despite these promising results, Maïano et al.'s (2011) study presents some noteworthy limitations that the present study seeks to address. First, although the original CESD is validated both in English (Radloff, 1977) and French (Morin et al., 2011), the CESD-ID currently only exists in French (Maïano et al., 2011), and has yet to be validated for English-speaking populations. Second, the correlations reported by these authors between the subscales sharing the same valence (depressive affect with somatic/retarded activity: r=.849; positive affect with interpersonal relationships: r=.732) were high enough to suggest some degree of conceptual overlap. More precisely, the CESD-ID might rather cover two components, one reflecting classical manifestations of depression (i.e., depressive mood) and one reflecting positive functioning (i.e., happiness). Unfortunately, this possibility has never been formally investigated, and the results reported by Maïano et al. (2011) have yet to be replicated. Third, Maïano et al. (2011) assessed if scores obtained on the CESD-ID were comparable (i.e., measurement invariance) across samples of boys and girls with ID, but not as a function of youth characteristics (e.g., ID level and comorbidities). They also did not assess the concurrent validity of this instrument with other indicators of socio-emotional functioning related to depression, or to informant-reported measures of depression. To address these limitations, the present study aims to (1) develop equivalent French and English versions of the CESD-ID, (2) contrast the a priori four-factor structure of the CESD-ID with the alternative two-factor structure suggested by Maïano et al.'s (2011) results; (3) test the measurement equivalence (or invariance; Millsap, 2011) of ratings obtained on this instrument as a function of youth sex, ID level and comorbidity, as well as across linguistic versions (also capturing potential cultural differences between Australia and Canada); and (4) test the convergent validity of ratings obtained on this instrument as a function of youth ratings of depression (obtained on another, longer, measure specific to youth with ID), anxiety, and loneliness, as well as of teachers' and parents' ratings of depression, general anxiety, and social avoidance.

The Present Study

Scale Development and Validation

The present study was designed to assess the psychometric properties of an improved version of the CESD-ID, the CESD-ID-*R*, among a sample of English-speaking Australian youth and French-speaking Canadian youth. Given the correlations reported by Maïano et al. (2011), we expected to identify two conceptually distinct and negatively correlated ($|r| < .500$) factors reflecting depressive mood (combining depressive affect and somatic/retarded activity) and happiness (combining positive affect and interpersonal relationships).

Generalizability Across Youth Characteristics

Our second objective was to ascertain that the psychometric properties of the CESD-ID-*R* remained unchanged (i.e., measurement invariance) as a function of various characteristics of youth with ID (sex, ID level, comorbidities, and country/language). Consistent with our expectation that the CESD-ID-*R* will be generalizable to all types of youth with ID, we expected items to function in the

same manner regardless of these characteristics.

Concurrent Validity

Our third objective was to assess the concurrent validity of the CESD-ID-*R* by testing associations with youth characteristics (sex, ID level, and comorbidities), as well as with youth, teacher, and parental ratings of youth's socio-emotional problems, including measures of depression, anxiety, and loneliness/social avoidance. Consistent with previous studies, we expect that levels of ID will not be associated with youth's levels of depression (Whitney et al., 2019), and pose as an open research question whether youth's levels of depression will be associated with their sex and with the presence of comorbid conditions (Austin et al., 2018; Chester et al., 2013; Hsieh et al., 2020; Whitney et al., 2019). We also expected that youth reporting higher levels of depression on the CESD-ID-*R* will present higher levels of depression, anxiety, and loneliness/social avoidance on alternative self- and informant-reported measures (Glenn et al., 2003; Klein et al., 2018; Rojahn et al., 2011). Consistent with each informant having a unique and complementary perspective of the reality (Turk et al., 2012), we expected scores on the CESD-ID-*R* to be more strongly related to youth's self-reports of socio-emotional difficulties than with teacher and parental reports of the same difficulties.

Methods

Participants

The study was conducted among a sample of youth simultaneously recruited in two countries (Canada and Australia) using identical procedures. The sample comprises 346 youth (39.02% girls) with mild (51.26%: IQ ranging from 35 to 49) and moderate (48.72%; IQ ranging from 50 to 70) levels of ID. ID classifications were determined using IQ scores available in the school records, in line with the DSM-IV (APA, 2000) which was the official classification system used in school records at the time of data collection. Participants were aged 11 to 22 years (M=15.69, SD=2.17), enrolled in secondary schools in Canada (French-speaking, n=115, 33.24%) and Australia (English-speaking, n=231, 66.76%). Teachers (n=241) and parents (n=161) also completed questionnaires. Additional details on the composition of the sample are reported in Table 2.

Procedure

All participants were recruited within schools or community organizations (see Table 2) that agreed to support this research proposal. Parents (or legal representatives) of all participating youth actively provided signed informed consent for their own, and their children's participation. Youth were met at their school (or at a time and location convenient for the parents for participants recruited outside of schools) by trained research assistants who explained the goals and procedures of the study, as well as youth's right not to participate or to withdraw from the study without consequences. Youth were asked to actively and voluntarily consent to the study. Using sample questions for each questionnaire section, research assistants explained how to use the response scales (all involving graphical displays and pictograms). Testing was realized in small groups including up to 8 youth with mild levels of ID or including 1 or 2 youth with moderate levels of ID for youth recruited in school. Youth recruited outside of school were tested individually. A read-aloud assisted procedure was utilized to maximize understanding, and youth were encouraged to ask questions. Sometimes, despite the available support, youth remained unable to understand a question. They were then instructed to select the "do not understand" option (4.1% to 6.3%; M=5.7%; treated as missing values). Parents and teachers were asked to complete a questionnaire sent to them by the school (or by the research team for participants recruited outside of schools). Authorization to conduct the study was obtained from the research ethics committees of the fourth, fifth, and last authors' institutions.

The consent procedure granted researchers access to school records for all youth, including youth's most recent level of intellectual functioning (only youth with an official school-based ID classification were recruited). The Wechsler (2003) Intelligence Scale for Children – Fourth Edition (WISC-IV) was the IQ test most frequently used by the schools in both countries. When the last IQ assessment in the school records was older than four years, a new IQ assessment was conducted by a registered psychologist using the WISC-IV, the Wechsler Adult Intelligence Scale-IV (Wechsler et al., 2018), or the Leiter international performance scale-revised (Roid & Miller, 1997), based on age and verbal ability. In Australia, 34 participants were re-assessed using the Wechsler version corresponding to their chronological age (31 WISC-IV and 3 WAIS-IV). In Canada, 59 participants were re-assessed, 50 using the Wechsler version (21 WISC-IV or 29 WAIS-IV) corresponding to their age, and 9 (with lower verbal expression skills) using the Leiter.

CESD-ID-*R*. First, the French items of the original CESD-ID (Maïano et al., 2011), were translated to English using a translation back-translation procedure by bilingual researchers. Second, items were maximally simplified and underwent minor reformulations to maximize clarity and linguistic equivalence. Third, pictograms were added to the items to improve their clarity. This adaptation was realized through a collaborative process involving bilingual researchers familiar with this process and population, as well as teachers, psychologists, and psycho-educators experienced in working with youth with ID. The complete process of test adaptation and pilot testing leading to the development of the CESD-ID-*R* is described in Appendix A of the online supplements. The verbal formulation and response scale of the CESD-ID-*R* items are reported in Appendix B (where the original items appear in parentheses), and the complete questionnaires are available from the corresponding author.

At both time points, students rated the 14 items from the CESD-ID-*R* on a response scale ranging from 1 (totally disagree) to 4 (totally agree). According to Maïano et al. (2011), these items form four distinct subscales: depressive affect (5 items; α=.820); positive affect (3 items; α=.734); somatic and retarded activity (4 items; α=.442); and interpersonal relationships (2 items; α=.760).

Outcomes: Youth Self-Reports. ***Depression*** was assessed using the 21 items from the Glasgow Depression Scale for people with Intellectual Disabilities (GDS-ID; Cuthill et al., 2003). For present purposes, these items were used to assess youth's depressive mood (16 items; α=.912; e.g., "I feel sad or depressed") and positive feelings (5 items; α=.753; e.g., "I enjoy the things I do and have fun"). Youth rated these items on a response scale ranging from 0 (never) to 4 (always). ***Anxiety*** was assessed using the 27 items from the Glasgow Anxiety Scale for people with Intellectual Disabilities (GAS-ID; Mindham & Espie, 2003). These items form three subscales assessing worries (10 items; α=.855; e.g., "I worry a lot"), specific fears (9 items; α=.836; e.g., "I am scared of dogs"), and physiological symptoms (8 items; α=.868; e.g., "When I am nervous or uncomfortable, my heart beats very fast"), rated on a response scale ranging from 0 (never) to 4 (always). ***Loneliness*** was assessed using 8 items from the School Loneliness Scale (Asher et al., 1984; α=.789; e.g., "I am lonely in my school"), rated on a scale ranging from 1 (totally disagree) to 5 (totally agree).

Outcomes: Informant Reports. Teachers and parents completed three subscales from the Anxiety, Depression, and Mood Screen (ADAMS; Esbensen et al., 2003) to assess youth's depressive mood (7 items; teachers: α=.871; parents: α=.899; e.g., "Over the last month, this student was sad"), general anxiety (7 items; teachers: α=.840; parents: α=.854; e.g., "Over the last month, this student was nervous"), and social avoidance (7 items; teachers: α=.813; parents: α=.810; e.g., "Over the last month, this student was withdrawn from other people"), on a response scale ranging from 1 (not a problem) to 5 (major problem).

Covariates. Youth's sex (0=female; 1=male), country of residence (0=Canada; 1=Australia), and ID level (0=mild; 1=moderate) were obtained via official school records. Among participants, 87 (25.14%) had a reported comorbid condition (coded 0=none; 1=yes).

Analyses

Model Estimation. All analyses were performed using Mplus 8.4 (Muthén & Muthén, 2019) and the robust Weight Least Square estimator (WLSMV), which outperforms Maximum Likelihood estimation with ordinal rating scales following asymmetric response thresholds such as those used in this study (Finney & Di Stephano, 2013). All models were estimated using all the available information, using missing data algorithms implemented in Mplus for WLSMV (Asparouhov & Muthén, 2010) and similar to full information maximum likelihood procedures (Enders, 2010). Measurement models based on youth self-reports (CESD-ID-*R* and outcomes) included 0.29% to 4.91% of missing data. Measurement models involving teacher and parent ratings (outcomes only) respectively included 0% to 1.32% and 0% to 1.15% of missing data.

Measurement Models. First, we estimated a one-factor confirmatory factor analytic (CFA) model in which all 14 items were used to define a single factor. Second, we estimated a four-factor CFA solution corresponding to the factor structure retained by Maïano et al. (2011). In this solution, each item was used to define one of four correlated factors representing depressive affect (items 2, 4, 6, 10, 13), positive affect factor (items 5, 8, 12), somatic and retarded activity (items 1, 3, 7, 9), and interpersonal relationships (items 11, 14). Third, we estimated a two-factor CFA model in which items were used to define two factors differing in their valence, as suggested by the correlations reported by Maïano et al. (2011): (a) depressive mood (items 1, 2, 3, 4, 6, 7, 10, 13); (b) happiness (items 5, 8, 9,

11, 12, 14). A priori correlated uniquenesses (CU) were included to the one-factor (15 CU among the 6 positively worded items) and four-factor (5 CU between item 9, which is positively worded and loaded on a negative factor, and 5 other positively-worded items which loaded on the positive factor) models to account for the methodological artifact caused by the presence of the positively worded items (reflecting the opposite of the depression construct; Marsh et al., 2010).

Measurement Invariance. The equivalence of the retained model was investigated via tests of measurement invariance conducted as a function of youth's sex (male or female), ID level (mild or moderate), presence of a comorbid condition (yes or no), and country (corresponding to linguistic versions: Canada-French or Australia-English). These tests were conducted in the following sequence (Millsap, 2011; Morin et al., 2011): (i) configural invariance (same model with no additional constraint); (ii) equal factor loadings (weak invariance); (iii) equal factor loadings and response thresholds (strong invariance); and (iv) equal factor loadings, response thresholds, and item uniquenesses (strict invariance).

Model Fit. Model fit was assessed using recommended fit indices (Hu & Bentler, 1999; Yu, 2002) and following recent interpretation guidelines (Marsh et al., 2005; Little, 2013): The Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Tucker-Lewis Index (TLI). According to current guidelines (Marsh et al., 2005) RMSEA≤.06 and CFI-TLI ≥.95 reflect an excellent fit, whereas RMSEA≤.08 and CFI/TLI ≥.90 reflect an adequate fit. We also report the composite reliability coefficients (ω; McDonald, 1970). To assess measurement invariance, RMSEA increases >.015 and CFI-TLI decreases >.010 reflect non-invariance (Chen, 2007; Cheung & Rensvold, 2002). Given the oversensitivity of the chi-square (χ^2) test of exact fit, and of chi-square difference tests ($\Delta\chi^2$) to minor (i.e., trivial) misspecifications and to sample size (Hu & Bentler, 1999; Marsh et al., 2005), we report only report these tests for descriptive purposes.

Convergent Validity. The convergent validity of the CESD-ID-*R* factors was tested via the estimation of correlations (Pearson) between scores on the latent continuous factors (thus entirely corrected for measurement errors, Bollen, 1989) from the retained measurement structure of the CESD-ID-*R*, youth's observed characteristics (sex, ID level, and comorbidities), and continuous factor scores (saved from preliminary CFA measurement models, and thus partially corrected for measurement errors; Skrondal & Laake,) reflecting youth (depressive mood, positive feelings, worries, fears, physiological symptoms, and loneliness) as well as teacher and parent (depressive mood, general anxiety, and social avoidance) outcome ratings. Results from the outcomes measurement models are reported in Tables S1 and S2 of the online supplements.

Results

Measurement Models

The fits of the alternative measurement models are reported in Table 3. The three solutions had an almost identical fit to the data, which for reasons of parsimony would appear to support the one-factor solution. However, the parameter estimates from this solution revealed that, even though this single factor was characterized by strong factor loadings from the negatively-valenced items (λ=.547 to .800), it was characterized by very weak loadings from the positively-valenced items (λ=-.018 to .106). In contrast, the two-factor solution resulted in mainly independent (*r*=.074, *p*=.157) and well-defined Depressive Mood (λ=.454 to .801) and Happiness (λ=488 to .900) factors. Finally, although the four-factor solution also resulted in well-defined factors (λ=.473 to .910), the sole item reflecting the opposite of its factor (item 9: "I can do lots of things" on the somatic and retarded activity factor) did not load on its a priori factor (λ=.002). These results replicated those reported by Maïano et al. (2011), revealing very high factor correlations between similarly-valenced factors: (a) Depressive affect with somatic and retarded activity (*r*=.898); (b) positive affect with interpersonal relationships (*r*=.970). Thus, the two-factor solution was retained. Parameter estimates and reliability from this solution are reported in Table 4.

Measurement Invariance

The results from the tests of measurement invariance are reported in Table 3. These results first support the complete measurement invariance (configural, weak, strong, and strict invariance) of this solution across samples of boys and girls, youth with mild and moderate levels of ID, and youth with or without comorbid conditions. Second, comparisons across countries also supported the configural, weak, and strong invariance of this model across samples of French-Speaking Canadian youth and English-speaking Australian youth, but not the strict invariance of this solution. Examination of the

parameter estimates associated with the model of strong invariance and of the modification indices associated with the model of strict invariance revealed that the non-invariance of items' uniquenesses appeared limited to three items. Relaxation of the invariance constraints placed on the uniquenesses of these three items led to a model of partial strict invariance, which was supported by the data. The results of this model reveal that levels of item-specific measurement errors were slightly higher in Canada than in Australia for three items associated with the Happiness factor: (a) item 5 ("I have a lot of hopes for the future"; .885 in Canada vs. .689 in Australia), (b) item 9 ("I can do a lot of things"; .764 vs. .466), and (c) item 11 ("People are nice to me"; .512 vs. 359).

Convergent Validity

The results of analyses of convergent validity (Table 5) indicate that boys reported lower levels than girls on the depressive mood and happiness factors. Relative to youth with mild levels of ID, those with moderate levels of ID also reported a higher level of happiness but did not differ in terms of depressive mood. Having a comorbid condition was not associated with ratings of depressive mood or happiness. Results also reveal higher levels of within-rater agreement (i.e., associations between youth's report on the CESD-ID-*R* and other youth-reported outcomes) than between-rater agreement (i.e., associations between youth's report on the CESD-ID-*R* and teacher- or parent-reported outcomes). First, the depressive mood factor was positively associated with most self-reported indicators of depression and anxiety, with the sole exception of the positive feelings factor. Second, the happiness factor shared negative associations with most self-reported indicators of depression and anxiety, with the exceptions of the loneliness and positive feelings factors. Third, the depressive mood factor was positively associated with teacher-reports of depressive mood and general anxiety, but not with their reports of social avoidance, whereas the happiness factor shared no associations with teachers' ratings. Fourth, none of the factors shared significant associations with any of the parental ratings, with the exception of happiness, which was unexpectedly associated with slightly higher levels of general anxiety rated by the parents.

Discussion

Our main objective was to validate the improved English and French versions of the Center for Epidemiologic Study Depression scale for youth with an Intellectual Disability, the CESD-ID-*R*. Our results showed that the 14 items from the CESD-ID-*R* followed a two-factor structure, matching the World Health Organization (WHO; 2014) representation of psychological health, and encompassing relatively orthogonal (i.e., uncorrelated) ratings of depressive mood and happiness.

The psychometric properties of this two-factor solution remained unchanged as a function of youth's sex, ID level, and comorbidities. Although there was some evidence of lack of measurement invariance between linguistic versions (English-Australia and French-Canada, see later discussion), these differences were not enough to call into question the fundamental equivalence of the English and French versions of CESD-ID-*R*. Finally, our results supported the concurrent validity of the CESD-ID-*R*, revealing associations between both factors and youth's self-reports on other measures of socio-emotional problems, although relations involving teachers and parental reports of socio-emotional difficulties were far more limited. This last observation reinforces that youth with ID can provide a reliable assessment of their internal state in a way that provides a non-redundant and complementary perspective to that obtained via informant reports.

Measurement Invariance

Tests of measurement invariance revealed that the CESD-ID-*R* functioned equally well irrespective of youth's sex, ID level, and comorbid conditions. Researchers can thus use this instrument to reliably assess and compare youth with ID irrespective of these personal characteristics. However, the reliability (uniquenesses) of youth's self-reports on three of the items associated with the happiness factor seemed to be slightly influenced by the linguistic version (English or French), or by the culture of the respective countries (Australia or Canada) where the study took place. More precisely, our results showed that the reliability of three items ("I have a lot of hopes for the future," "I can do a lot of things," "People are nice to me") was slightly higher in English (Australia) than in French (Canada), which might be explained by the higher complexity of the French language. Given that the lack of strict (i.e., uniquenesses) invariance was limited to a subset of items from the happiness factor, it is not sufficient to preclude the use of this CESD-ID-*R* factor for cross-cultural (or cross-linguistic) comparisons. However, these results suggest that researchers would do well to rely on latent variable methodologies to control for this source of lack of measurement invariance in their studies, and that practitioners should

use caution when using this specific factor to conduct comparisons across individuals from different linguistic or cultural backgrounds.

Convergent Validity

Results support the concurrent validity of the depressive mood and happiness factors, which both demonstrated associations with youth's characteristics and self-reports on indicators of socio-emotional problems. Matching previous results (Costello et al., 2011; Hsieh et al., 2020), girls with ID reported higher levels of depressive mood than that reported by boys with ID. Girls with ID also reported higher levels of happiness than their male counterparts. These results match those obtained among TD populations (Chaplin & Aldao, 2013), suggesting females' greater tendency to use both positive and negative emotional regulation strategies (Nolen-Hoeksema, 2012). Moreover, although this result might seem paradoxical, these two factors were essentially uncorrelated, meaning that they were relatively independent from one another, suggesting that youth with ID could thus feel both depressed and happy at the same time. Consequently, depressive mood and happiness seem to tap into distinct, complementary, and non-redundant aspects of youth with ID's emotional functioning (WHO, 2014) which is consistent with the results from previous validations of the CESD conducted among TD populations (Morin et al., 2011). Yet, this result is not consensual, as other studies have also found that the factor structure of the CESD assessed among TD youth could be best represented by depressive mood and happiness conceptualized as part of a same continuum (Siddaway et al., 2017; Wood et al., 2010). Such results are at odds with those obtained in the present study, suggesting that, even though these two states might be negatively correlated with one another among TD populations, they seem to be relatively independent from one another among youth with ID. These mixed results call for further investigation of how depression and happiness might or might not coexist in all youth, with or without ID. Besides, and pending replication, results from this study suggest that researchers seeking to understand youth with ID's internal states should consistently assess both aspects (depressive mood and happiness) separately, rather than on a single global CESD-ID-*R* score.

Although previous studies conducted among youth with ID were equivocal regarding the association between comorbid condition and depressive symptoms (Hsieh et al., 2020; Whitney et al., 2019), our results revealed no such association. Likewise, our results suggested a lack of association between youth's ID levels and their levels of depressive mood (Whitney et al., 2019). However, youth having a moderate compared to a mild level of ID reported slightly higher levels of happiness. This observation is consistent with studies suggesting that people with borderline levels of ID might be less happy than those with mild or moderate ID (Matikka & Ojanen, 2004).

Associations with youth's self-reports of socio-emotional problems also supported the convergent validity of the CESD-ID-*R.* As expected (Glenn et al., 2003; Klein et al., 2018; Masi et al., 2002), youth who reported higher levels of depressive mood on the CESD-ID-*R* also reported a higher level of depressive mood on the GDS-ID, higher levels of anxiety on the three factors of the GAS-ID, and higher levels of school loneliness. Likewise, youth who reported higher levels of happiness on the CESD-ID-*R* reported lower levels of depressive mood on the GDS-ID and lower levels of anxiety on the three factors of the GAS-ID. However, youth self-reports of happiness on the CESD-ID-*R* were not related to their self-reports of school loneliness. This result could be explained by the fact that, whereas people suffering from depression tend to avoid social interactions across contexts (Rojahn et al., 2011), happy people might fulfill their need for social connectedness across a variety of social contexts, in addition to school settings (Satici et al., 2016). More surprising was the lack of association between youth's self-reports of happiness on the CESD-ID-*R* and their self-reports of positive feelings on the GDS-ID, thus the convergent validity of this factor might be lower than that of the depressive mood factor. However, this lack of association could also be explained by the fact that the CESD-ID-*R* Happiness factor and the GDS-ID Positive feelings factor appear to tap into slightly different internal states. Thus, whereas the former seems to capture feelings of hope and joy about oneself (e.g., I have lots of hopes for the future; I can do lots of things; I feel that people like me), the latter seems to capture feelings of enjoyment about others and the world (e.g., I enjoy the things I do and have fun; I enjoy talking to people and being with other people; When something good happens to me, I feel happy). However, and pending replication, our results suggest that the Happiness factor of the CESD-ID-*R* and the Positive feelings factor of the GDS-ID seem to reflect distinct, and possibly complementary, facets of youth's positive internal states. Overall, youth's self-reports of socio-emotional problems were more strongly related to their level of depressive mood (r=.218 to .543) than to their level of happiness (r=-

.165 to -.254), which could reflect the tendency of depression, anxiety, and loneliness symptoms to aggregate (Caspi et al., 2014).

Consistent with youth being best placed to assess their own internal states (Milevicuite & Hartley, 2015; Smith, 2007; Turk et al., 2012) and with previous recommendations highlighting the importance of providing them with a way to express their voice in research focusing on their internal reality (Hartley & MacLean, 2006), our results suggested that youth's self-reports on the CESD-ID-*R* were mainly independent of teachers' and parents' reports of their socio-emotional difficulties. Thus, we found some congruence between youth's ratings of depressive mood (but not happiness) on the CESD-ID-*R* and their teachers' ratings of depressive mood and general anxiety. However, our results also highlight a lack of associations between youth's CESD-ID-*R* ratings and parental reports of socio-emotional difficulties, with the exception of parental reports of general anxiety, which were paradoxically associated with higher self-reported levels of happiness. These results reinforce that teachers might more accurately identify students at risk for internalizing problems than parents (Dwyer et al., 2006). Teachers' experience in working with a variety of at-risk youth might make them more sensitive to subtle variations in mood otherwise missed by parents, especially when asked to rate internal manifestations rather than observable behaviors (Kemper et al., 2003). This result is also consistent with youth's tendency to disclose less information about themselves to their parents during adolescence (Keijsers & Poulin, 2013). Furthermore, a lack of agreement between youth, parents, and teachers in the assessment of depressive symptoms is a commonly observed phenomenon, even in research conducted among non-clinical samples of TD youth (Salbach-Andrea et al., 2008; Youngstrom et al., 2000). Among these youth, this lack of congruence is generally explained by informants' limited abilities to assess the emotional states of someone else (Smith, 2007), an interpretation that can easily be applied to youth with ID. Still, properly evaluating if these discrepancies reflect youth's more accurate evaluation of their internal states or a potential lack of self-awareness could be assessed by comparing CESD-ID-*R* scores to structured diagnostic interviews, which would be an interesting question for future research. However, for the moment, the presence of convergence between youth self-reports of different types of difficulties seems consistent with the idea that their self-reports of depression are meaningful, at least from their perspective.

Limitations

First, although we found that our results among a sample of youth with ID were similar to results obtained among samples of TD youth, no comparison sample of TD youth was considered, thus making any claim of generalizability or difference tentative at best. Second, the study relied on youth from two countries sharing a very similar culture and is thus unable to account for possible cultural differences. The study still assessed the equivalence of linguistic (English and French) versions of the CESD-ID-*R*. Thus, the fact that these two versions were used in different countries (Australia and Canada) makes it impossible to disentangle differences due to language or to slight culture variations. Furthermore, although our sample size is relatively large compared to other studies of youth with ID, it remains at the lower bound when using Chen's (2007) guidelines for tests of measurement invariance. Future work is thus needed to verify the generalizability of our findings to larger samples of youth with ID from a greater variety of countries and cultures. Third, the study remains cross-sectional in nature and thus unable to inform questions related to the directionality of the observed associations and test-retest reliability and stability. To better understand how depression is predicted by and predicts socio-emotional outcomes, future studies should rely on fully longitudinal designs, making it possible to explicitly consider change over time. Longitudinal studies are also needed to assess the test-retest reliability and longitudinal stability of youth's ratings on the CESD-ID-*R*. Finally, despite the promising conclusions from this study, caution is required when using youth's self-reports on the happiness factor. Indeed, when compared to the depressive mood factor, the happiness factor only showed moderate associations with the outcomes. Further investigations are thus needed to better document the convergent validity (or lack thereof) of ratings on this factor.

Conclusion

Our results are encouraging regarding the ability of the CESD-ID-*R* to accurately capture depressive mood and happiness among youth with mild to moderate levels of ID in the context of epidemiological studies conducted among English- and French-speaking boys and girls irrespective of comorbid conditions. The high rates of depression observed among youth with ID are currently accompanied by a limited number of self-report instruments validated for this population in English or

French. The CESD-ID-*R* provides a valuable epidemiological tool for researchers and practitioners to identify youth with ID suffering from depression and evaluate the efficacy of interventions. Important next steps in the validation of this instrument would involve verifying whether: (a) the CESD-ID-*R* is able to provide comparable scores among populations of youth with ID and TD youth; (b) scores obtained on the CESD-ID-*R* are comparable to those obtained on the CESD among populations of TD youth. To conclude, the CESD-ID-*R* empowers youth with ID by providing them with an avenue for self-expression, voice, and agency rather than merely capturing informant reports of their internal states. The consequences of depression among youth call for more studies assessing the concurrent associations between CESD-ID-*R* scores and a more extensive set of outcomes. Such studies will help further understand how youth's perceptions of their internal states affect their day-to-day functioning and adjustment.

References

American Psychological Association [APA]. (2000). *Diagnostic and Statistical Manual, Fourth Edition* (DSM-IV). American Psychological Association.

American Psychological Association [APA]. *APA Dictionary of Psychology*. American Psychological Association.

Asher, S.R., Hymel, S., & Renshaw, P.D. (1984). Loneliness in children. *Child Development, 55*, 1456-1464.

Asparouhov, T., & Muthén, B.O. (2010). *Weighted least square estimation with missing data*. http://www.statmodel.com/download/GstrucMissingRevision.pdf

Austin, K.L., Hunter, M., Gallagher, E., & Campbell, L.E. (2018). Depression and anxiety symptoms during the transition to early adulthood for people with intellectual disabilities. *Journal of Intellectual Disability Research, 62*, 407-721.

Berg, K.L., Shiu, C.S., Msall, M.E., & Acharya, K. (2015). Victimization and depression among youth with disabilities in the US child welfare system. *Child: Care, Health and Development, 41*, 989-999.

Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley.

Caspi, A., Houts, R.M., Belsky, D.W., Goldman-Mellor, S.J., Harrington, H., Israel, S., Meirer, M.H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T.E. (2014). The p Factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science, 2*, 119-137.

Chaplin, T.M. & Aldao, A. (2013). Gender differences in emotion expression in children: A meta-analytic review. *Psychological Bulletin, 139*, 735-765.

Charlot, L., Deutsch, C., Hunt, A., Fletcher, K., & McIlvane, W. (2007). Validation of the Mood and Anxiety Semi-structured (MASS) Interview for patients with intellectual disabilities. *Journal of Intellectual Disability Research, 51*, 821-834.

Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504.

Chester, R., Chaplin, E., Tsakanikos, E., McCarthy, J., Bouras, N., & Craig, T. (2013). Gender differences in self-reported symptoms of depression and anxiety in adults with intellectual disabilities. *Advances in Mental Health and Intellectual Disabilities, 7*, 191-200.

Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Costello, E.J., Copeland, W., & Angold, A. (2017). Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults? *The Journal of Child Psychology and Psychiatry, 52*, 1015-1025.

Cuthill, F.M., Espie, C.A., & Cooper, S.A. (2003). Development and psychometric properties of the Glasgow Depression Scale for people with a learning disability. *British Journal of Psychiatry, 182*, 347-353.

Dagnan, D., Jahoda, A., McDowell, K., Masson, J., Banks, P., & Hare, D. (2008). The psychometric properties of the hospital anxiety and depression scale adapted for use with people with intellectual disabilities. *Journal of Intellectual Disability Research, 52*, 942-949.

Dovgan, K., Mazurek, M.O., & Hansen, J. (2019). Measurement invariance of the child behavior checklist in children with autism spectrum disorder with and without intellectual disability: Follow-up study. *Research in Autism Spectrum Disorder, 58*, 19-29.

Dwyer, S.B., Nicholson, J.M., & Battistuta, D. (2006). Parent and teacher identification of children at risk of developing internalizing or externalizing mental health problems: A comparison of screening methods. *Prevention Science*, *7*, 343-357.

Enders, C.K. (2010). *Applied missing data analysis*. Guildford.

Esbensen, A.J., Rojahn, J., Aman, M.G., Ruedrich, S. (2003). Reliability and validity of an assessment instrument for anxiety, depression, and mood among individuals with mental retardation. *Journal of Autism and Developmental Disorders*, *33*, 617-629.

Esbensen, A.J., Seltzer, M.M., Greenberg, J.S., & Benson, B.A. (2005). Psychometric evaluation of a self-report measure of depression for individuals with mental retardation. *American Journal of Mental Retardation*, *110*, 469-481.

Finney, S.J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G.R. Hancock & R.O. Mueller (Eds.), Quantitative methods in education and the behavioral sciences: Issues, research, and teaching. *Structural equation modeling: A second course* (pp. 439-492). IAP Information Age Publishing.

Garber, J., & Rao, U. (2014). Depression in children and adolescents. In M. Lewis & K.D. Rudolph (Eds.), *Handbook of Developmental Psychopathology, Third Edition* (pp. 489-520). Springer.

GBD 2019 Diseases and Injuries collaborators (2020). Global Burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, *396*, 1204-1222.

Glenn, E., Bihm, E.M., & Lammers, W.J. (2003). Depression, anxiety, and relevant cognitions in persons with mental retardation. *Journal of Autism and Developmental Disorders*, *33*, 69-76.

Hamers, P.C.M., van Ool, J.S., Festen, D.A.M., Hendriksen, J.G.M., Bindels, P.J.E., & Hermans, H. (2018). Reliability and validity of the Dutch Anxiety, Depression and Mood Scale in adult aged <50 years with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, *3*, 568-574.

Hartley, S.L., & MacLean, W.E. (2006). A review of the reliability and validity of Likert scales for people with intellectual disability. *Journal of Intellectual Disability Research*, *50*, 813–827.

Hermans, H. & Evenhuis, H.M. (2010). Characteristics of instruments screening for depression in adults with intellectual disabilities: Systematic review. *Research in Developmental Disabilities, 31*, 1109-1120.

Hermans, H., Jelluma, N., van der Pas, F.H., & Evenhuis, H.M. (2012). Feasibility, reliability and validity of the Dutch translation of the Anxiety, Depression and Mood Scale in older adults with intellectual disability. *Research in Developmental Disabilities*, *33*, 315-323.

Hermans, H., Wieland, J., Jelluma, N., van der Pas, F.H., & Evenhuis, H.M. (2013). Reliability and validity of the Dutch version of the Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID). *Journal of Intellectual Disability Research*, *57*, 728-736.

Hsieh, K., Scott, H.M., & Murthy, S. (2020). Associated risk factors for depression and anxiety in adults with intellectual and developmental disabilities: Five-year follow up. *American Journal on Intellectual and Developmental Disabilities*, *125*, 49-63.

Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Institut de la statistique du Québec (ISQ; 2018). *Enquête québécoise sur la santé mentale des jeunes du secondaire 2016-2017. Résultats de la deuxième édition. L'adaptation sociale et la santé mentale des jeunes, Tome 2.* www.stat.gouv.qc.ca.

Keijsers, L., & Poulin, F. (2013). Developmental changes in parent–child communication throughout adolescence. *Developmental Psychology*, *49*, 2301–2308

Kemper, T.S., Nader, R.G., Repper, K.K., & Kistner, J.A. (2003). Mother-child agreement on reports of internalizing symptoms among children referred for evaluation of ADHD. *Journal of Psychopathology and Behavioral Assessment, 25*, 239-250.

Klein, A.M., Houtkamp, E.O., Salemink, E., Baartmans, J.M.D., Rinck, M., & van der Molen, M.J. (2018). Differences between self- and peer-rated likability in relation to social anxiety and depression in adolescents with mild intellectual disabilities. *Research in Developmental Disabilities, 80*, 44-51.

Lindsay, W.R. & Skene, D.D. (2007). The Beck Depression Inventory II and the Beck Anxiety Inventory in people with intellectual disabilities: Factor analyses and group data. *Journal of*

Applied Research in Intellectual Disabilities, *20*, 401-408.
Little, T.D. (2013). *Longitudinal Structural Equation Modeling*, The Guildford Press.
Maïano, C., Coutu, S., Tracey, D., Bouchard, S., Lepage, G., Morin, A.J.S., Moullec, G. (2018). Prevalence of anxiety and depressive disorders among youth with intellectual disabilities: A systematic review and meta-analysis. *Journal of Affective Disorders*, *236*, 230-242.
Maïano, C., Morin, A.J.S., & Bégarie, J. (2011). The Center for Epidemiologic Studies Depression Scale: Factor validity and reliability in a French sample of adolescents with intellectual disability. *Research in Developmental Disabilities*, *32*, 1872-1883.
Marsh, H.W., Scalas, L.F., & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg self-esteem scale: Traits, ephemeral artifacts, and stable response styles. *Psychological Assessment*, *22*, 366-381.
Marsh, H.W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McArdle (Eds.), *Contemporary Psychometric* (pp.275-340). Erlbaum.
Martin, L., Hirdes, J.P., Fries, B.E., & Smith, T.F. (2007). Development and psychometric properties of an assessment for persons with intellectual disability – The interRAI ID. *Journal of Policy and Practice in Intellectual Disabilities*, *4*, 23-29.
Masi, G., Brovedani, P., Mucci, M., & Favilla, L. (2002). Assessment of anxiety and depression in adolescents with mental retardation. *Child Psychiatry and Human Development*, *32*, 227-237.
Matikka, L.M., & Ojanen, M. (2004). Happiness in persons with intellectual disabilities. *European Journal of Mental Disability*, *29*, 3-18.
Matson, J.L., Baglio, C.S., Smiroldo, B.B., Hamilton, M., Packlowskyi, T., Williams, D., & Kirkpatrick-Sanchez, S. (1996). Characteristics of autism as assessed by the Diagnostic Assessment for the Severely Handicapped-II (DASH-II). *Research in Developmental Disabilities*, *17*, 135-143.
Matson, J.L., & Bamburg, J.W. (1998). Reliability of the Assessment of Dual Diagnosis (ADD). *Research in Developmental Disabilities*, *19*, 89–95.
McDonald, R.P. (1970). The theoretical foundations of principal factor analysis and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, *23*, 1-21.
Meins, W. (1993). Assessment of depression in mentally retarded adults: Reliability and validity of the Children's Depression Inventory (CDI). *Research in Developmental Disabilities*, *14*, 299–312.
Meins, W. (1996). A new depression scale designed for use with adults with mental retardation. *Journal of Intellectual Disability Research*, *40*, 222–226.
Mileviciute, I., & Hartley, S.L. (2013). Self-reported versus informant-reported depressive symptoms in adults with mild intellectual disability. *Journal of Intellectual Disability Research*, *59*, 158-169.
Millsap, R. (2011). *Statistical Approaches to Measurement Invariance*. Taylor & Francis.
Mindham, J. & Espie, C.A. (2003). Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID): Development and psychometric properties of a new measure for use with people with mild intellectual disability. *Journal of Intellectual Disability Research*, *47*, 22-30.
Morin, A.J.S., Moullec, G., Maïano, C., Layet, L., Just, J.L., & Ninot, G. (2011). Psychometric properties of the Center for Epidemiologic Studies Depression Scale (CES-D) in French clinical and nonclinical adults. *Epidemiology and Public Health*, *59*, 327-340.
Muthén, L., & Muthén, B. (2019). *Mplus User's Guide*. Muthén & Muthén.
Nolen-Hoeksema, S. (2012). Emotion regulation and psychopathology: The role of gender. *Annual Review of Clinical Psychology*, *8*, 161-187.
Olivier, E., Azarnia, P., Morin, A.J.S., Houle, S.A., Dubé, C. Tracey, D., & Maïano, C. (2020). The moderating role of teacher-student relationships on the association between peer victimization and depression in students with intellectual disabilities. *Research in Developmental Disabilities*, *98*, 1-11.
Perez-Achiaga, N., Nelson, S., & Hassiotis, A. (2009). Instruments for the detection of depressive symptoms in people with intellectual disabilities. *Journal of Intellectual Disabilities*, *13*, 55-76.
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.
Prosser, H., Moss, S., Costello, H., Simpson, N., Patel, P., & Rowe, S. (1998). Reliability and validity

of the Mini PAS-ADD for assessing psychiatric disorders in adults with intellectual disability. *Journal of Intellectual Disability Research, 42*, 264–272.

Powell, R. (2003). Psychometric properties of the Beck Depression Inventory and the Zung Self-Rating Depression Scale in adults with mental retardation. *Mental Retardation, 41*, 88-95.

Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385-401.

Reynolds, W.M. & Baker, J.A. (1988) Assessment of depression in persons with mental retardation. *American Journal on Mental Retardation*, *93*, 93-103.

Roid, G.H., & Miller, L.J. (1997). *Leiter International Performance Scale-Revised*. Stoelting.

Rojahn, J., Rowe, E.W., Kasdan, S., Moore, L., & van Ingen, D.J. (2011). Psychometric properties of the Aberrant Behavior Checklist, the Anxiety, Depression and Mood Scale, the Assessment of Dual Diagnosis and the Social Performance Survey Schedule in adults with intellectual disabilities. *Research in Developmental Disabilities*, *32*, 2309-2320.

Ross, E., & Oliver, C. (2003). Preliminary analysis of the psychometric properties of the Mood, Interest & Pleasure Questionnaire (MIPQ) for adults with severe and profound learning disabilities. *British Journal of Clinical Psychology*, *42*, 81–93.

Salbach-Andrea, H., Lenz, K., & Lehmkuhl, U. (2008). Patterns of agreement among parent, teacher and youth ratings in a referred sample. *European Psychiatry*, *24*, 345-351.

Santor, D.A., Gregus, M., & Welch, A. (2006). Focus article: Eight decades of measurement in depression. *Measurement: Interdisciplinary Research and Perspectives*, *4*, 135-155.

Satici, S.A., Uysakm R., & Deniz, M.E. (2016). Linking social connectedness to loneliness: The mediating role of subjective happiness. *Personality and Individual Differences*, *97*, 306-310.

Siddaway, A.P., Wood, A.M., & Taylor, P.J. (2017). The Center for Epidemiologic Studies-Depression (CES-D) scale measures a continuum from well-being to depression: Testing two key predictions of positive clinical psychology. *Journal of Affective Disorders*, *213*, 180-186.

Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, *66*, 563-576

Smith, S.R. (2007). Making sense of multiple informants in child and adolescent psychopathology, a guide for clinicians. *Journal of Psychoeducational Assessment*, *25*, 139-149.

Sturmey, P., Newton, J.T., Cowley, A., Bouras, N., & Holt, G. (2005). The PAS-ADD Checklist: Independent replication of its psychometric properties in a community sample. *British Journal of Psychiatry, 186,* 319–323.

Turk, V., Khattran, S., Kerry, S., Corney, R., & Painter, K. (2012). Reporting of health problems and pain by adults with an intellectual disability and by their carers. *Journal of Applied Research in Intellectual Disabilities*, *25*, 155-165.

Watson, J.E., Aman, M.G., & Singh, N.N. (1988). The psychopathology instrument for mentally retarded adults: Psychometric characteristics, factor structure, and relationship to subject characteristics. *Research in Developmental Disabilities*, *9*, 277-290.

Weinberger, A.H., Gbedemah, M., Martinez, A.M., Nash, D., Galea, S., & Goodwin, R.D. (2018). Trends in depression prevalence in the USA from 2005 to 2015: Widening disparities in vulnerable groups. *Psychological Medicine, 48,* 1308–1315.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth Edition*. Pearson.

Wechsler, D., Coalson, D.L., & Raiford, S E. (2008). *WAIS-IV technical and interpretive manual*. Pearson.

Whitney, D.G., Shapiro, D.N., Paterson, M.D., & Warschausky, S.A. (2019). Factors associated with depression and anxiety in children with intellectual disabilities. *Journal of Intellectual Disability Research*, *63*, 408-417.

World Health Organization (2014). *Mental Health.* www.who.int/features/factfiles/mental_health/en/.

Wood, A.M., Taylor, P.J., & Joseph, S. (2010). Does the CES-D measure a continuum from depression to happiness? Comparing substantive and artifactual models. *Psychiatry Research*, *177*, 120-123.

Youngstrom, E., Loever, R., & Stouthamer-Loeber, M. (2000). Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of Consulting and Clinical Psychology*, *68*, 1038-1050.

Yu, C.Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*. University of California.

Table 1

Review of Existing Depression Scales Validated among Individuals with ID

Instrument	Validation studies for a population with ID	Rated by	Setting	Validated for (age group)	Length	Available for TD
Anxiety, Depression, and Mood Scale (ADAMS)	Esbensen et al. (2003) Hamers et al. (2018) Hermans et al. (2012) Rojahn et al. (2011)	Professional Caregiver-report	Residential Treatment / clinical setting	Adolescents Adults	7 items	no
Beck Depression Inventory (BDI)	Lindsay & Skene (2007) Powell (2003)	Self-report	Residential Treatment / clinical setting	Adults	21 items	yes
Center for Epidemiologic Studies Depression Scale (CESD)	Maïano et al. (2011)	Self-report	Various / Community	Adolescents	14 items	yes
Child Behavior Checklist (CBCL) – Anxiety/Depressed subscale	Dovgan et al. (2019)	Parent-report	Various / Community	Children Adolescents	13 items	yes
Comprehensive Psychopathological Rating Scale – Depression Scale	Meins (1996)	Researcher	Treatment / clinical setting	Adults	9 items	yes
Glasgow Depression Scale for People with Learning Disabilities (GDS-ID)	Cuthill et al. (2003) Hermans et al. (2013)	Self-report	Various / Community	Adults	20 items	no
Mood, Interest & Pleasure Questionnaire (MIPQ)	Ross & Oliver (2003)	Informant-report	Various / Community	Adults	25 items	no
Reiss Screen for Maladaptive Behaviour (RSMB)	Havercamp & Reiss (1997)	Caregiver-rated	Treatment / clinical setting	Adults	10 items	no
Self-Reported Depression Questionnaire (SRDQ)	Reynolds & Baker (1988) Esbensen et al. (2005)	Self-report	Residential Treatment / clinical setting	Adolescents Adults	32 items	no
Zung Self-Rating Depression Scale	Powell (2003)	Self-report	Residential Treatment / clinical setting	Adults	20 items	yes

Note. Validation studies included in this table were all published in English, although they could refer to measures developed in other languages. This table excludes interview protocols or checklists, instruments exclusively designed for service providers, and older instruments which have not been re-validated since the publication of the DSM-IV and DSM-V. As such, the following instruments are not included in the table: Assessment of Dual Diagnosis (ADD; e.g., Matson & Bamburn, 1998); Diagnostic Assessment for the Severely Handicapped-II (e.g., Matson et al., 1996); Hospital Anxiety and Depression Scale for People with ID (e.g., Dagnan et al., 2008); InterRAI-Intellectual Disability (e.g., Martin et al., 2007); Mini Psychiatric Assessment Schedule for Adults with Developmental Disabilities (Mini-PAS-ADD; e.g., Posser et al., 1998); Mood and Anxiety Semi-structured (MASS) Interview (e.g., Charlot et al., 2007); Psychiatric Assessment Schedule for Adults with Developmental Disabilities (PAS-ADD) Checklist (e.g., Sturmey et al., 2005); Psychopathology Instrument for Mentally Retarded Adults (e.g., Watson et al., 1988). Although it has been used with adults with ID (e.g., Meins, 1993), the Child Depression Inventory is also excluded from this Table as we found no record of studies having validated this instrument among samples of youth with ID. For a more comprehensive review of existing scales, please see Hermans & Evenhuis (2010) and Perez-Achiaga et al. (2009).

Table 2

Descriptive statistics per country.

	Canada	Australia
Sex		
Male	48.7%	67.1%
Female	51.3%	32.9%
ID level		
Mild	33.0%	59.3%
Moderate	67.0%	40.7%
Comorbid condition		
None	55.9%	51.6%
Autism Spectrum	12.2%	34.0%
Genetic syndrome	31.9%	11.3%
Both	0.0%	3.1%
School type		
Special school	31.0%	0.0%
Special classroom in regular school	69.0%	92.6%
Regular classroom in regular school	0.0%	7.4%
Recruitment setting		
School	91.5%	100.0%
Community	8.5%	0.0%
Immigration status		
1st or 2nd generation immigrant[1]	17.0%	41.5%
3rd + generation immigrant	83.0%	58.5%
Parents' highest education level		
No schooling	3.2%	1.2%
Primary schooling	1.1%	1.2%
Secondary schooling	54.8%	49.4%
Vocational training	4.3%	23.5%
University (bachelor and up)	36.6%	24.7%

1: At least one parent born abroad.

Table 3

Model Fit of Alternative Measurement Models and Measurement Invariance Models.

Model	χ^2	*df*	CFI	TLI	RMSEA	RMSEA 90% CI	$\Delta\chi^2$	Δdf	ΔCFI	ΔTLI	ΔRMSEA
Main Measurement Models											
1. One-factor CFA	187.897*	62	.976	.964	.077	.064-.089					
2. Four-factor CFA	192.093*	66	.975	.966	.074	.062-.087	4.196	4	+.001	+.002	-.003
3. Two-factor CFA	213.680*	76	.973	.968	.072	.061-.084	21.587*	10	-.002	+.002	-.002
Measurement Invariance (Two-factor CFA): Sex											
1. Configural invariance	292.790*	152	.970	.964	.073	.060-.086	–	–	–	–	–
2. Weak invariance	301.808*	164	.970	.967	.070	.057-.082	9.582	12	.000	+.003	-.003
3. Strong invariance	310.365*	183	.973	.973	.063	.051-.075	13.836	19	+.003	+.006	-.007
4. Strict invariance	335.606*	197	.970	.972	.064	.052-.075	33.300*	14	-.003	-.001	+.001
Measurement Invariance (Two-factor CFA): ID Level											
1. Configural invariance	270.363*	152	.975	.970	.070	.056-.083	–	–	–	–	–
2. Weak invariance	284.966*	164	.974	.972	.068	.055-.081	16.025	12	-.001	+.002	-.002
3. Strong invariance	300.540*	183	.975	.975	.064	.050-.076	22.231	19	+.001	+.003	-.004
4. Strict invariance	346.458*	197	.968	.971	.069	.057-.081	56.356*	14	-.007	-.004	+.005
Measurement Invariance (Two-factor CFA): Comorbidities											
1. Configural invariance	288.754*	152	.952	.942	.098	.081-.115	–	–	–	–	–
2. Weak invariance	299.438*	164	.952	.947	.094	.077-.110	12.325	12	.000	+.005	-.004
3. Strong invariance	317.744*	183	.953	.953	.089	.072-.105	24.743	19	+.001	+.006	-.005
4. Strict invariance	338.233*	197	.950	.954	.087	.071-.103	28.503*	14	-.003	+.001	-.002
Measurement Invariance (Two-factor CFA): Country											
1. Configural invariance	322.078*	152	.949	.938	.080	.068-.093	–	–	–	–	–
2. Weak invariance	309.608*	164	.956	.951	.072	.059-.084	103.088*	12	+.007	+.013	-.008
3. Strong invariance	354.214*	183	.948	.949	.074	.062-.085	56.571	19	-.008	-.002	+.002
4. Strict invariance	470.734*	197	.917	.924	.090	.079-.100	7.721*	14	-.031	-.025	+.016
5. Partial Strict invariance	391.151*	193	.940	.944	.077	.066-.088	43.331*	10	-.008	-.005	+.003

Note. **p* < .01; χ^2: Chi square test of model fit and associated degrees of freedom (*df*) reported for descriptive purposes; CFI: Comparative Fit Index; TLI: Tucker–Lewis Index; RMSEA: Root Mean Square Error of Approximation and 90% Confidence Interval (CI); (Excellent fit: RMSEA≤.06 and CFI-TLI ≥.95; Adequate fit: RMSEA≤.08 and CFI/TLI ≥.90) ; Δ: Change according to the previous retained model (Non-invariance: RMSEA increases >.015 and CFI-TLI decreases >.010); $\Delta\chi^2$: Chi square difference test calculated using the Mplus DIFFTEST option (Muthén & Muthén, 2019) reported for descriptive purposes.

Table 4

Standardized Factor Loadings (λ) and Uniquenesses (δ) from the Two-factor Measurement Model

		Dep. Mood λ	Happiness λ	δ
1	I am bothered by things that don't usually bother me	.545		.703
2	I think about bad things	.759		.424
3	I have a hard time keeping my mind on what I am doing	.695		.517
4	I feel depressed	.777		.397
5	I have lot of hopes for the future		.488	.762
6	I feel that my life is a failure	.801		.358
7	I sleep badly	.682		.535
8	I am happy		.875	.235
9	I can do a lot of things		.657	.569
10	I feel lonely	.734		.461
11	People are nice to me		.818	.331
12	I enjoy life		.900	.190
13	I feel sad	.780		.391
14	I feel that people like me		.844	.288
Composite reliability (ω)		.898	.898	
Scale Score Reliability (α)		.848	.858	

Note. All results are significant at $p \leq .05$; ω: omega coefficient of composite reliability (McDonald, 1970); α: Cronbach alpha coefficient of scale score reliability.

Table 5

Convergent Validity (Pearson Correlations)

	Depressive Mood *r* [95% CI]	Happiness *r* [95% CI]
Youth Characteristics		
Sex (0 = female; 1 = male)	-.18 [-.14; -.03]**	-.21 [-.16; .06]**
ID Level (0 = mild)	.07 [-.03; .10]	.12 [.00; .11]*
Comorbidities (0 = none)	.06 [-.04; .10]	-.03 [-.09; .06]
Youth-rated Outcomes		
Depression-Depressive Mood (GDS-ID)	.54 [.44; .62]**	-.24 [-.34; -.13]**
Depression-Positive Feelings (GDS-ID)	.04 [-.06; .13]	-.10 [-.18; .01]
Anxiety-Worries (GAS-ID)	.44 [.33; .52]**	-.25 [-.35; -.14]**
Anxiety-Fears (GAS-ID)	.46 [.34; .53]**	-.17 [-.26; -.05]**
Anxiety-Physiological (GAS-ID)	.52 [.41; .59]**	-.18 [-.28; -.06]**
Loneliness	.22 [.09; .29]**	-.03 [-.11; .06]
Teacher-rated Outcomes		
Depressive Mood (ADAMS)	.26 [.09; .30]**	.06 [-.06; .14]
General Anxiety (ADAMS)	.25 [.09; .32]**	.07 [-.05; .16]
Social Avoidance (ADAMS)	.13 [-.01; .23]	-.11 [-.19; .02]
Parent-rated Outcomes		
Depressive Mood (ADAMS)	-.05 [-.16; .09]	.10 [-.04; .20]
General Anxiety (ADAMS)	-.07 [-.19; .08]	.16 [.02; .26]*
Social Avoidance (ADAMS)	-.09 [-.21; .06]	-.04 [-.17; .10]

Note. * $p \leq .01$; ** $p \leq .05$. *r*: Pearson's correlation. CI: Confidence intervals; GDS-ID: Glasgow Depression Scale for people with Intellectual Disabilities; GAS-ID: Glasgow Anxiety Scale for people with Intellectual Disabilities; ADAMS: Anxiety, Depression, and Mood Screen.

Online Supplements for Validation of a Revised Version of the Center for Epidemiologic Studies Depression Scale for Youth with Intellectual Disabilities (CESD-ID-*R*)

Appendix A
Scale Development and Pilot Testing

Objectives

Starting from the original French version of the CESD-ID developed by Maïano et al. (2011), the first objective of these pilot studies was to develop an equivalent English version of this instrument. As part of these procedures, we also implemented further adaptations to the item and response format to improve the clarity and ease of application of this instrument, based on recommendations related to the use of self-report questionnaires among people with ID (Finlay & Lyons, 2001, 2002). This preliminary adaptation of the CESD-ID-*R* was then tested among a first sample of youth with ID, which led to further adaptations. The final adaptation was then tested again among a second sample of youth with ID.

Method

Participants and Procedures. The pilot sample comprised 34 youth (aged between 13 to 21 years; 35% girls) with mild to moderate-severe ID, including 20 English-speaking Australians and 14 French-speaking Canadians. A first subsample of 18 youth (N = 10 in Australia and 8 in Canada) was solicited to evaluate the format and clarity of a preliminary version of the CESD-ID-*R*. A second subsample of 16 youth (N = 10 in Australia and 6 in Canada) was solicited to assess the format and clarify of the final version of the CESD-ID-*R*. The procedures used in this pilot study were identical to those used in the main study and received approval from the same research ethics committees. However, in this pilot process, the CESD-ID-*R* was administered individually, at school, by trained research assistants using a read-aloud assisted procedure to maximise youth's understanding and to facilitate discussion. This administration was mainly focused on assessing the level of understanding of the youth and the ease with which they could respond to the items.

Measures. A preliminary assessment of the appropriateness of the format and clarity of the items was conducted by all members of the research team familiar with the use of self-report questionnaires among youth with ID. This preliminary assessment confirmed that the items, already adapted by Maïano et al. (2011) for youth with ID, were appropriate for use as a self-reported questionnaire. This initial version was first adapted to English by two bilingual members of the research team using a translation back-translation procedure. Then, this preliminary English version was back translated into French by two other bilingual members of the research team and compared with the original French version. Discrepancies were resolved by adapting the English items. During this process, decisions were taken and discussed by the research team members in committee until a consensus was reached. To further increase youth's understanding of the sentences, words from the items were also associated with pictograms (presented above the words). Additionally, a "do not understand the statement" option was added to the response scale for situations in which respondents remained unable to understand the item. This whole process was conducted in collaboration with school personnel (i.e., teachers, psychologists, and psycho-educators) familiar with youth with ID.

Results

The responses provided by the first subsample of youth revealed that some words used in some of the items lack of precision or remained hard to understand for youth with ID (more specifically by those with more severe levels of ID). These results also revealed that the response scale was easy to understand for all participants, but that some participants did not use the exact wording of the verbal anchors of this response scale when answering items, preferring to use a simpler "yes" or "no". Therefore, the problematic words were replaced by simpler words preserving the same meaning, and the verbal anchors of the response scale were revised to also include "No" and "Yes" above the graphical faces depicting the response scales, corresponding to the following response categories: "*No, I* totally disagree", "*No, I* disagree", "*Yes, I* agree", and "*Yes, I* totally agree". Finally, a template comprising graphical displays and pictograms was developed to explain to the youth how to use the answer scale. This final version of the CESD-ID-*R* was administered to the second subsample of youth. Results supported the adequacy of the final English and French adapted versions of the CESD-ID-*R* and proved their suitability for use as self-report instruments among youth with ID.

References

Finlay, W.M.L., & Lyons, E. (2001). Methodological issues in interviewing and using self-report questionnaires with people with mental retardation. *Psychological Assessment, 13*, 319–335.

Finlay, W.M.L., & Lyons, E. (2002). Acquiescence in interviews with people with mental retardation. *Mental Retardation*, *40*, 14–29.

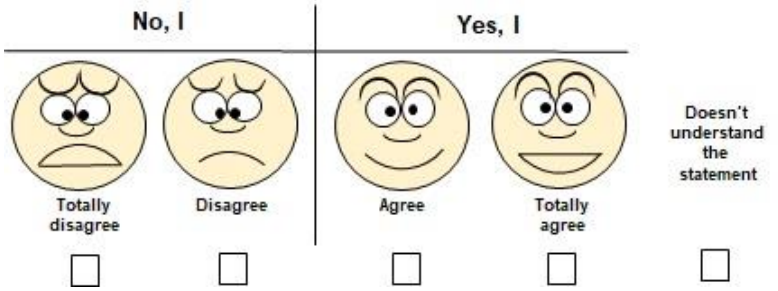
Appendix B

Complete List of Items for Center for Epidemiologic Studies Depression Scale for Youth with Intellectual Disabilities, Revised Version (CESD-ID-R)

	English Version	*French Version*
1	I am bothered by things that don't usually bother me	Je suis embêté(e) par des choses qui d'habitude ne m'embêtent pas
2	I think about bad things	Je pense à de mauvaises choses
3	I have a hard time keeping my mind on what I am doing (*I have a hard time keeping my mind on what I do*)	J'ai de la difficulté è me concentrer sur ce que je fais (*J'ai des difficultés à faire attention à ce que je fais*)
4	I feel depressed	Je me sens déprimé(e)
5	I have lot of hopes for the future (*I feel hopeful about the future*)	J'ai plein d'espoir dans l'avenir
6	I feel that my life is a failure	Je sens que ma vie est ratée.
7	I sleep badly	Je dors mal
8	I am happy	Je suis heureux(se)
9	I can do a lot of things	J'arrive à faire plein de choses
10	I feel lonely	Je me sens seul(e)
11	People are nice to me (*People are nice*)	Les gens sont gentils avec moi (*Les gens sont gentils*)
12	I enjoy life	J'ai du plaisir dans la vie (*Je profite de la vie*)
13	I feel sad	Je me sens triste
14	I feel that people like me	Je sens que les gens m'aiment bien
	No, I: 1. Strongly disagree 2. Disagree Yes, I: 3. Agree 4. Totally agree Doesn't understand the statement	Non, je suis : 1. Pas du tout d'accord 2. Pas d'accord Oui je suis : 3. D'accord 4. Tout à fait d'accord Ne comprend pas l'énoncé

Note. Items from the original CESD-ID that were modified in the CESD-ID-R are in parenthesis.

Graphical Response Scale: English



Graphical Response Scale: French

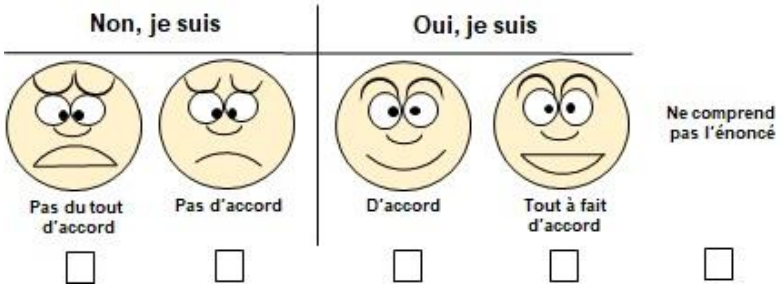


Table S1

Measurement Model for the Outcomes self-reported by the Youth

	Depressive Mood (GDS-ID)		*Positive Feelings (GDS-ID)*		*Worries (GAS-ID)*		*Fears (GAS-ID)*		*Psychosomatic (*
Items	λ	δ	λ	δ	λ	δ	λ	δ	λ
1	.850	.278	.785	.384	.706	.501	.665	.558	.706
2	.764	.416	.772	.404	.545	.703	.676	.543	.758
3	.547	.701	.754	.431	.612	.625	.699	.511	.771
4	.685	.531	.525	.725	.691	.522	.569	.677	.692
5	.780	.392	.623	.611	.828	.315	.616	.620	.754
6	.668	.554			.802	.356	.678	.541	.680
7	.701	.509			.707	.499	.803	.356	.812
8	.611	.626			.739	.454	.824	.321	.825
9	.111	.988			.461	.788	.834	.305	
10	.633	.599			.687	.528			
11	.704	.504							
12	.770	.407							
13	.788	.379							
14	.832	.307							
15	.787	.381							
16	.850	.277							
Reliability									
ω	.940		.824		.897		.901		.912
α	.912		.753		.855		.836		.868

Note. All results are significant at $p \leq .05$; GDS-ID: Glasgow depression scale for people with intellectual disabilities; GAS-ID: Glasgow Anxiety scale for people with intellectual disabilities; ω: omega coefficient of composite reliability (McDonald, 1970); α: Cronbach alpha coefficient of scale score reliability; Goodness of fit: χ^2=1937.118 (df =1463), p<.001; RMSEA=.030 [.026-.033]; CFI=.966; TLI=.964.

Table S2

Measurement Model for the Outcomes Rated by the Teachers and the Parents

	Depressive Mood				General Anxiety				Social Avoidance			
	Teacher		Parent		Teacher		Parent		Teacher		Parent	
Items	λ	δ	λ	δ	λ	δ	λ	δ	λ	δ	λ	δ
1	.631	.602	.657	.569	.919	.155	.895	.199	.742	.449	.657	.543
2	.934	.128	.805	.352	.653	.573	.820	.327	.901	.188	.805	.204
3	.873	.238	.864	.254	.903	.185	.916	.160	.575	.669	.864	.578
4	.811	.343	.721	.479	.881	.225	.806	.350	.905	.180	.721	.184
5	.488	.762	.695	.518	.874	.236	.928	.139	.569	.676	.695	.518
6	.909	.174	.766	.413	.784	.385	.760	.423	.779	.394	.766	.405
7	.876	.233	.802	.356	.467	.782	.668	.554	.835	.303	.802	.190
Reliability												
ω	.925		.906		.922		.940		.908		.915	
α	.871		.899		.840		.854		.813		.810	

Note. All results are significant at $p \leq .05$; ω: omega coefficient of composite reliability (McDonald, 1970); α: Cronbach alpha coefficient of scale score reliability; Goodness of fit of the teacher model: χ^2=628.447 (df =186), p<.001; RMSEA=.092 [.084-.100]; CFI=.944; TLI=.937. Goodness of fit of the parent model: χ^2=380.562 (df =186), p<.001; RMSEA=.077 [.066-.088]; CFI=.950; TLI=.943.