UC Santa Barbara
UC Santa Barbara Previously Published Works

Title

The Me and My School Questionnaire: Examining the cross-cultural validity of a children's self-report mental health measure

Permalink

https://escholarship.org/uc/item/4j93w69g

Journal

International Journal of School & Educational Psychology, 9(1)

ISSN

2168-3603

Authors

Moffa, Kathryn
Wagle, Rhea
Dowdy, Erin
et al.

Publication Date

2021-01-02

DOI

10.1080/21683603.2019.1650858

Peer reviewed

The Me and My School Questionnaire:

Examining the Cross-Cultural Validity of a Children's Self-Report Mental Health Measure

Kathryn Moffa, Rhea Wagle, Erin Dowdy

University of California Santa Barbara

Olympia Palikara, Susana Castro

University of Roehamptom

Danielle Dougherty Michael J. Furlong

University of California Santa Barbara

Under Review January 2019

Abstract

The Me and My School Questionnaire (M&MS) is a brief self-report measure of elementary school students' social, emotional, and behavioral challenges. This study examined its factor structure and measurement invariance for elementary school students in the United States (U.S.; $N = 784$) and the U.K. ($N = 538$). Results replicated the two-factor structure. Convergent and discriminant validity, test-retest reliability, and measurement invariance for girls and boys were examined in the U.S. sample. Partial measurement invariance was established when comparing factor structures of the U.S. and U.K. samples. Implications for mental health monitoring, and for comparative international research are discussed.

Keywords: elementary school, mental health, screening, United States, United Kingdom, emotional and behavioral problems, self-report

The Me and My School Questionnaire:

Examining the Cross-Cultural Validity of a Children's Self-Report Mental Health Measure

Mental health problems in children and adolescents are common with onset often occurring in late childhood and early adolescence (Merikangas et al., 2010). One-time prevalence rates of youth mental health disorders in the United States (U.S.) are reported to range from 13% to 40% (Center for Disease Control and Prevention, 2013; Kessler, Petukhova, Sampson, Zaslavsky, & Wittchen, 2012), and it is estimated that 25% of school-age children will meet criteria for an emotional or behavioral disorder at some point during their childhood years (Forness, Kim, & Walker, 2012). In the United Kingdom (U.K.), approximately one in 10 children and youth, between the ages of 5-16, experience a psychiatric disorder (Green et al., 2005). Despite the high need for childhood mental health services, it is estimated that only 20% to 45% of children with mental health difficulties in the U.S. actually receive formal care (Costello, He, Sampson, Kessler, & Merikangas, 2014; National Institute of Mental Health, 2009). Similarly, in the U.K, only 25% of the children and youth with clinically diagnosed psychiatric disorders will receive the needed specialist care (Ford et al., 2007). Across contexts, many children and adolescents need support because untreated mental health challenges can lead to long-term negative outcomes. Student mental health difficulties, for example, are linked with poor academic achievement and school attendance (DeSocio & Hootman, 2004), school dropout (Freudenberg & Ruglis, 2007), and involvement in the juvenile justice system (Cocozza & Skowyra, 2000, Joint Commissioning Panel for Mental Health, 2013). Additionally, school-age children with behavioral difficulties are at increased risk for school maladjustment and antisocial activity later in life (Schofield, Bierman, Heinrichs, & Nix, 2008).

Prevention efforts support universal screening for mental health as an efficient way to

identify youth in need of mental health services (Husky, Kaplan et al., 2011; Husky, Sheridan, McGuire, & Olfson, 2011). Mental health screening/monitoring in schools offers the opportunity to search for, and find, youth who experience social, emotional, and behavioral difficulties so that they can be connected with appropriate supports and resources (Dowdy, Ritchey, & Kamphaus, 2010; Rones & Hoagwood, 2000). Schools are an important context for universal screening because they are the place where students most often receive mental health supports (Marks & LaRosa, 2012). Early and universal screening in schools has the potential to identify students with subthreshold psychological symptoms that do not meet the criteria for formal psychological disorders (Flett & Hewitt, 2013). When school-based screening includes a self-report component, it has the additional advantage of identifying students with internalizing distress, who are less likely to be referred for services than students who are behaviorally disruptive (Weist, Rubin, Moore, Adelsheim, & Wrobel, 2007). Husky, Sheridan, and colleagues (2011) found that most students identified and referred for mental health services through universal screening received treatment; that is, school-based mental health screening successfully linked students to needed interventions.

Even with the promising evidence on the utility of mental health screening in schools, few schools implement these procedures. Bruhn, Woods-Groves, and Huddle (2014) estimated that only 12% of U.S. schools utilized universal mental health screeners. More pointedly, only 25% of school mental health practitioners assess students' social-emotional needs through evidence-based mental health assessment and screening practices, whereas 80% rely on academic data, such as grades, attendance, and office discipline referrals (Connors, Arora, Curtis, & Stephan, 2015). These commonly used school archival data, including teacher referrals, under-identify at-risk students when compared with structured or standardized rating scales (Eklund &

Dowdy, 2014).

It is challenging for elementary schools to identify age-appropriate, validated measures for mental health screening. Most self-report, universal screening measures are developed and normed with students ages 11 and older (Levitt, Saka, Romanelli, & Hoagwood, 2007). For younger students, teacher-report or parent-report screeners are more frequently utilized (Levitt et al., 2007). Although parent and teacher perspectives are useful, self-report screening measures offer the advantage of obtaining the child-centric perspective. With respect to considering the cognitive complexity needed for children to meaningfully respond to mental health self-report measures, Riley (2004) found that children as young as eight years old could provide valid responses on self-report questionnaires regarding their health. This supports research that explores the psychometric properties of mental health screeners for elementary school-age children. To aid in the implementation of early preventive universal screening efforts, there is a need for age-appropriate self-report measures that are efficient, accurate, and feasible for universal administration in school contexts. The present study seeks to contribute to childhood mental health research and practice by examining the psychometric properties of a child self-report screening measure (Me and My School questionnaire; Wolpert et al., 2011).

The Me and My School Questionnaire

The Me and My School Questionnaire (M&MS) was designed and developed as part of a nationwide, targeted mental health intervention in the U.K. to address the need for a brief, meaningful self-report screening measure for use with children (Wolpert et al., 2011). The M&MS is a self-report measure of general mental health that includes both emotional and behavioral difficulties (Patalay, Deighton, Fonagy, Vostanis, & Wolpert, 2014). It was developed with children 8 to 12 years old. Initial validation evidence in support of the instrument included

an exploratory factor analysis (EFA), a differential item functioning analysis (DIF), and a

confirmatory factor analysis (CFA). Items without salient factor loadings, with high cross-

loadings, and with differential functioning across gender, economic background, and special

educational needs were deleted. The analyses supported a two-factor model: *Emotional*

Difficulties and *Behavioral Difficulties* (Deighton et al., 2012). Additionally, the measure had

good internal consistency and external validity when compared with the *Strengths and*

Difficulties Questionnaire (SDQ; Deighton et al., 2012). Subsequent analyses identified cut

scores that adequately discriminated clinical from community samples (Patalay et al., 2014).

Sharpe and colleagues (2017) administered the KIDSCREEN-10 questionnaire (Ravens-Sieberer

et al., 2006) and M&MS to assess children's appraisals of their quality of life and mental health

problems, respectively. Internal consistency of the two factors on the M&Ms were within

acceptable limits (i.e., $\alpha =.78$ and .79, respectively) and between-construct correlations were in

the expected direction, with students who reported fewer mental health problems endorsing

higher quality of life, and students who reported more mental health problems endorsing lower

quality of life.

Current Study

Initial M&MS validation studies supported its use as a promising instrument that supports

universal mental health screening with younger children in the U.K. However, the validity of the

M&MS for use with children from the U.S. has not yet been examined. An investigation of the

validity of the M&MS outside of the U.K. would build evidence of its psychometric properties

and the generalizability of its use across cultural contexts. The current study's primary aim was

to test the generalizability of the two-factor structure of the M&MS with a sample of U.S.

elementary school children. In that the primary purpose of the current study was to investigate

the usefulness of the M&MS with U.S. students, stability and measurement invariance between U.S. girls and boys were examined. Evidence of concurrent validity of the M&MS with other screening measures was explored. Lastly, the two-factor structure of the M&MS was tested with a new comparison sample of U.K. school children to inform its potential use for cross-cultural collaborations that monitor and support youth mental health.

Method

Participants

U.S. sample. U.S. participants were students from seven primary schools in California ($N = 784$). In 2016, 415 students participated in the mental health screening. In 2017, 369 additional students participated in the mental health screening. Across both years, the U.S. sample consisted of fourth-graders (27%), 36% fifth-graders (36%), and sixth-graders (37%). All students were between the ages of 8 and 12. Students' self-reported cultural group/ethnicity were as follows: White (40.9%), Latinx/Hispanic (21.7%), Native American or American Indian (8.7%), Asian American (5.5%), Black or African American (2.7%), Pacific Islander (1.3%), and identified as "other" (15.4%). Thirty students chose not to report ethnicity. Approximately 51% of students identified as girls, 49% as boys, and 0.3% did not include this information.

A subsample of students participated in a stability (test-retest) reliability analyses ($n = 138$). Students were from three schools and consisted of fourth-graders (29.9%), fifth-graders (74.6%), and sixth-graders (1.4%). Self-reported cultural group/ethnicity included: White (47.1%), Latinx/Hispanic (23.9%), Native American or American Indian (6.5%), Asian American (2.2%), Black or African American (1.4%), Pacific Islander (0.7%), and identified as "other" (12.3%). The stability subsample included 56% girls and 44% boys.

U.K. sample. Students from three primary schools in the U.K. participated in a mental

health screening in 2017 (*N* = 538). The sample included students in Year 3 (17.8%), Year 4

(26.8%), Year 5 (31.8%), and Year 6 (23.2%). All students were between the ages of 8 and 12.

Two students did not include school year information. As there were many options for cultural

group/ethnicity included in the U.K. questionnaire, only the most commonly identified cultural

groups are reported here: Indian (34.8%), White/British (13.0%), other White background

(8.7%), Pakistani (8.2%), any other Asian background (8.2%), and any other ethnic background

(5.0%). The other 22.1% of students identified as Bangladeshi, Black Caribbean, Black African,

Chinese, Mixed: White/Black African, Mixed: White/Asian, Mixed: Any other Mixed

Background, White Caribbean, White: Irish, Any other White background, Traveler: Irish

Heritage, and Gypsy/Roma. Approximately 50% of students identified as girls, 49% as boys,

with two students not providing gender information.

Primary Measure

   Me and My School Questionnaire. The Me and My School Questionnaire (M&MS;

Deighton et al., 2012) is a 16-item self-report survey assessing emotional and behavioral

difficulties. It was designed so that younger children could easily read and comprehend its items

(Deighton et al., 2012). Previous factor analyses conducted in England identified two factors:

Emotional Difficulties and *Behavioral Difficulties* (Deighton et al., 2012). Ten items loaded onto

the *Emotional Difficulties* factor (e.g., *I feel lonely*) and six items loaded onto the *Behavioral*

Difficulties factor (e.g., *I get very angry*). A three-point response option is employed: *never*,

sometimes, and *always*. One Behavioral Difficulties item (*I am calm*) is reverse coded with

higher subscale scores indicating more behavioral difficulties. For the current study's samples,

the internal consistencies were: Emotional Difficulties (U.S., α =.76; U.K., α =.78) and

Behavioral Difficulties (U.S., α =.68; U.K. = .73).

Validity Analyses Measures

Social Emotional Health Survey–Primary. The Social Emotional Health Survey–Primary (SEHS-P) is a measure of positive psychological traits among youth in Grades 4–6 (Furlong, You, Renshaw, O'Malley, & Rebelez, 2013).The present study used three SEHS-P subscales comprised of four items each: *gratitude* (e.g., "Do you feel thankful to go to your school?"), *zest* (e.g., "Do you wake up in the morning excited to go to school?"), and *optimism* (e.g., "Do you feel positive that good things will happen to you at school?"). These traits have previously been associated with children's well-being and positive school functioning (Boman, Furlong, Shocet, Lilles, & Jones, 2009; Froh et al., 2011; Weber, Wagner, & Ruch, 2016). The SEHS-P has evidence supporting measurement invariance across genders and is positively correlated with youths' prosocial behavior, sense of school membership, and sense of school safety (Furlong et al., 2013). Items were measured using a six-point response scale (1 = *no, never*, 2 = *no, almost never*, 3 = *yes, sometimes*, 4 = *yes, often*, 5 = *yes, very often*, and 6 = *yes, always*). In the current study, mean scores on the gratitude, zest, and optimism subscales were used in validity analyses with the U.S. sample. Internal consistency reliability estimates for each subscale were: gratitude (α =.70), zest (α =.78), and optimism (α =.77).

Bully victimization. Two items were utilized to measure students' experiences of bullying victimization at school: *I get hit, kicked or pushed at school*, and *I get called names or threatened by other students at school*. Responses to these items were on a five-point scale (1 = *never*, 2 = *almost never*, 3 = *sometimes*, 4 = *often*, and 5 = *very often*) and used in analyses of construct validity with each latent construct of the M&MS questionnaire. For the current study, internal consistency of the items for the U.S. sample was: α =.74.

Procedures

In the U.S., students from seven California elementary schools participated in school-based universal mental health screening during the 2016–2017 and 2017–2018 school years. Letters were sent to parents requesting permission for their children to take the survey. Signed active consent was provided. Students were also asked to give assent prior to responding. Students completed an online survey (using Qualtrics®, a secure survey application) during a regular school day. Teachers were provided a script to read to the students that explained the content of the survey, the confidentiality policies, and each child's right to decline or discontinue participation at any time. The survey was designed to be easy for students to read, with large font and only one item per screen presentation. Procedures for the study were approved by the authors' university institutional review board.

In the U.K., a team of university researchers collected data at three primary schools between February and June 2017. In accordance with the university research ethics committee, passive consent forms were provided to students' parents and guardians so that they could opt out of survey participation. Instructions for survey completion were given by graduate research students prior to, and during, survey administration. Students completed paper versions of the survey.

Statistical Analyses

Data screening. Prior to conducting factor analyses, data were screened using SPSS 24. All M&MS items were categorical variables scored on a three-point scale, so they did not meet normality expectations (as measured by skewness, kurtosis, boxplots, and histograms). Additionally, bivariate correlations were calculated for each of the M&MS items. Correlations, means, and standard deviations are presented in Table 1. All items hypothesized to load onto the same factor were statistically significantly correlated with one another ($p < .01$), and no items

were multicollinear with other items.

Confirmatory factor analysis with U.S. sample. Confirmatory factor analysis (CFA) was conducted with the U.S. sample to investigate if the two-factor structure previously found among children in the U.K. also fit the item responses of the U.S. sample. Since all items were categorical variables, a robust mean and variance adjusted weighted least squares (WLSMV) estimator was used for all analyses (Brown, 2006). Goodness of model fit was assessed by the chi-square test of model fit, comparative fit index (CFI), Tucker-Lewis index (TLI), and root-mean-square error of approximation (RMSEA). There is some evidence that CFI, RMSEA, the TLI, and the robust WLSMV χ^2 statistic work well with categorical data (Flora & Curran, 2004; Hutchinson & Olmos, 1998; Yu & Muthén, 2002). Given that the use of the Weighted Root Mean Square Residual (WRMR) as a reliable fit statistic is inconsistent (Newsom, 2015), this statistic was not used in evaluating the model fit. Based on Hu and Bentler's (1999) recommendations, a cutoff value of .90 was established for the CFI and TLI, and a cutoff value of .08 was used for the RMSEA. Factor loadings of .30 and higher were considered adequate.

Multiple group analysis with boys and girls. To investigate the measurement invariance of the M&MS across boys and girls in the U.S. sample, multiple group analysis was conducted with M*plus* software 7.4 (Muthén & Muthén, 1998-2015). After confirming that the model fit the sample for boys and girls separately, multiple-group analysis was conducted to evaluate measurement invariance between U.S. boys and girls. This method consists of a series of steps that gradually constrain parameters to be equal between groups to determine whether there are differences in the way the scale functions across groups. If model fit substantially decreases, this contradicts measurement invariance. Tests of configural, metric, and scalar invariance were used to determine if the latent constructs are equivalent for girls and boys. First,

configural invariance was tested by allowing all parameters to be freely estimated, which establishes that the same set of indicators, or variables, measures the same factors across groups (Cheung & Rensvold, 2002). Results from configural invariance served as the baseline against which the other models were compared. Next, metric invariance was tested to determine if values of the factor loadings are equivalent across groups by constraining factor loadings to equality. Finally, scalar invariance was tested by constraining thresholds to be equal, with partial invariance pursued by freeing selected thresholds if necessary. If fit does not significantly worsen, this suggests that the likelihood from moving from one response choice to the next is equal between groups.

Measurement invariance is determined based on chi-square difference tests, ΔRMSEA, and ΔCFI. If the difference in chi-square values between the models for the boys and girls is significant, $p < .05$, then measurement invariance cannot be assumed. However, since chi-square has demonstrated sensitivity to large sample sizes (Hu & Bentler, 1998), ΔRMSEA $< .015$ and ΔCFI $< .01$ (Chen, 2007; Cheung & Rensvold, 2002) were also considered. If measurement invariance is established at the configural, metric, and scalar stages, then differences in factor mean scores for girls and boys can be examined and, if statistically significant, can be interpreted as actual differences in scores on the latent constructs rather than differences in factor structure between groups.

Stability. The stability of scores at approximately one-year follow up was examined with a sample of U.S. students ($n = 138$). The correlations between total scores on each subscale at time 1 (T1) and time 2 (T2) were calculated, as previous recommendations for score interpretation are based on subscale total scores.

Convergent and discriminant validity. Convergent and discriminant validity of the

M&MS were assessed. A structural path model was specified to examine the relations between each factor of the M&MS and the mean score for the gratitude, optimism, and zest subscales of the SEHS-P and each item measuring bullying victimization. Fit was assessed using the same criteria detailed above.

Multiple group analysis with the U.K. and U.S. samples. To investigate the applicability of the M&MS for cross-cultural research, measurement invariance of the M&MS questionnaire with U.S. and U.K. students was examined. The same steps taken to examine measurement invariance across genders were followed for cross-cultural multiple group analysis.

Results

Confirmatory factor analysis with U.S. sample. CFA was conducted on the U.S. sample to investigate evidence for the two-factor structure previously supported with samples of U.K. school students. Fit statistics indicated the two-factor model had good fit, RMSEA = .064, 90% CI [.058, .071], CFI = .923, and TLI = .910. Factor loadings on the *Emotional Difficulties* and *Behavioral Difficulties* factors were acceptable for all items. The item "I am shy" on the *Emotional Difficulties* factor demonstrated minimally acceptable fit (i.e., just above .30; see Table 2).

Multiple group analysis with boys and girls. Following CFA on the U.S. sample, CFA was performed separately with the sample of U.S. girls and boys. Fit information for the two-factor structure with both subsamples indicated adequate fit (see Table 3). All factor loadings for the subsample of U.S. boys were above .40 and statistically significant, $p < .001$. Overall, factor loadings for the subsample of U.S. girls were strong, although the factor loading for the item "I am shy" was below .30. The configural model was tested next by freeing all parameters. The model demonstrated good fit: $\chi^2 (208) = 455.31$, $p < .001$, CFI = .942, and RMSEA = .055, 90%

CI [.048, .062] (Table 3). To test equality of factor loadings across the subsamples of girls and

boys, the metric model was estimated by constraining factor loadings to be equal. The fit slightly

improved: χ^2 (222) = 466.59, p < .001, CFI = .943, and RMSEA = .053, 90% CI [.046, .060]

(Table 3). Although $\Delta\chi^2$ was significant, Pendergast, von der Embse, Kilgus, and Eklund (2017)

recommended that ΔCFI and ΔRMSEA are adequate for assessing measurement invariance

between two groups. Lastly, the intercepts were constrained to be equal across groups to test the

scalar model. The resulting model indicated adequate fit, $\Delta\chi^2$ = 42.671, ΔCFI = .010, and

ΔRMSEA = .001.

Stability. The one-year stability coefficients were as follows: Emotional Difficulties (r =

.65, p < .001) and Behavioral Difficulties (r = .47, p < .001).

Convergent and discriminant validity. Construct validity of the M&MS Emotional

Difficulties and Behavioral Difficultie*s* factors were explored by examining the simultaneous

relations among the M&Ms factors and gratitude, zest, optimism, and bullying victimization. As

expected, path analyses indicated negative relations with Emotional and Behavioral Difficulties

and subscale mean scores for gratitude (R^2 = .12), zest (R^2 = .10), and optimism (R^2 = .27).

Additionally, as expected, these factors were positively related to both forms of bullying

victimization (R^2_{Item1} = .20 and R^2_{Item2} = .19). The overall model demonstrated good fit to the

data, χ^2 = 630.60, df = 173, p < .001; SRMR = .05; RMSEA = .058, 90% CI [.053, .063]. Figure

1 presents the standardized coefficients of the path model.

Multiple-group analysis with the U.K. and U.S. samples. A multiple-group analysis

was conducted to evaluate the measurement invariance for U.S. and U.K. students. First, CFA

was conducted with the overall sample, which included all students from the U.S. and U.K.

samples. Fit statistics indicated good fit for the two-factor model previously supported among

other U.K. samples, χ^2 (103) = 663.066, $p < .001$, RMSEA = .064, CFI = .926 (Table 4). CFA was conducted with each sample separately, with both models producing good fit (see Table 4). Furthermore, all factor loadings were greater than .30 and statistically significant for both samples ($p < .001$). Thus, a two-factor model with Emotional Difficulties and Behavioral Difficulties was supported for the U.K. and U.S. samples.

Next, the configural model was estimated, freeing all parameters and testing invariance of factor structure. The model produced adequate fit: χ^2 (206) = 755.14, $p < .001$, CFI = .926, and RMSEA = .064, 90% CI [.059, .069] (Table 4). The metric model was estimated to explore equality of factor loadings across samples. Fit statistics indicated that constraining factor loadings to be equal did not significantly worsen fit, and in fact improved fit: ΔCFI = .018 and ΔRMSEA = -.011 (Table 4). Finally, the scalar model was estimated by constraining the intercepts to be equal across groups. The resulting model did not indicate adequate fit, $\Delta\chi^2$ = 37.913, ΔCFI = .036, and ΔRMSEA = .012. Partial invariance was explored by iteratively freeing thresholds based on modification indices. Eight thresholds were allowed to be freely estimated until fit statistics indicated no significant change in model fit. All freely estimated thresholds were those estimating the probability of an item response changing from never to sometimes. While partial invariance was found due to the estimation of thresholds, only eight thresholds out of 48 were freed, suggesting that factor means may be compared between the U.S. and U.K. samples.

Discussion

The current study investigated the use of the self-report M&MS as a brief measure of distress with U.S. children in upper elementary school grades. Although psychometric properties of the M&MS with children in the U.K. were previously reported, examination of its use with the

U.S. children in the current study further evaluated its generalizability as a measure with reliability and validity evidence supporting its use in school-based mental health screening. The current study's first analysis evaluated the original two-factor structure of the M&MS with a sample of U.S. elementary school students. Measurement invariance with girls and boys was also investigated. Next, the stability of the M&MS subscale scores was tested among a subsample of the U.S. students. Convergent and discriminant validity were examined with measures of wellness and bullying victimization. Finally, cross-national measurement invariance with the U.S. sample and a new sample of U.K. elementary school children was tested to evaluate the generalizability of the M&MS as a emotional-behavioral distress measure.

Core Psychometric Properties

Structural validity. Confirmatory factor analysis replicated the original two-factor M&MS structure (Deighton et al., 2012) with U.S. elementary school-age children. This finding provides initial support for interpreting both factors of the M&MS (i.e., *Emotional Difficulties* and *Behavioral Difficulties*) as meaningful for U.S. students.

Stability. Results from tests of one-year stability on subscale scores suggested the measure yields relatively stable scores for both subscales. This indicates that the measure assesses more than just recent events that might cause distress, and may measure enduring symptoms of emotional and behavioral difficulties.

Convergent and discriminant validity. The Emotional and Behavioral Difficulties factors and mean scores on measures of positive social-emotional health (i.e., gratitude, zest, and optimism) were negatively related, with small to moderate effects (i.e., R^2) of the Emotional and Behavioral Difficulties factors on gratitude, zest, and optimism. Students reporting more emotional-behavioral difficulties were less likely to endorse high levels of positive social-

emotional health than students reporting fewer emotional-behavioral problems. The effect sizes

provide evidence that there is value in asking primary school children about both their emotional

and behavioral difficulties and indicators of well-being. Additionally, the Emotional and

Behavioral Difficulties factors were both positively related to being the victim of bullying, as

shown by the small effects of the Emotional and Behavioral Difficulties factors on bullying

victimization. The modest effect size is expected because a number of circumstances can

precipitate social, emotional, and behavioral challenges, only one of them being bullying (Rhee,

Lee, & Jung, 2017). Conversely, literature on childhood resilience suggests that not all students

who are the victims of bullying experience at-risk or clinical levels of emotional and behavioral

problems (Moore & Woodcock, 2017).

    Invariance analyses. Analyses provided evidence of measurement invariance for U.S.

girls and boys. This finding provides evidence that the factor mean scores on the M&MS can be

compared across genders with the assumption that items hold the same meaning and similar

weight for U.S. girls and boys. Invariance testing for the U.S. and U.K. samples produced partial

measurement invariance (due to differences in eight item-response thresholds). This level of

invariance may justify making comparisons in factor scores between U.S. and U.K. children;

however, replication is needed prior to making factor mean comparisons.

Implications and Applications

    The results of the present study provide evidence supporting the use of the M&MS as a

brief measure of upper elementary school-age children's self-reported emotional and behavioral

difficulties for use among English-speaking U.S. and U.K. children. These findings contribute to

the body of psychometric evidence in support of the M&MS for use as a universal

screening/monitoring measure of emotional distress with U.S. and U.K. upper elementary

school-age children. Findings that the M&MS has partial invariance across national samples indicate that caution should be used prior to making factor mean comparisons between U.S. and U.K. students. Still, results from the current study warrant more generalization investigations when the M&MS is considered for use in international studies of youth mental health.

Limitations and Future Directions

To build evidence for using the M&MS for school-based wellness screening, replication with other U.S. samples is needed; this optimally should include research to establish normative information. Further work is also needed to examine the convergent validity of the M&MS with other measures of social, emotional, and behavioral distress, and particularly with other U.S. samples. Finally, future research needs to address possible issues with the factor structure of the M&MS. The factor loading for the item "I am shy" was barely adequate in CFA conducted with the U.S. sample, which is consistent with previous findings among U.K. samples (Deighton et al., 2012). It is possible that this item does not accurately reflect emotional difficulties with the same strength as the other items. While measurement invariance across cultures is difficult to establish, it is a critical next step to move forward with cross-cultural and cross-national research utilizing the M&MS to assess youths' social-emotional difficulties. This study provides initial psychometric evidence in support of the M&MS for assessing emotional and behavioral difficulties in children, which is a critical step to early identification and intervention.

References

Boman, P., Fulrong, M. J., Shocet, I., Lilles, E., & Jones, C. (2009). Optimism and the school

context. In R. Gilman, E. S. Huebner, & M. J. Furlong (Eds.), *Handbook of positive*

psychology in schools (pp. 51–64). New York, NY: Routledge.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York, NY:

Guilford.

Bruhn, A. L., Woods-Groves, S., & Huddle, S. (2014). A preliminary investigation of emotional

and behavioral screening practices in K-12 schools. *Education & Treatment of Children,*

37, 611–634. doi:10.1353/etc.2014.0039

Centers for Disease Control and Prevention (CDC). (2013). Mental health surveillance among

children–United States, 2005–2011. *Morbidity and Mortality Weekly Report, 62*(02), 1–35.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance.

Structural Equation Modeling, 14, 464–504. doi:10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing

measurement invariance. *Structural Equation Modeling, 9*, 233–255.

doi:10.1207/S15328007SEM0902_5

Cocozza, J. J., & Skowyra, K. R. (2000). Youth with mental health disorders: Issues and

emerging responses. *Juvenile Justice, 7*, 3–12.

Connors, E. H., Arora, P., Curtis, L., & Stephan, S. H. (2015). Evidence-based assessment in

school mental health. *Cognitive and Behavioral Practice, 22*, 60–73.

doi:10.1016/j.cbpra.2014.03.008

Costello, E., He, J., Sampson, N., Kessler, R., & Merikangas, K. (2014). Services for adolescents

with psychiatric disorders: 12-month data from the National Comorbidity Survey-

Adolescent. *Psychiatric Services, 65,* 359–366. doi:10.1176/appi.ps.201100518

Deighton, J., Tymms, P., Vostanis, P., Belsky, J., Fonagy, P., Brown, A., … & Wolpert, M.

(2012). The development of a school-based measure of child mental health. *Journal of*

Psychoeducational Assessment, 31, 247–257. doi:10.1177/073482291246570

DeSocio, J., & Hootman, J. (2004). Children's mental health and school success. *The Journal of*

School Nursing, 20, 189–196. doi:10.1177/10598405040200040201

Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-based screening: A population-

based approach to inform and monitor children's mental health needs. *School Mental*

Health, 2, 166–176. doi:10.1007/s12310-010-9036-3

Eklund, K., & Dowdy, E. (2014). Screening for behavioral and emotional risk versus traditional

school identification methods. *Disability and Psychoeducation Studies, 6,* 40–49.

doi:10.1007/s12310-013-9109-1

Flett, G. L., & Hewitt, P. L. (2013). Disguised distress in children and adolescents "flying under the

radar": Why psychological problems are underestimated and how schools must respond.

Canadian Journal of School Psychology, 28, 12–27. doi:10.1177/0829573512468845

Flora D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation

for confirmatory factor analysis with ordinal data. *Psychological Methods, 9,* 466–491.

doi:10.1037/1082-989X.9.4.466

Forness, S. R., Kim, J., & Walker, H. M. (2012). Prevalence of students with EBD: Impact on

general education. *Beyond Behavior, 21,* 3–9. doi:10.1177/1063426611401624

Freudenberg, N., & Ruglis, J. (2007). Reframing school dropout as a public health issue. *Preventing*

Chronic Disease, 4. Available from, www.cdc.gov/pcd/issues/2007/oct/07_0063.htm

Froh, J. J., Fan, J., Emmons, R. A., Bono, G., Huebner, E. S., & Watkins, P. (2011). Measuring

gratitude in youth: Assessing the psychometric properties of adult gratitude scales in

children and adolescents. *Psychological Assessment, 23*, 311–324. doi:10.1037/a0021590

Ford, T., Hamilton, H., Meltzer, H., & Goodman, R. (2007). Child mental health is everybody's

business: The prevalence of contact with public sector services by type of disorder among

British school children in a three-year period. *Child and Adolescent Mental Health*, *12*,

13–20. doi:10.1111/j.1475-3588.2006.00414.x

Furlong, M. J., You, S., Renshaw, T. L., O'Malley, M. D., & Rebelez, J. (2013). Preliminary

development of the Positive Experiences at School Scale for elementary school children.

Child Indicators Research, 6, 753–775. doi:10.1007/s12187-013-9193-7

Green, H., McGinnity, Á., Meltzer, H., Ford, T., & Goodman, R. (2005). *Mental health of*

children and young people in Great Britain, 2004. Available from,

https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-

children-and-young-people-in-great-britain/mental-health-of-children-and-young-people-

in-great-britain-2004

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure

analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A*

Multidisciplinary Journal, 6, 1–55. doi:10.1080/10705519909540118

Husky, M. M., Kaplan, A., McGuire, L., Flynn, L., Chostowsky, C., & Olfson, M. (2011).

Identifying adolescents at risk through voluntary school-based mental health screening.

Journal of Adolescents, 34, 505–511. doi:10.1016/j.adolescence.2010.05.018

Husky, M. M., Sheridan, M., McGuie, L., & Olfson, M. (2011). Mental health screening and

follow-up care in public high schools. *Journal of the American Academy of Child and*

Adolescent Psychiatry, 50, 881–891. doi:10.1016//jaac.2011.05.013

Hutchinson, S. R., & Olmos, A. (1998). Behavior of descriptive fit indexes in confirmatory

    factor analysis using ordered categorical data. *Structural Equation Modeling: A*

    Multidisciplinary Journal, *5,* 344–364. doi:10.1080/10705519809540111

Joint Commissioning Panel for Mental Health (2013). *Guidance for commissioners of children*

    and adolescent mental health services. Available from, jcpmh-rehab-guide.pdf

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H. U. (2012).

    Twelve-month and lifetime prevalence and lifetime morbidity risk of anxiety and mood

    disorders in the United States. *International Journal of Methods in Psychiatric Research,*

    21, 169–184. doi:10.1002/mpr.1359

Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental

    health problems in schools: The status of instrumentation. *Journal of School Psychology,*

    45, 167–191. doi:10.1016/j.jsp.2006.11.005

Merikangas, K. R., He, J., Burnstein, M., Swanson, S. A., Avenevoli, S., Cui, L., … Swendsen,

    J. (2010). Prevalence of mental disorders in U.S. adolescents: Results from the National

    Comorbidity Study – Adolescent Supplement (NSC-A). *Journal of the American*

    Academy Child and Adolescent Psychiatry, 49, 980–989. doi:10.1016/j.jaac.2010.05.017

Moore, B., & Woodcock, S. (2017). Resilience, bullying, and mental health: Factors associated

    with improved outcomes. *Psychology in the Schools, 54*, 689–702.

    doi:2048/10.1002/pits.22028

Muthén, B.O., du Toit, S., & Spisic, D. (1997). *Robust inference using weighted least squares*

    and quadratic estimating equations in latent variable modeling with categorical and

    continuous outcomes. Unpublished manuscript.

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus user's guide* (7th ed.). Los Angeles, CA:

Muthén & Muthén.

Newsom, J. (2015). *Practical approaches to dealing with nonnormal and categorical variables.*

Available from, http://www.upa.pdx.edu/IOA/newsom/semclass/ho_estimate2.pdf

Patalay, P., Deighton, J., Fonagy, P., Vostanis, P., & Wolpert, M. (2014). Clinical validity of the

Me and My School questionnaire: A self-repot mental health measure for children and

adolescents. *Child and Adolescent Psychiatry and Mental Health, 8*(17), 1–7.

doi:10.1186/1753-2000-8-17

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2018). Measurement

equivalence: A non-technical primer on categorical multi-group confirmatory factor

analysis in school psychology. *Journal of School Psychology, 60*, 65–82.

doi:10.1016/j.jsp.2016.11.002

Ravens-Sieberer, U., Gosch, A., Abel, T., Auguier, P., Bellach, B. M., Bruil, J., … Rajmil, L.

(2006). Quality of life in children and adolescents: A European public health perspective.

Soz Praventivmed, 46, 294–302.

Rhee, S., Lee, S., & Jung, S. (2017). Ethnic differences in bullying victimization and

psychological distress: A test of an ecological model. *Journal of Adolescence, 60*, 155–

160. doi:2048/10.1016/j.adolescence.2017.07.013

Riley, A. W. (2004). Evidence that school-age children can self-report on their health.

Ambulatory Pediatrics, 4, 371–376. doi:10.1367/A03-178R.1

Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review.

Clinical Child and Family Psychology Review, 3, 223–241. doi:10.1023/A:1026425104386

Schofield, H. L., Bierman, K. L., Heinrichs, B., & Nix, R. L. (2008). Predicting early sexual

activity with behavior problems exhibited at school entry and in early adolescence.

Journal of Abnormal Child Psychology, 36, 1175–1188. doi:10.1007/s10802-008—9252-6

Sharpe, H., Patalay, P., Vostanis, P., Belsky, J., Humphrey, N., & Wolpert, M. (2017). Use, acceptability and impact of booklets designed to support mental health self-management and help seeking in school: Results of a large randomized controlled trial in England. *European Child & Adolescent Psychiatry, 26*, 315–324. doi:10.1007/s00787-016-0089-3

Weber, M., Wagner, L., & Ruch, W. (2016). Positive feelings at school: On the relationships between students' character strengths, school-relate affect, and school functioning. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-being, 17*, 341–355. doi:10.1007/s10902-014-9597-1

Weist, M. D., Rubin, M., Moore, E., Adelsheim, S., & Wrobel, G. (2007). Mental health screening in schools. *Journal of School Health, 77*, 53–58. doi:10.1111/j.1746-1561.2007.00167.x

Wolpert, M., Deighton, J., Patalay, P., Martin, A., Fitzgerald-Yau, N., Demir, E., … Leyland, A. (2011). *Me and My School: Findings from the National Evaluation of Targeted Mental Health in Schools 2008-2011* (Research Report RR232). London, UK: Department for Children, Schools and Families.

Yu, C. Y., & Muthén, B. O. (2002). *Evaluation of model fit indices for latent variable models with categorical and continuous outcomes* (Technical Report). Los Angeles, CA: University of California at Los Angeles, Graduate School of Education & Information Studies.

Table 1

Descriptive Statistics of Item Responses on the M&MS with U.S. and U.K. Students

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Lonely	—															
2. Unhappy	.31**	—														
3. Likes	.33**	.23**	—													
4. Cry	.32**	.24**	.24**	—												
5. School	.30**	.32**	.17**	.23**	—											
6 Worry	.32**	.29**	.24**	.33**	.45**	—										
7. Sleep	.27**	.28**	.16**	.19**	.26**	.23**	—									
8. Wake up	.19**	.17**	.18**	.18**	.20**	.24**	.42**	—								
9. Shy	.20**	.11**	.13**	.20**	.23**	.27**	.08*	.08*	—							
10. Scared	.32**	.33**	.20**	.33**	.34**	.37**	.23**	.21**	.22**	—						
11. Angry	.30**	.27**	.23**	.31**	.21**	.28**	.27**	.23**	.09**	.29**	—					
12. Temper	.25**	.29**	.23**	.25**	.24**	.25**	.28**	.21**	.07	.28**	.51**	—				
13. Hurt	.06	.13**	.18**	.10**	.06	.07	.11**	.13**	-.04	.13**	.23**	.26**	—			
14. Calm	-.18**	-.21**	-.19**	-.18**	-.15**	-.15**	-.22**	-.14**	.01	-.16**	-.26**	-.27**	-.16**	—		
15. Hit	.10**	.12**	.18**	.14**	.10**	.08*	.17**	.17**	-.004	.14**	.24**	.36**	.45**	-.22**	—	
16. Break	.08*	.08*	.15**	.07*	.01	.09*	.18**	.17**	.033	.13**	.18**	.26**	.30**	-.15**	.33**	—
M	1.47	1.68	1.39	1.39	1.60	1.69	1.63	1.88	1.73	1.58	1.56	1.49	1.10	2.32	1.15	1.13
SD	.57	.51	.57	.54	.59	.60	.66	.64	.59	.54	.54	.56	.30	.56	.37	.34

$p < .05.$ $^{**}p < .01.$

Table 2

Standardized Factor Loadings for the U.S. Students

Item	Emotional Difficulties	Behavioral Difficulties
1. I feel lonely.	.67	
2. I am unhappy.	.65	
3. Nobody likes me.	.54	
4. I cry a lot.	.62	
5. I worry when I am at school.	.63	
6. I worry a lot.	.68	
7. I have problems sleeping.	.59	
8. I wake up in the night.	.51	
9. I am shy.	.33	
10. I feel scared.	.68	
11. I get very angry.		.80
12. I lose my temper.		.82
13. I do things to hurt people.		.71
14. I am calm. (reverse coded)		.51
15. I hit others when I am angry.		.74
16. I break things on purpose.		.57

All *p* < .001.

Table 3

Fit Statistics for Multiple Group Analysis across Gender with U.S. Students

	χ^2	*df*	$\Delta\chi^2$	Δdf	RMSEA (90% CI)	ΔRMSEA	CFI	ΔCFI
Single Group Solutions								
Overall Sample	437.94	103	—	—	.064 (.058, .071)	—	.923	—
Boys (*n* = 384)	268.21	103	—	—	.065 (.055, .074)	—	.905	—
Girls (*n* = 398)	201.75	103	—	—	.049 (.039, .059)	—	.962	—
Measurement Invariance								
Configural	455.31	208	—		.055 (.048, .062)	—	.942	—
Metric	466.59	222	36.517	14	.053 (.046, .060)	-.002	.943	.001
Scalar	535.88	248	42.671	26	.054 (.048, .061)	.001	.930	.010

All χ^2 significant at *p* < .001.

Table 4

Fit Statistics for Multiple Group Analysis across U.S. and U.K. Students

	χ^2	df	$\Delta\chi^2$	Δdf	RMSEA (90% CI)	ΔRMSEA	CFI	ΔCFI
Single Group Solutions								
Overall Sample	663.06	103	—	—	.064 (.060, .069)	—	.926	—
U.S. (*n* = 779)	435.64	103	—	—	.064 (.058, .071)	—	.923	—
U.K. (*n* = 536)	319.46	103	—	—	.063 (.058, .070)	—	.934	—
Measurement Invariance								
Configural	755.14	206	—	—	.064 (.059, .069)	—	.927	—
Metric	636.98	222	172.75	14	.053 (.048, .058)	.011	.945	.018
Scalar (Thresholds)$_a$	730.16	242	35.52	20	.055 (.051, .060)	.002	.935	-.010

[a] Reported fit statistics for scalar model represent fit with eight thresholds freely estimated.
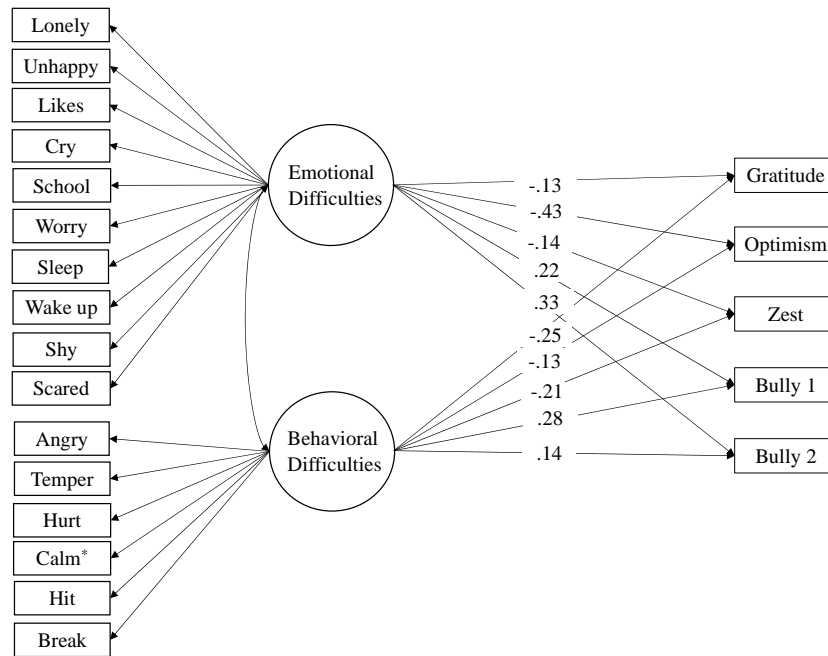
All χ^2 significant at *p* < .001.

Figure 1. Me and My School Questionnaire convergent validity model with U.S. students (all

standardized path coefficients are significant at *p* < .001).

*Reverse coded