Received: 10 August 2017 | Revised: 27 December 2017 | Accepted: 5 January 2018

DOI: 10.1002/mpr.1610

ORIGINAL ARTICLE

WILEY

Validation of the Mood and Feelings Questionnaire (MFQ) and Short Mood and Feelings Questionnaire (SMFQ) in New Zealand help-seeking adolescents

Hiran Thabrew | Karolina Stasiak | Lynda-Maree Bavin | Chris Frampton | Sally Merry

Department of Psychological Medicine, University of Auckland, Auckland, New Zealand

Correspondence
Hiran Thabrew, Department of Psychological Medicine, University of Auckland, Level 12 Support Block, Auckland Hospital, Grafton, Auckland 1149, New Zealand.
Email: h.thabrew@auckland.ac.nz

Funding information
New Zealand Ministry of Health

Abstract

Objective: This study examines the reliability and validity of the Mood and Feelings Questionnaire (MFQ) and Short Mood and Feelings Questionnaire (SMFQ) for measuring depression in New Zealand help-seeking adolescents.

Method: A sample of 183 adolescents completed the 33-item MFQ, which includes all 13 items on the SMFQ, at three time points during a trial of a computerized intervention for depression.

Results: Both the MFQ and SMFQ demonstrated good to excellent Cronbach's alphas, moderate to strong item-total score correlations, moderate to strong correlations with quality of life and anxiety measures, and strong correlations with the clinician-rated Children's Depression Rating Scale—Revised and the Reynolds Adolescent Depression Scale 2 at all time points, indicating good reliability and content, convergent, and concurrent validities, respectively. Favoring sensitivity over specificity, the optimal cut-off value for differentiating depressed from nondepressed cases for the MFQ was ≥28 and for the SMFQ was ≥12. Both instruments demonstrated satisfactory diagnostic accuracy and sensitivity to change.

Conclusion: The MFQ and SMFQ are free and simple instruments that can be used to identify depression and measure symptom change in New Zealand help-seeking adolescents.

KEYWORDS

adolescent, assessment, depression, mood and feelings questionnaire, validation

1 | INTRODUCTION

Depression is a common and significant health issue among New Zealand adolescents. During the most recent national school-based survey of adolescents in years 9 to 13 (Youth 12), 16% of female students and 9% of male students reported symptoms of depression that are likely to be clinically significant (i.e., likely to have an impact on a student's daily life; T. M. Fleming, Clark, & Denny, 2014). Depression in adolescents is associated with a number of adverse health-related, psychological, social, and economic outcomes (Weiss et al., 2005) and is one of the leading contributors to adolescent suicidal behavior and to completed suicide (Brent & Birmaher, 2002; J. E. Fleming, Boyle, & Offord, 1993; Lewinsohn, Rohde, & Seeley, 1998; Rao et al., 1995). Even mild-to-moderate depressive symptoms, if untreated, pose a significant risk for the young person's life trajectory, with a high likelihood of recurrence and development of other mental health problems in early adulthood (Harrington, Fudge, Rutter, Pickles, & Hill, 1990; Kovacs, 1996; Lewinsohn et al., 1998). Consequently, the early detection of depression and the demonstration of therapeutic efficacy of treatment regimens by change in symptom score are both important (Reynolds & Johnston, 1994; Reynolds & Stark, 1987). Psychometrically valid and reliable instruments are required to assess both.

Currently, internationally validated instruments for measuring depression in adolescents include, among others, the Reynolds Adolescent Depression Scale—Second Edition (RADS-2; Reynolds, 2002), the Children's Depression Rating Scale—Revised (CDRS-R; Poznanski et al., 1984), the Moods and Feelings Questionnaire (MFQ; Costello & Angold, 1988), and the Short Moods and Feelings Questionnaire (SMFQ; Angold, Costello, Messer, & Pickles, 1995). The MFQ is recommended by National Institute for Health and Clinical Excellence (2005) for the screening of depression in children and adolescents, is free to use, and has been validated with clinical and nonclinical samples (Sund,

Larsson, & Wichstrøm, 2001; Burleson Daviss et al., 2006; Wood, Kroll, Moore, & Harrington, 1995). The SMFQ is an appealing alternative to the MFQ because it takes half the time to complete; however, validation studies of the SMFQ have produced varied findings with regard to sensitivity, specificity, and recommended clinical cut-offs. Neither the MFQ nor the SMFQ has been validated with a New Zealand population, and it is uncertain whether the results of international studies are applicable to New Zealand adolescents, particularly those from indigenous (Maori) communities (Myers & Winters, 2002a). The significance of cross-cultural validation and adaptation of the content or cut-off points for psychometric instruments has been recommended by a number of previous researchers (Adewuya, Ola, Dada, & Fasoto, 2006; Beaton, Bombardier, Guillemin, & Ferraz, 2000). Gender is also known to affect the discriminatory validity of certain scales (Cherpitel, 1995). The current study aims to assess the reliability and validity of the MFQ and SMFQ in New Zealand with a help-seeking adolescent sample. Specifically, it examines the internal reliability, content validity, convergent validity, concurrent validity, diagnostic accuracy, and sensitivity to change of these measures. Optimal cut-offs for differentiating depressed from nondepressed cases, and by ethnicity and gender, will also be determined.

2 | METHODS

2.1 | Participants

Of 187 participants recruited for a computerized cognitive behavior therapy (CCBT) intervention trial for adolescent depression, 183 completed all items of the MFQ and 186 completed all items of the SMFQ at baseline. Demographic characteristics of both samples are presented in Table 1. All participants were help-seeking adolescents, aged 12 to 19 years, presenting to primary care youth clinics, general practices, and school-based counselling services in seven urban and provincial locations around New Zealand. Participants were required to have mild to moderate symptoms of depression at eligibility screening (indicated either by obtaining a score of 10–19 on the Patient Health Questionnaire – 9 [PHQ-9; Kroenke, Spitzer, & Williams, 2001] or by the clinician and adolescent suggesting that treatment was warranted), to have obtained at least 1 year of schooling in English, and to have access to a computer to complete the CCBT program. Exclusion criteria were severe depression, being assessed by their clinician as being at high risk of suicide or self-harm, having a score of 7 on Item 12 (morbid ideation) or a score of 5 or higher on Item 13 (suicidal ideation) on the CDRS-R, having a raw score of less than 30 on the CDRS-R (suggesting that a depressive disorder is unlikely), physical limitations or intellectual disability that would prevent the use of the CCBT program, the presence of a primary mental health disorder that was not depression, and currently or recently (in the past 3 months) receiving treatment with cognitive behavioral therapy, interpersonal therapy, or antidepressant medication. Further details about the participants, procedure, and findings of this intervention study have been published previously (Merry et al., 2012).

2.2 | Measures

The MFQ (Costello & Angold, 1988) is a 33-item questionnaire assessing depressive symptoms over the past 2 weeks, with responses rated on a 3-point scale (0 = *not true*, 1 = *sometimes*, and 2 = *true*). Total scores range from 0 to 66, with a score ≥29 commonly used in the literature (e.g., Burleson Daviss et al., 2006; Kent, Vostanis, & Feehan, 1997) to indicate clinically significant depression. The MFQ was originally designed for use with 8- to 18-year olds, is based on DSM-III-R symptom criteria, and has been recommended by National Institute for Health and Clinical Excellence (2005) for the screening of depression in children and adolescents. There are matching versions completed by the parent (MFQ-P) and the child or adolescent (MFQ-C,

TABLE 1 Baseline characteristics of the samples

Characteristic	MFQ sample at baseline assessment (*N* = 183)	SMFQ sample at baseline assessment (*N* = 186)
Gender		
Female	66.1%	66.1%
Male	33.9%	33.9%
Mean age	15.6 years (range 11.7–21, SD 1.59)	15.6 years (range 11.7–21, SD 1.59)
Ethnicity		
New Zealand European	59.0%	59.7%
Maori	24.0%	23.7%
Pacific people	8.2%	8.1%
Asian[a]	6.6%	6.4%
Other	2.2%	2.1%
Depressive symptoms		
CDRS-R (*n* = 180)	42.6 (26–71, 10.9)	42.6 (26–71, 10.8)
RADS-2 (*n* = 175)	75.0 (42–107, 13.7)	—
MFQ (*n* = 183)	28.4 (2–61, 12.5)	28.4 (2–61, 12.5)
SMFQ (*n* = 186)	—	12.1 (0–25, 6.0)

Note. Unless expressed as percentages, values are the means, ranges, and standard deviations. Three participants completed Short Mood and Feelings Questionnaire (SMFQ) questions but not all Mood and Feelings Questionnaire (MFQ) questions. CDRS-R = Children's Depression Rating Scale—Revised; RADS-2 = Reynolds Adolescent Depression Scale—Second Edition.

[a]Chinese, Japanese, Korean, Indian, Malaysian, Fijian Indian, Cambodian, and Thai.

used in this and most studies). It is freely available (Costello & Angold, 1988) and takes approximately 10 min to complete. It has been validated with clinical (Burleson Daviss et al., 2006; Kent et al., 1997; Wood et al., 1995) and nonclinical (Banh et al., 2012; Burleson Daviss et al., 2006; Sund, Larsson, & Wichstrøm, 2001) samples and is widely used in clinical and research settings.

The SMFQ (Angold et al., 1995) is a 13-item shortened version of the 33-item MFQ (Costello & Angold, 1988) and was developed in response to the need for a brief depression measure to reduce participant burden in research trials, while still retaining strong criterion validity (Angold et al., 1995). It assesses depressive symptoms over the past 2 weeks. Responses are rated on a 3-point scale (0 = *not true*, 1 = *sometimes*, and 2 = *true*). Like the MFQ, the SMFQ has matching versions completed by the parent (SMFQ-P) and the child or adolescent (SMFQ-C), and both are freely available (Costello & Angold, 1988). The SMFQ takes approximately 3 to 5 min to complete and has been validated with clinical (e.g., Kuo, Stoep, & Stewart, 2005) and nonclinical (e.g., Angold et al., 1995; Rhew et al., 2010; Sharp, Goodyer, & Croudace, 2006; Thapar & McGuffin, 1998; Turner, Joinson, Peters, Wiles, & Lewis, 2014) samples. Although it has been found to strongly correlate with the MFQ (r = .96, Angold et al., 1995; r = .95, Kuo et al., 2005), validation studies of the SMFQ have produced varied findings with regard to sensitivity and specificity. Recommended clinical cut-offs also vary between ≥4 (Rhew et al., 2010), ≥6 (Katon, Russo, Richardson, McCauley, & Lozano, 2008), ≥8 (Angold et al., 1995; Thapar & McGuffin, 1998), ≥10 (Kuo et al., 2005), and ≥11 (Patton et al., 2008; Turner et al., 2014).

The RADS-2 (Reynolds, 2002) is a 30-item, self-rated questionnaire assessing depressive symptoms over the past 2 weeks, with responses rated on a 4-point scale from 1 (*almost never*) to 4 (*most of the time*). Total scores range from 30 to 120, with scores ≥76 indicating that clinically significant depression may be present (Reynolds, 2002). The RADS-2 includes six critical items. Endorsement of four or more of these critical items suggests that the respondent is potentially at risk, even with a total depression score below the cut-off. In the current sample, responses to these critical items did not change any of the participants' diagnoses from that indicated by their total RADS-2 score. The RADS-2 is almost identical to the RADS. The RADS is recommended as a screening instrument for depression in community adolescent samples (Myers & Winters, 2002b) and has been validated in New Zealand, with consistency and validity findings being as strong as in the original validation conducted by Reynolds, and in overseas studies (Walker et al., 2005). Internal consistency of the RADS-2 for the current sample was excellent (Cronbach's α coefficient = .91). Both the RADS and the RADS-2 have a charge per use.

The CDRS-R (Poznanski et al., 1984) can be used for screening, as an aid for diagnosis, and to measure the severity of depression and is widely used in clinical research (Myers & Winters, 2002b). The CDRS-R was developed for children but is often used with adolescents (Myers & Winters, 2002b), and it has been found to have good reliability, validity, and sensitivity to change in children and adolescent populations (Mayes, Bernstein, Haley, Kennard, & Emslie, 2010; Myers & Winters, 2002b; Patel, Pharm, DelBello, & Strakowski, 2006; Poznanski & Mokros, 1996). It is an adaption of the adult Hamilton Depression Rating Scale (Hamilton, 1960). The CDRS-R contains 17 items that are each rated from 1 to 7 (or 1 to 5 on a select few questions) by a trained clinician, based on a 30 min interview with the child or adolescent, giving a total possible score of between 17 and 113. A score of ≥44 indicates that a depressive disorder is likely (Poznanski & Mokros, 1996). Internal consistency of the CDRS-R for the current sample was good (α = .84). Charges incurred include the purchase of the CDRS-R manual (currently US$118 each) and administration booklets (currently US$2 each).

The Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (Endicott, Nee, Yang, & Wohlberg, 2006) is a 15-item quality of life measure designed for use with children and adolescents. The Spence Children's Anxiety Scale (Spence, 1998) is a 44-item measure of childhood anxiety. Both measures were found to have good internal consistency in the current sample (α = .89 and α = .86, respectively).

2.3 | Data collection

Participants were tested at three time points: baseline, postintervention, and follow-up. The postintervention assessment was completed approximately 2 months after baseline, and the follow-up was completed 5 months after the start of the intervention.

2.4 | Analytical strategy

Data were analyzed using SPSS v23.0. All analyses included only the respondents who completed all items of the applicable measures. Internal reliability of the MFQ and SMFQ was evaluated by examining the Cronbach's alpha, which measures the average correlations between all of the scale items, with an α ≥ .90 considered as excellent, .85 ≤ α < .90 as good, .80 ≤ α < .85 as moderate, and .75 ≤ α < .80 as fair (Ponterotto & Ruckdeschel, 2007). Content validity was assessed by examining the item-total score correlations of the MFQ and SMFQ. Convergent validity was assessed by examining the correlations of the MFQ and SMFQ with scales of theoretically related constructs; specifically, the Spence Children's Anxiety Scale (Spence, 1998) and the Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (Endicott et al., 2006). Concurrent validity was assessed by examining the correlations between the MFQ and the clinician-rated CDRS-R and between the MFQ and the RADS-2. For the SMFQ, concurrent validity was assessed by examining the correlations between the SMFQ and the MFQ and clinician-rated CDRS-R.

The diagnostic accuracy and optimal cut-off value for diagnosing depression were assessed at baseline because this time point was prior to the study's intervention; hence, most representative for depression screening in primary care settings and for assessing eligibility for inclusion in depression intervention studies. Diagnostic accuracy was assessed at baseline by examining the percentage agreement of the MFQ's and SMFQ's (the index tests) diagnosis of depression against the CDRS-R's (the criterion/reference standard) diagnosis of depression (using the manual's cut-off value of ≥44 for likely depression), and for the SMFQ, this was also gauged by measuring alignment with the MFQ.

Diagnostic accuracy of the MFQ and SMFQ was also evaluated by examining the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the curve (AUC) using

receiver operating characteristic (ROC) curve analyses. Sensitivity is the proportion of the subjects with a condition who are accurately identified by the index test as positive for that condition (true positive rate), whereas specificity is the proportion of the subjects without a condition who are accurately identified by the index test as negative for that condition (true negative rate; Zou, O'Malley, & Mauri, 2007). PPV is the percentage of the subjects identified by the index test as having the condition who are true positive cases, whereas NPV is the percentage of the subjects identified by the index test as not having the condition who are true negative cases (Murphy & Davidshofer, 2005). ROC curves plot the sensitivity against 1—specificity (Fawcett, 2006; Zou et al., 2007). The AUC is a measure of a test's overall ability to accurately classify subjects with or without a condition (Brooks & Kutcher, 2001). An AUC of .5 corresponds to the accuracy of the test being equivalent to random chance, whereas an AUC of 1.0 represents perfect diagnostic accuracy (Zou et al., 2007). An AUC > .9 is considered "high," .7 < AUC ≤ .9 considered "moderate," and an AUC ≤ .7 considered "low" accuracy for differentiating cases from noncases (Henderson, 1993). Sensitivity to change is an instrument's ability to detect changes in the severity of a condition over time (Brooks & Kutcher, 2001).

The optimal MFQ and SMFQ cut-off values for the total sample, and by gender and ethnicity (for the two largest groups: NZ European and Maori), were determined by examining the relative sensitivities and specificities. Given that the MFQ and SMFQ may be used for the identification of depression in primary care settings and for assessing eligibility for inclusion in depression intervention studies, the favoring of sensitivity in the trade-off between specificity and sensitivity, and targets of 80% as a minimum sensitivity and 70% as a minimum specificity (as per Pettersson, Boström, Gustavsson, & Ekselius, 2015), was used as the basis for determining the optimal MFQ and SMFQ cut-off values.

Sensitivity to change was assessed by examining the agreement between the MFQ/SMFQ and the CDRS-R (as the criterion) for change in diagnosis from baseline to post-treatment, examining the correlations between the MFQ/SMFQ and CDRS-R in change scores from baseline to post-treatment, and comparing the treatment group effect sizes (ES) of the MFQ/SMFQ to the CDRS-R. The MFQ's sensitivity to change was also assessed against the RADS-2 and the SMFQ (using the CDRS-R as the criterion).

3 | RESULTS

3.1 | Demographics

Demographic characteristics are presented in Table 1.

3.2 | Internal reliability

Internal reliability was assessed by examining the Cronbach's alpha of the MFQ and SMFQ at each assessment time point. Based on Ponterotto and Ruckdeschel's (2007) classifications, the Cronbach's alphas of the MFQ were excellent for each time point (α = .91 to.93), and the SMFQ's were good (α = .88 to .89).

3.3 | Content validity

Content validity was assessed by examining the item-total score correlations of the MFQ and SMFQ. Across all three time points, 29 of the 33 MFQ items had item-total correlations of at least .40. The four items with item-total correlations smaller than .40 at one or more time points were, "I ate more than usual" (r = .03 to r = .27), "I didn't want to see my friends" (r = .36 to r = .51), "I worried about aches and pains" (r = −.26 to r = .41), and "I slept a lot more than usual" (r = .09 to r = .17). For the SMFQ, across all three time points, all 13 items had item-total correlations of at least .50.

3.4 | Convergent validity

Convergent validity was assessed by examining the correlation, at baseline, between the MFQ/SMFQ and two scales measuring constructs theoretically related to depression. The MFQ was significantly and strongly correlated with both the Spence Children's Anxiety Scale (r = .62, p < .001) and the Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (r = −.73, p < .001). The SMFQ was moderately correlated with the Spence Children's Anxiety Scale (r = .57, p < .001) and strongly correlated with the Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (r = −.77, p < .001).

3.5 | Criterion validity

Criterion validity was assessed by examining the concurrent validity and diagnostic accuracy of the MFQ and SMFQ. At each time point, MFQ scores were significantly and strongly correlated with the CDRS-R (r = .66 to .71, p < .001) and with the RADS-2 (r = .83 to .85, p < .001). The correlations between the MFQ and the CDRS-R were either the same or higher than the correlations between the RADS-2 and the CDRS-R at each time point (r = .66 to.69, p < .001). Based on an MFQ cut-off value of ≥29, which has previously been used with psychiatric (Kent et al., 1997), delinquent (Kuo et al., 2005), and research and psychiatric clinic (Burleson Daviss et al., 2006) samples, agreement between the MFQ and the CDRS-R (using ≥44 as the cut-off) for differentiating depressed from nondepressed subjects was 76.1% at baseline. In comparison, agreement between the RADS-2 (at the manual's recommended cut-off ≥76; Reynolds, 2002) and the CDRS-R was 73.0%. Sensitivity was 81.3% for the MFQ and 79.2% for the RADS-2, specificity was 72.4% for the MFQ and 68.6% for the RADS-2, PPV was 67.8% for the MFQ and 64.0% for the RADS-2, and NPV was 84.4% for the MFQ and 82.4% for the RADS-2. Hence, the MFQ had a higher diagnostic accuracy than the RADS-2 in our New Zealand sample. For the SMFQ, scores at each time point were strongly and significantly correlated with the CDRS-R (r = .65 to .70, p < .001) and with the MFQ (r = .95 to .96, p < .001). The correlations between the SMFQ and the CDRS-R were very similar (within .01) to the correlations between the MFQ and the CDRS-R at each time point.

Diagnostic accuracy was also assessed using the sample's sensitivity-favored optimal cut-off values for the SMFQ (≥ 12) and MFQ (≥ 28). Agreement between the SMFQ and the CDRS-R for differentiating depressed from nondepressed subjects at baseline was 74.9%, and agreement between the MFQ and the CDRS-R was 76.1%. At

these optimal cut-off values, sensitivities for the MFQ and SMFQ, respectively, were 84.0% and 84.2%, specificities were 70.5% and 68.2%, PPVs were 67.0% and 65.3%, and NPVs were 86.1% and 85.9%. Hence, the SMFQ had a very similar diagnostic accuracy to the MFQ in our New Zealand sample.

3.6 | Optimal cut-offs

The optimal cut-off value for differentiating depressed from nondepressed cases was examined across the total sample, and by gender and ethnicity (see Tables 2 and 3). Based on the cut-off achieving or most closely approximating, a target minimum sensitivity of 80% and a target minimum specificity of 70% and favoring sensitivity over specificity, the MFQ was optimized at ≥28 for the sample, with a corresponding sensitivity of 84.0% and specificity of 70.5% (see Table 2). The AUC of the ROC curve was .86, 95% CI [.81, .91], with .70 < AUC ≤ .90 regarded as moderate (Henderson, 1993). The SMFQ was optimized at ≥12 for the sample, with a corresponding sensitivity of 84.2% and specificity of 68.2% (see Table 3) and an AUC of .86, CI [.80, .91].

The total samples' MFQ optimal cut-off value of ≥28 (when favoring sensitivity) did not differ between the two largest ethnicity groups (NZ European and Maori) for which ROC analyses were conducted; however, the optimal cut-off did differ slightly by gender. When applying the same criteria for the target minimum sensitivity and specificity while favoring sensitivity, the optimal cut-off value was ≥28 for females (with a sensitivity of 85.4% and a specificity of 66.2%) and ≥29 for males (with a sensitivity of 81.5% and a specificity of 82.4%). For the SMFQ, there was a similarly subtle difference in optimal cut-off values by gender. The optimal cut-off value was ≥12 for males (with a sensitivity of 85.2% and a specificity of 74.3%) and ≥13 for females (with a sensitivity of 79.6% and a specificity of 75.0%). The optimal cut-off value also differed slightly by ethnicity. The optimal cut-off value was ≥12 for New Zealand European (with a sensitivity of 81.4% and a specificity of 65.7%) and ≥11 for Maori (with a sensitivity of 94.4% and a specificity of 76.0%) adolescents.

3.7 | Sensitivity to change

Sensitivity to change was assessed by examining the agreement between the MFQ/SMFQ (index tests) and the CDRS-R (criterion) for change in diagnosis from baseline to post-treatment, examining the correlations between the MFQ/SMFQ and CDRS-R in change scores, comparing the treatment group effect sizes of the MFQ/SMFQ and CDRS-R, and comparing the MFQ's sensitivity to change against the RADS-2. The agreement of the MFQ with the CDRS-R for change in diagnosis of likely depression was 64.4% for the sensitivity favored MFQ cut-off value of ≥28. The MFQ and CDRS-R change scores were strongly correlated (r = 0.64, p < .001). Comparing the MFQ's sensitivity to change against the RADS-2, the agreement between the RADS-2 (at the recommended cut-off ≥76; Reynolds, 2002) and the CDRS-R for the change in diagnosis for likely depression (64.2%) was similar, and the correlation of the RADS-2 and the CDRS-R change scores (r = .61, p < .001) was also similar. The MFQ had a treatment group

TABLE 2 Sensitivities, specificities, and predictive values for different cut-offs on the MFQ, using the CDRS-R (≥44) as the reference standard for a likely depressive disorder

Cut-off	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Total sample				
≥27	85.3 (75.3–92.4)	67.6 (57.8–76.4)	65.3 (55.0–74.7)	86.6 (77.3–93.1)
≥28[a]	84.0 (73.7–91.5)	70.5 (60.8–79.0)	67.0 (56.6–76.4)	86.1 (76.9–92.6)
≥29	81.3 (70.7–89.4)	72.4 (62.8–80.7)	67.8 (57.1–77.3)	84.4 (75.3–91.2)
Males				
≥27	81.5 (61.9–93.7)	76.5 (58.8–89.3)	73.3 (54.1–87.7)	83.9 (66.3–94.6)
≥28	81.5 (61.9–93.7)	79.4 (62.1–91.3)	75.9 (56.5–89.7)	84.4 (67.2–94.7)
≥29[a]	81.5 (61.9–93.7)	82.4 (65.5–93.2)	78.6 (59.1–91.7)	84.9 (68.1–94.9)
Females				
≥27	87.5 (74.8–95.3)	63.4 (51.1–74.5)	61.8 (49.2–73.3)	88.2 (73.1–95.6)
≥28[a]	85.4 (72.2–93.9)	66.2 (54.0–77.0)	63.1 (50.2–74.7)	87.0 (75.1–94.6)
≥29	81.3 (67.4–91.1)	67.6 (55.5–78.2)	62.9 (49.7–78.8)	84.2 (72.1–92.5)
New Zealand European				
≥27	83.3 (68.6–93.0)	66.2 (53.4–77.4)	61.4 (47.6–74.0)	86.0 (73.3–94.2)
≥28[a]	81.0 (65.9–91.4)	70.8 (58.2–81.4)	64.2 (49.8–76.9)	85.2 (72.9–93.4)
≥29	78.6 (63.2–89.7)	73.9 (61.5–84.0)	66.0 (51.2–78.8)	84.2 (72.1–92.5)
Maori				
≥27	88.9 (65.3–98.6)	76.0 (54.9–90.6)	72.7 (49.8–89.3)	90.5 (69.6–98.8)
≥28[a]	88.9 (65.3–98.6)	76.0 (54.9–90.6)	72.7 (49.8–89.3)	90.5 (69.6–98.8)
≥29	88.9 (65.3–98.6)	76.0 (54.9–90.6)	72.7 (49.8–89.3)	90.5 (69.6–98.8)

Note. MFQ = Mood and Feelings Questionnaire; CDRS-R (≥44) = Children's Depression Rating Scale—Revised, with a cut-off value ≥44 indicating a likely depressive disorder; PPV = positive predictive value; NPV = negative predictive value; CI = confidence interval.

[a]Optimal cut-off value based on achieving, or most closely approximating, a minimum sensitivity of 80% and a minimum specificity of 70%, and favoring sensitivity.

TABLE 3 Sensitivities, specificities, and predictive values, for different cut-offs on the SMFQ, using the CDRS-R (≥44) as the reference standard for a likely depressive disorder

Cut-off	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Total sample				
≥11	90.8 (81.9–96.2)	63.6 (53.7–72.6)	63.9 (54.1–72.9)	90.7 (81.7–96.2)
≥12[a]	84.2 (74.0–91.6)	68.2 (58.5–76.9)	65.3 (55.0–74.6)	85.9 (76.6–92.5)
≥13	77.6 (66.6–86.4)	76.6 (67.5–84.3)	70.2 (59.3–79.7)	82.8 (73.9–89.7)
Males				
≥11	92.6 (75.7–99.1)	65.7 (47.8–80.9)	67.6 (50.2–82.0)	92.0 (74.0–99.0)
≥12[a]	85.2 (66.3–95.8)	74.3 (56.7–87.5)	71.9 (53.3–86.3)	86.7 (69.3–96.2)
≥13	74.1 (53.7–88.9)	80.0 (63.1–91.6)	74.1 (53.7–88.9)	80.0 (63.1–91.6)
Females				
≥11	89.8 (77.8–96.6)	62.5 (50.3–73.6)	62.0 (49.7–73.2)	90.0 (78.2–96.7)
≥12	83.7 (70.3–92.7)	65.3 (53.1–76.1)	62.1 (49.3–73.4)	85.5 (73.3–93.5)
≥13[a]	79.6 (65.7–89.8)	75.0 (63.4–84.5)	68.4 (54.8–80.1)	84.4 (73.1–92.2)
New Zealand European				
≥11	90.7 (77.9–97.4)	59.7 (47.0–71.5)	59.1 (46.3–71.1)	90.9 (78.3–97.5)
≥12[a]	81.4 (66.6–91.6)	65.7 (53.1–76.9)	60.3 (46.6–73.0)	84.6 (71.9–93.1)
≥13	74.4 (58.8–86.5)	76.1 (64.1–85.7)	66.7 (51.6–79.6)	82.3 (70.5–90.8)
Maori				
≥11[a]	94.4 (72.7–99.9)	76.0 (55.0–90.6)	73.9 (51.6–89.8)	95.0 (75.1–99.9)
≥12	88.9 (65.3–98.6)	80.0 (59.3–93.2)	76.2 (52.8–91.8)	90.9 (70.8–98.9)
≥13	77.8 (52.4–93.7)	80.0 (59.3–93.2)	73.7 (48.8–90.9)	83.3 (62.6–95.3)

Note. SMFQ = Short Mood and Feelings Questionnaire; CDRS-R (≥44) = Children's Depression Rating Scale—Revised, with a cut-off value ≥44 indicating a likely depressive disorder; PPV = positive predictive value; NPV = negative predictive value; CI = confidence interval.

[a]Optimal cut-off value based on achieving, or most closely approximating, a target minimum sensitivity of 80% and specificity of 70%, while favoring sensitivity.

effect for change in depressive symptom scores (ES = 0.19) smaller than the CDRS-R (ES = 0.28) and similar to the RADS-2 (ES = 0.20).

The agreement of the SMFQ with the CDRS-R for change in diagnosis of likely depression was 66.7% for the sensitivity-favored SMFQ cut-off value of ≥12. The SMFQ and CDRS-R change scores were strongly correlated (r = 0.62, p < .001), and the SMFQ's treatment group effect for change in depressive symptom scores (ES = 0.23) approximated that of the CDRS-R's (ES = 0.28).

Comparing the MFQ and SMFQ's sensitivity to change, the agreement with the CDRS-R for the change in diagnosis for likely depression and the correlation with the CDRS-R in change scores were similar; however, the SMFQ's treatment group effect size more closely approximated the CDRS-R's effect size than did the MFQ.

4 | DISCUSSION

The MFQ demonstrates satisfactory validity, internal consistency, diagnostic accuracy, and sensitivity to change in a help-seeking adolescent sample with mild–moderate depressive symptoms and may be an appropriate instrument to use with New Zealand adolescents who are at risk of depression. In the current study, the sensitivity (84.0%) and specificity (70.5%) at the optimal MFQ cut-off value of ≥28 are above Pettersson et al.'s (2015) minimum recommended values. Given the current study's obtained sensitivity of 84%, most adolescents with a depressive disorder who are screened using the MFQ in a primary care setting will be identified and ideally referred for further evaluation. Furthermore, given the positive predictive value of 67%, the majority of adolescents who are screened in a primary care setting and identified as likely to have a depressive disorder will be true positive cases.

The concurrent and change score correlations between the MFQ and the CDRS-R were the same or slightly stronger than between the RADS-2 and the CDRS-R. Furthermore, using the CDRS-R as the criterion, the diagnostic accuracy of the MFQ was higher than that of the RADS-2 in our New Zealand sample. These results suggest that the MFQ may effectively be used in place of the RADS-2 when assessing adolescents in clinical or research settings. In contrast to the RADS-2 and the CDRS-R, which have a charge per use, the MFQ is free to use.

The MFQ's optimal cut-off value of ≥28 that was observed in the present study is similar to the ≥29 cut-off score that has previously been found to optimally differentiate between depressed and nondepressed cases in youth psychiatric (Kent et al., 1997), delinquent (Kuo et al., 2005), and research and psychiatric clinic (Burleson Daviss et al., 2006) samples and is also similar to Wood et al.'s (1995) optimal MFQ cut-off value of ≥27 for the diagnosis of major depression in a sample of adolescent psychiatric outpatients. In the present study, the MFQ's optimal cut-off value did not differ between the two largest ethnicity groups (NZ European and Maori) for which ROC analyses were conducted and only slightly differed between male (≥29) and female (≥28) adolescents. Our findings are similar to those of Banh et al. (2012), who concluded that the MFQ appeared to measure depression equivalently for male and female students, and across ethnicities (Whites, African Americans, Asian Americans, Hispanic

Americans), for their student sample from Seattle, Washington, and are also similar to Wood et al. (1995), who reported that the AUC did not significantly differ by gender in their adolescent psychiatric outpatient sample. The present study, therefore, provides further support to the international literature for the MFQ essentially being equivalent across gender and ethnicity.

The SMFQ's validity, internal consistency, diagnostic accuracy, and sensitivity to change were also satisfactory and very similar to that obtained with the MFQ. Both the SMFQ and MFQ are free to use, but the SMFQ takes approximately half the time to complete, and this is likely to appeal to adolescents (Stasiak et al., 2013). If a new measure is being assessed to potentially replace an existing measure because it is cheaper, faster, or easier to use, then there should be a high correlation between the two tests of .80 or above (Streiner, 1993). The very strong correlation between the 13-item SMFQ and the full 33-item MFQ exceeded the .80 recommended minimum. These results suggest that the SMFQ may effectively be used in place of the MFQ when assessing adolescents in primary care or research settings.

At the optimal SMFQ cut-off value of ≥12 in the present study, the corresponding sensitivity (84%) exceeds, and the specificity (68%) very closely approximates, Pettersson et al.'s (2015) minimum recommended values of 80% and 70%, respectively. The optimal cut-off value differed only slightly between NZ European (≥12) and Maori (≥11) and between males (≥12) and females (≥13). McKenzie et al. (2011) and Turner et al. (2014) also found that demographic variables had little impact on the SMFQ's discriminatory validity. These cut-off scores are similar to those suggested by Turner et al. (2014; ≥11 in a community sample of 17 to 18 year olds) and by Kuo et al. (2005; ≥10 in a delinquent sample of 15 to 16 year olds). They are, however, substantially higher than those suggested by Thapar and McGuffin (1998; ≥8 in a nonclinical sample of twins aged 8 to 16 years) and by Rhew et al. (2010; ≥4 in a student sample of 11 to 13 year olds).

Strengths of the current study include being the first to assess the reliability and validity of the MFQ and SMFQ in any New Zealand adolescent population; and, to our knowledge, the first to examine the SMFQ's sensitivity to change. Another strength is the systematic examination of a range of psychometric properties for both the MFQ and SMFQ. Limitations of this study include the lack of generalizabilty of results beyond a help-seeking population, the MFQ's diagnostic accuracy being assessed using the CDRS-R (which is not a "gold standard" clinical interview) as the criterion, SMFQ questions not being asked independent of MFQ questions, and the relatively small group sizes (particularly for those of Maori ethnicity and for males). Mitigating some of these limitations, the CDRS-R is a widely used clinician-rated scale with good reliability and validity (Mayes et al., 2010; Myers & Winters, 2002b; Poznanski & Mokros, 1996) that has previously been validated against the Kiddie Schedule for Affective Disorders and Schizophrenia (Yee et al., 2015), and it is reassuring that McKenzie et al. (2011) and Turner et al. (2014) found that demographic variables had little impact on the SMFQ's discriminatory validity. Banh et al. (2012) similarly concluded that the MFQ appeared to measure depression equivalently across gender and ethnicities. Replication with a larger help-seeking and nonhelp-seeking sample is recommended to provide more conclusive support for the cut-off values obtained for each subgroup and the wider adolescent population. We also recommend that future research assesses the SMFQ separate from the MFQ to further test that it is a psychometrically sound standalone instrument.

5 | CONCLUSION

Both the MFQ and SMFQ are free and easy to administer assessment instruments, which have now demonstrated satisfactory validity, internal consistency, diagnostic accuracy, and sensitivity to change in a New Zealand sample. Both instruments can justifiably be used by researchers and healthcare professionals to aid diagnosis, and to measure remission, or change in depressive symptoms in New Zealand help-seeking adolescents. Our findings provide further support to the extant literature for the MFQ's equivalency across ethnicities and the first international evidence of the MFQ's sensitivity to change.

ACKNOWLEDGEMENTS

The SPARX study was generously funded by the New Zealand Ministry of Health.

ORCID

Hiran Thabrew http://orcid.org/0000-0002-8755-6217

REFERENCES

Adewuya, A. O., Ola, B. A., Dada, A. O., & Fasoto, O. O. (2006). Validation of the Edinburgh postnatal depression scale as a screening tool for depression in late pregnancy among Nigerian women. *Journal of Psychosomatic Obstetrics and Gynecology*, *27*(4), 267–272.

Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, *5*(4), 237–249.

Banh, M. K., Crane, P. K., Rhew, I., Gudmundsen, G., Vander Stoep, A., Lyon, A., & McCauley, E. (2012). Measurement equivalence across racial/ethnic groups of the mood and feelings questionnaire for childhood depression. *Journal of Abnormal Child Psychology*, *40*(3), 353–367. https://doi.org/10.1007/s10802-011-9569-4

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191.

Brent, D. A., & Birmaher, B. (2002). Adolescent depression. *New England Journal of Medicine*, *347*(9), 667–671. https://doi.org/10.1056/NEJMcp012249

Brooks, S. J., & Kutcher, S. (2001). Diagnosis and measurement of adolescent depression: A review of commonly utilized instruments. *Journal of Child and Adolescent Psychopharmacology*, *11*(4), 341–376. https://doi.org/10.1089/104454601317261546

Burleson Daviss, W., Birmaher, B., Melhem, N. A., Axelson, D. A., Michaels, S. M., & Brent, D. A. (2006). Criterion validity of the mood and feelings questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of Child Psychology and Psychiatry*, *47*(9), 927–934. https://doi.org/10.1111/j.1469-7610.2006.01646.x

Cherpitel, C. J. (1995). Analysis of cut points for screening instruments for alcohol problems in the emergency room. *Journal of Studies on Alcohol*, *56*(6), 695–700.

Costello, E. J., & Angold, A. (1988). Scales to assess child and adolescent depression: Checklists, screens, and nets. *Journal of the American Academy of Child and Adolescent Psychiatry*, *27*(6), 726–737. https://doi.org/10.1097/00004583-198811000-00011

Endicott, J., Nee, J., Yang, R., & Wohlberg, C. (2006). Pediatric quality of life enjoyment and satisfaction questionnaire (PQ-LES-Q): Reliability and validity. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*(4), 401–407. https://doi.org/10.1097/01.chi.0000198590.38325.81

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Fleming, J. E., Boyle, M. H., & Offord, D. R. (1993). The outcome of adolescent depression in the Ontario child health study follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*(1), 28–33. https://doi.org/10.1097/00004583-199301000-00005

Fleming, T. M., Clark, T., & Denny, S. (2014). Stability and change in the mental health of New Zealand secondary school students 2007–2012: Results from the national adolescent health surveys. *Australian and New Zealand Journal of Psychiatry*, *48*(5), 472–480. https://doi.org/10.1177/0004867413514489

Hamilton, M. A. (1960). Rating scale for depression. *Journal of Neurological and Neurosurgical Psychiatry*, *23*, 56–62.

Harrington, R., Fudge, H., Rutter, M., Pickles, A., & Hill, J. (1990). Adult outcomes of childhood and adolescent depression: I. Psychiatric status. *Archives of General Psychiatry*, *47*(5), 465–473. https://doi.org/10.1001/archpsyc.1990.01810170065010

Henderson, A. R. (1993). Assessing test accuracy and its clinical consequences: A primer for receiver operating characteristic curve analysis. *Annals of Clinical Biochemistry*, *30*(6), 521–539.

Katon, W., Russo, J., Richardson, L., McCauley, E., & Lozano, P. (2008). Anxiety and depression screening for youth in a primary care population. *Ambulatory Pediatrics*, *8*(3), 182–188. https://doi.org/10.1016/j.ambp.2008.01.003

Kent, L., Vostanis, P., & Feehan, C. (1997). Detection of major and minor depression in children and adolescents: Evaluation of the mood and feelings questionnaire. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *38*(5), 565–573. https://doi.org/10.1111/j.1469-7610.1997.tb01543.x

Kovacs, M. (1996). Presentation and course of major depressive disorder during childhood and later years of the life span. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(6), 705–715. https://doi.org/10.1097/00004583-199606000-00010

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kuo, E. S., Stoep, A. V., & Stewart, D. G. (2005). Using the short mood and feelings questionnaire to detect depression in detained adolescents. *Assessment*, *12*(4), 374–383. https://doi.org/10.1177/1073191105279984

Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1998). Major depressive disorder in older adolescents: Prevalence, risk factors, and clinical implications. *Clinical Psychology Review*, *18*(7), 765–794. https://doi.org/10.1016/S0272-7358(98)00010-5

Mayes, T. L., Bernstein, I. H., Haley, C. L., Kennard B. D., & Emslie G. J. (2010). Psychometric properties of the children's depression rating scale-revised in adolescents. *Journal of Child and Adolescent Psychopharmacology*, *20*(6), 513–516. https://doi.org/10.1089/cap.2010.0063

McKenzie, D. P., Toumbourou, J. W., Forbes, A. B., Mackinnon, A. J., McMorris, B. J., Catalano, R. F., & Patton, G. C. (2011). Predicting future depression in adolescents using the Short Mood and Feelings Questionnaire: A two-nation study. *Journal of affective disorders*, *134*(1), 151–159.

Merry, S. N., Stasiak, K., Shepherd, M., Frampton, C., Fleming, T., & Lucassen, M. F. (2012). The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: Randomised controlled non-inferiority trial. *BMJ (Clinical Research Ed.)*, *344*, e2598. https://doi.org/10.1136/bmj.e2598

Murphy, K. R., & Davidshofer, C. O. (2005). *Psychological testing: Principles and applications* (6th ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.

Myers, K., & Winters, N. C. (2002a). Ten-year review of rating scales. I: Overview of scale functioning, psychometric properties, and selection. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*(2), 114–122. https://doi.org/10.1097/00004583-200202000-00004

Myers, K., & Winters, N. C. (2002b). Ten-year review of rating scales. II: Scales for internalizing disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*(6), 634–659. https://doi.org/10.1097/00004583-200206000-00004

National Institute for Health and Clinical Excellence (2005) Depression in children and young people: Identification and management in primary, community and secondary care. Retrieved from http://www.nice.org.uk (CG28)

Patel, N. C., DelBello, M. P., Kowatch, R. A., & Strakowski, S. M. (2006). Preliminary study of relationships among measures of depressive symptoms in adolescents with bipolar disorder. *Journal of Child and Adolescent Psychopharmacology*, *16*(3), 327–335.

Patton, G. C., Olsson, C., Bond, L., Toumbourou, J. W., Carlin, J. B., Hemphill, S. A., & Catalano, R. F. (2008). Predicting female depression across puberty: A two-nation longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*(12), 1424–1432. https://doi.org/10.1097/CHI.0b013e3181886ebe

Pettersson, A., Boström, K. B., Gustavsson, P., & Ekselius L. (2015). Which instruments to support diagnosis of depression have sufficient accuracy? A systematic review. *Nordic Journal of Psychiatry*, *69*(7), 497–508. https://doi.org/10.3109/08039488.2015.1008568

Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and Motor Skills*, *105*(3 I), 997–1014. https://doi.org/10.2466/PMS.105.3.997-1014

Poznanski, E. O., Grossman, J. A., Buchsbaum, Y., Banegas, M., Freeman, L., & Gibbons, R. (1984). Preliminary studies of the reliability and validity of the children's depression rating scale. *Journal of the American Academy of Child Psychiatry*, *23*(2), 191–197. https://doi.org/10.1097/00004583-198403000-00011

Poznanski EO & Mokros HB (1996) *Manual for the children's depression rating scale - revised*. Los Angeles, CA: Western Psychological Services

Rao, U. M. A., Ryan, N. D., Birmaher, B., Dahl, R. E., Williamson, D. E., Kaufman, J., … Nelson, B. (1995). Unipolar depression in adolescents: Clinical outcome in adulthood. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*(5), 566–578. https://doi.org/10.1097/00004583-199505000-00009

Reynolds, W. M. (2002). *Reynolds adolescent depression scale: Professional manual* (2nd ed.). Odessa, FL: Psychological Assessment Resources.

Reynolds, W. M., & Johnston, H. F. (1994). *Handbook of depression in children and adolescents* (Vol. 616). New York: Plenum Press.

Reynolds, W. M., & Stark, K. D. (1987). School-based intervention strategies for the treatment of depression in children and adolescents. *Special Services in the Schools*, *3*(3–4), 69–88. https://doi.org/10.1300/J008v03n03_06

Rhew, I. C., Simpson, K., Tracy, M., Lymp, J., McCauley, E., Tsuang, D., & Vander Stoep, A. (2010). Criterion validity of the short mood and feelings questionnaire and one-and two-item depression screens in young adolescents. *Child and Adolescent Psychiatry and Mental Health*, *4*(1), 8. https://doi.org/10.1186/1753-2000-4-8

Sharp, C., Goodyer, I. M., & Croudace, T. J. (2006). The short mood and feelings questionnaire (SMFQ): A unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7-through 11-year-old children. *Journal of Abnormal Child Psychology*, *34*(3), 379–391. https://doi.org/10.10007/s10802-006-9027-x

Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behavior Research and Therapy*, *36*(5), 545–566. https://doi.org/10.1016/S0005-7967(98)00034-5

Stasiak, K., Parkin, A., Seymour, F., Lambie, I., Crengle, S., Pasene-Mizziebo, E., & Merry, S. (2013). Measuring outcome in child and adolescent mental health services: Consumers' views of measures. *Clinical Child Psychology and Psychiatry*, *18*(4), 519–535.

Streiner, D. L. (1993). A checklist for evaluating the usefulness of rating scales. *Canadian Journal of Psychiatry*, *38*, 140–148.

Sund, A. M., Larsson, B., & Wichstrøm, L. (2001). Depressive symptoms among young Norwegian adolescents as measured by the mood and feelings questionnaire (MFQ). *European Child and Adolescent Psychiatry*, *10*(4), 222–229. https://doi.org/10.1007/s007870170011

Thapar, A., & McGuffin, P. (1998). Validity of the shortened mood and feelings questionnaire in a community sample of children and adolescents: A preliminary research note. *Psychiatry Research*, *81*, 259–268. https://doi.org/10.1016/S0165-1781(98)00073-0

Turner, N., Joinson, C., Peters, T. J., Wiles, N., & Lewis, G. (2014). Validity of the short mood and feelings questionnaire in late adolescence. *Psychological Assessment*, *26*(3), 752–762. https://doi.org/10.1037/a0036572

Walker, L., Merry, S., Watson, P. D., Robinson, E., Crengle, S., & Schaaf, D. (2005). The Reynolds adolescent depression scale in New Zealand adolescents. *Australian and New Zealand Journal of Psychiatry*, *39*(3), 136–140. https://doi.org/10.1080/j.1440-1614.2005.01534.x

Weiss, J. W., Mouttapa, M., Chou, C. P., Nezami, E., Johnson, C. A., Palmer, P. H., … Unger, J. B. (2005). Hostility, depressive symptoms, and smoking in early adolescence. *Journal of Adolescence*, *28*(1), 49–62. https://doi.org/10.1016/j.adolescence.2004.03.009

Wood, A., Kroll, L., Moore, A., & Harrington, R. (1995). Properties of the mood and feelings questionnaire in adolescent psychiatric outpatients: A research note. *Journal of Child Psychology and Psychiatry*, *36*(2), 327–334. https://doi.org/10.1111/j.1469-7610.1995.tb01828.x

Yee, A. M., Algorta, G. P., Youngstrom, E. A., Findling, R. L., Birmaher, B., Fristad, M. A., & LAMS Group. (2015). Unfiltered administration of the YMRS and CDRS-R in a clinical sample of children. *Journal of Clinical Child & Adolescent Psychology*, *44*(6), 992–1007.

Zou, K. H., O'Malley, A. J., & Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, *115*(5), 654–657. https://doi.org/10.1161/CIRCULATIONAHA.105.594929

How to cite this article: Thabrew H, Stasiak K, Bavin L-M, Frampton C, Merry S. Validation of the Mood and Feelings Questionnaire (MFQ) and Short Mood and Feelings Questionnaire (SMFQ) in New Zealand help-seeking adolescents. *Int J Methods Psychiatr Res*. 2018;27:e1610. https://doi.org/10.1002/mpr.1610