American Psychological Association

Psychological Assessment

1040-3590/19/$12.00

2019, Vol. 31, No. 1, 114–119
http://dx.doi.org/10.1037/pas0000640

BRIEF REPORT

Borderline Personality Features Scale for Children-11: Measurement Invariance Over Time and Across Gender in a Community Sample of Adolescents

Salome Vanwoerden, Lorra Garey, and Tayler Ferguson
University of Houston

Jeff R. Temple
University of Texas Medical Branch

Carla Sharp
University of Houston

The Borderline Personality Features Scale for Children (BPFS-C) was recently shortened using item response theory to an 11-item version that is optimal for use in epidemiological studies and repeated assessment over time. Only 1 study has examined invariance of the BPFS-C-11 items across gender and no study has done so over time. The present study employed a longitudinal design to address this gap by evaluating measurement invariance across gender and over time during the transition into adulthood in a diverse community-based sample of 755 adolescents (56% female). Results indicated measurement variance for items measuring personal relationships and impulsivity/recklessness, with females having a greater probability of endorsing items regarding relationship instability and males more likely to endorse impulsivity, despite an equal position on the latent trait. Overall, there was partial measurement invariance of a single dimension of borderline features between males and females and full longitudinal invariance of this factor through the transition into young adulthood. The current findings provide empirical evidence supporting the reliability of BPFS-C-11 scores as a measure of borderline pathology (BP) during late adolescence and early adulthood.

Public Significance Statement
This study suggests that the reports on the Borderline Personality Features Scale for Children function similarly over time from mid to late adolescence and across gender with the exception of different endorsement of items relating to interpersonal functioning and impulsivity/recklessness. This study contributes to a growing body of evidence suggesting that BP can be measured reliably in the transition into adulthood.

Keywords: Borderline Personality Features Scale for Children, measurement, adolescence

Borderline personality disorder (BPD) is a debilitating illness characterized by pervasive instability in interpersonal relationships, identity, behavior, and affect (American Psychiatric Association, 2013). There is a lack of studies using rigorous methodology to evaluate the gender and longitudinal invariance of scores on measures developed for youth (Chanen et al., 2008). Relevant to this goal is the principle of heterotypic continuity, which is when symptoms of the same disorder manifest differently over time despite level of the latent trait. Understanding heterotypic continuity is important because the majority of our understanding

This article was published Online First August 6, 2018.

Salome Vanwoerden, Lorra Garey, and Tayler Ferguson, Department of Psychology, University of Houston; Jeff R. Temple, Department of Behavioral Health and Research, University of Texas Medical Branch; Carla Sharp, Department of Psychology, University of Houston.

This research was supported, in part, by Award K23HD059916 (PI: Jeff R. Temple) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) and 2012-WG-BX-0005 (PI: Jeff R. Temple) from the National Institute of Justice (NIJ). The content is solely the responsibility of the authors and does not necessarily represent the official views of NICHD or NIJ. The Borderline Personality Features Scale for Children-11 (BPFS-C-11): Measurement invariance over time and across gender in a community sample of adolescents.

Correspondence concerning this article should be addressed to Carla Sharp, Department of Psychology, University of Houston, 3695 Cullen Boulevard Room 126, Houston, TX 77204-5022. E-mail: csharp2@central.uh.edu

about mental disorders in youth is based on downward extensions of research conducted among adults (Rutter, 2008), which is especially true for borderline pathology (BP; Sharp & Tackett, 2014).

There are only three *DSM–IV* (BPD criterion were unchanged in *DSM–5* Section II) based self-report measures of BP developed specifically for youth: the Personality Assessment Inventory for Adolescents Borderline Scale (PAI-A-BOR; Morey, 2007), the Borderline Personality Features Scale for Children (BPFS-C; Crick, Murray–Close, & Woods, 2005), and the 11-item version of the BPFS-C (BPFS-C-11; Sharp, Steinberg, Temple, & Newlin, 2014). The BPFS-C-11 was developed using item response theory analyses and has been validated in two separate samples showing good criterion validity of test score interpretations, internal consistency, and test–retest reliability of scores (Fossati, Sharp, Borroni, & Somma, 2016; Sharp, Steinberg et al., 2014). However, the potential of the BPFS-C-11 as a brief measure for repeated assessments to track symptom change at the population level or through intervention remains unclear. Measurement invariance is a method for evaluating whether a test functions in the same way across varied conditions when they are irrelevant to the underlying attribute (Millsap, 2011). Determining invariance across conditions ensures that when differences are found in test performance, these result from true differences on the attribute being measured and not as a function of measurement bias. It also can shed light on whether BP presents in a consistent way over time and across gender more generally. By doing so, the issue of whether heterotypic continuity exists for BP can be evaluated.

Although previous studies conducted in early to midadolescence have found longitudinal invariance of measures of BP (Haltigan & Vaillancourt, 2016; Wright, Zalewski, Hallquist, Hipwell, & Stepp, 2016), no studies have evaluated this for *DSM*-based measures of BP in the transition from adolescence to young adulthood, specifically, and no study has tested longitudinal measurement invariance of the BPFS-C-11. This is significant given that certain features of BPD mirror typical maturational changes through adolescence including emotion regulation and social cognition (Blakemore, 2008). Therefore, it is necessary to determine the utility of measures through this transition. Moreover, few studies have evaluated invariance across gender in youth. Among adults, it has been found that it may be easier for males to endorse anger and impulsivity (Aggen, Neale, Røysamb, Reichborn-Kjennerud, & Kendler, 2009; Sharp, Michonski et al., 2014) and that females more easily endorse self-harm and suicidality (Hoertel, Peyre, Wall, Limosin, & Blanco, 2014) and affective instability (Aggen et al., 2009). The single study that examined gender invariance of the BPFS-C-11 in adolescents found invariance of scores on the BPFS-C-11 in adolescents age 14–20 (Fossati et al., 2016). However, using the full BPFS-C, Haltigan and Vaillancourt (2016) found that during early adolescence, females were more likely to endorse items of interpersonal and intrapersonal dysfunction, despite level on the latent trait; and items related to affective instability and impulsivity were variant across males and females.

To this end, the aim of the current study was to evaluate measurement invariance of BPFS-C-11 scores in a large community-based sample of adolescents. Specifically, we were interested in evaluating longitudinal measurement invariance from late adolescence into young adulthood (ages 16–19) as well as measurement invariance across gender. We expected to find partial measurement invariance across genders for items focusing on impulsivity and relationship functioning, given that the BPFS-C-11 does not include any items evaluating self-harm and suicidality. While one previous study conducted in early to midadolescence found longitudinal invariance of a youth measure of BP (Haltigan & Vaillancourt, 2016), we expected to find partial longitudinal invariance because the transition into young adulthood is a significant developmental transition marked by change in environment (i.e., out of high school), change in peer and family relationships, and identity development (Blakemore & Mills, 2014).

Method

Participants were 781 adolescents who participated in an ongoing longitudinal study (Temple, Shorey, Tortolero, Wolfe, & Stuart, 2013). Adolescents were recruited from seven public schools. Study recruitment and assessment occurred during school hours in classes with mandated attendance when participants were in the ninth or tenth grade. Assessment continued annually for 4 years. All students present in the selected classes were eligible to participate and there was a response rate of 62%. At the time of recruitment, the sample had a mean age of 15.84 years (SD = .68), was 56% female, and identified their race as White (64.7%), African American (23.8%), Asian/Pacific Islander (3.8%), and Other/Mixed Race (7.7%); 50% of White individuals identified as Hispanic. Twenty-six participants did not complete assessments across all years and were therefore excluded from the sample; these participants were significantly older (M_{age} = 15.19, SD = 0.80) compared with those who completed (M_{age} = 14.83, SD = 0.67; $t(26.23)$ = 26.23, p = .03), but did not differ in gender (χ^2 = .34, p = .56). Looking at each year of the study, missing data was variable. In Year 1, n = 727 (93%) participated (M_{age} = 15.81, SD_{age} = 0.67; 55% female); in Year 2, n = 670 (86%) participated (M_{age} = 16.80, SD_{age} = 0.65; 56% female); in Year 3, n = 618 (79%) participated (M_{age} = 17.77, SD_{age} = 0.65; 58% female); in Year 4, n = 507 (65%) participated (M_{age} = 18.79, SD_{age} = 0.65; 63% female). In Year 3, missing data was greater among males, $\chi^2(1, N = 755) = 4.61$, p = .03. In Years 2–3, those with missing data were older (Year 2: $t(753) = 2.84$, p = .01; year 3: $t(753) = 4.94$, $p < .001$). The final sample consisted of 755 adolescents (M_{age} = 15.83, SD_{age} = .67, 56% female).

The current study was approved by the appropriate institutional review board. Recruitment occurred during regular school hours. Research staff presented the study to students and answered any questions, and take-home packets with study information and parental consent forms were sent home with students. Students who returned with parental consent provided assent and completed assessments during school hours. Participants were compensated with $10 (Years 1–3) and $20 (Year 4) gift cards for participating.

The Borderline Personality Feature Scale for Children–11 (BPFS-C-11; Sharp, Steinberg, et al., 2014) is an 11-item measure of borderline personality features for children 9- to 18-years-old. The BPFS-C-11 includes indicators of BP such as affective instability, identity problems, and negative relationships. Items responses are on a 5-point Likert scale ranging from *not true at all* to *always true*. Studies have demonstrated construct validity of BPFS-C-11 test score interpretations via positive relations with other measures of BPD and positive relations with measures of correlates of BPD including emotion dysregulation (Sharp, Stein-

berg et al., 2014). In the current sample, Cronbach's alpha was .86, .85, .86, and .90 in Years 1–4, respectively.

Item level distributions were examined for normality and it was found that distributions of items were within the normal range (skewness range: |0.08|–|1.23|; kurtosis range: |0.08|–|1.17|). Thus, we used maximum likelihood (ML) estimation to account for missing data. Table 1 lists the items and levels of endorsement within the full sample and within each gender separately. Fit of each model was examined using multiple fit indices (Kline, 2011). Specifically, the root mean square error of approximation (RMSEA) was examined, with values of less than .08 indicating reasonable fit and values above .10 suggesting poor fit (Browne & Cudeck, 1993). Second, the comparative fit index (CFI; Bentler, 1990) was also used, with values between 0.95 and 1.00 indicating excellent fit and values between .90 and .95 indicating acceptable fit (Hu & Bentler, 1999). Third, the standardized root-mean-square residual (SRMR) was evaluated, with values less than .08 indicating acceptable fit (Hu & Bentler, 1999).

First, measurement invariance was examined across gender at each year of the study, taking the form of a hierarchical set of multigroup CFAs, in which each subsequent model imposed additional constraints of equality across gender. The baseline model tested configural invariance to test whether the factor structure of BP scores was the same for boys and girls. Next, metric invariance was tested to evaluate whether the magnitude of factor loadings were equal across gender. Finally, scalar invariance was tested to evaluate whether item intercepts were equal across gender. Based on gender invariance results, longitudinal invariance was examined in the full sample (if full gender invariance was demonstrated) and/or within each gender separately (if full invariance was not found). To test longitudinal invariance, a hierarchical set of models with increasing levels of equality constraints over time within persons were evaluated. Configural invariance was first tested to determine whether the same factor and pattern of factor loadings existed over time. Next, metric invariance was tested to determine whether factor loadings were equal over time. Finally, scalar invariance was tested to examine whether intercepts were equal over time. When necessary, empirically and theoretically warranted modifications were evaluated and employed iteratively to improve model fit (Steenkamp & Baumgartner, 2000), which is common in cases of constraining several model parameters.

The χ^2 difference test was employed to assess differences in model fit. A nonsignificant difference in model comparison indicating model invariance for the more constrained model (Anderson & Gerbing, 1988). Considering χ^2 difference tests are susceptible to similar problems as the χ^2, including sample size dependency (Kline, 2011), additional fit indices were used to evaluate difference in model fit. Specifically, CFI change of less than 0.010 and RMSEA change of less than 0.015 (Chen, 2007) provided statistical evidence for invariance between the less constrained and more constrained model; for instances wherein a discrepancy for invariance was observed across evaluative statistics, CFI change and RMSEA change were considered more accurate tests for model comparison and used as primary indicators of invariance.

Results

A one-factor model at baseline yielded good model fit, $\chi^2(55) = 1875.20$, $p < .001$; RMSEA = .06, 90% CI [.05, .07]; CFI = .93; SRMR = .04. Standardized loadings ranged from .44 (Item 11) to .72 (Item 7) with an average standardized loading of .59 across all items. Gender invariance of BPFS-C-11 scores were conducted (see Table 2) to determine whether longitudinal invariance analyses should be conducted for each gender separately or within the full sample. At each of the four years, the single factor structure of the BPFS-C-11 demonstrated configural invariance across genders. Next, metric invariance was demonstrated across genders.

The model of scalar invariance at each year fit significantly worse than the metric invariance model. Modification indices were examined separately at each year and were tested one-by-one until the resulting partial scalar model did not fit significantly worse than the metric model. Overall and relative fit for each of these models are presented in Table 2 along with brief item content for each item that was freed. Modification indices for analyses across all years showed that Item 20 (*Lots of times, my friends and I are really mean to each other*) was the source of misfit and should be allowed to vary across genders. Item intercepts that were freed differed across the 4 years—in Year 1, four items were freed; in

Table 1
BPFS-C-11 Items and Levels of Endorsement

Item description	Male		Female		Full sample	
	Mean (*SD*)	Range	Mean (*SD*)	Range	Mean (*SD*)	Range
1. (2) I feel very lonely.	2.18 (1.09)	1–5	2.40 (1.15)	1–5	2.30 (1.13)	1–5
2. (6) I want to let people know they've hurt me.	2.47 (1.20)	1–5	3.13 (1.28)	1–5	2.84 (1.29)	1–5
3. (8) My feelings are very strong.	2.73 (1.26)	1–5	3.10 (1.24)	1–5	2.94 (1.26)	1–5
4. (9) There is something important missing.	2.70 (1.31)	1–5	2.96 (1.36)	1–5	2.84 (1.34)	1–5
5. (11) I'm careless with things that are important.	1.97 (1.04)	1–5	1.96 (1.12)	1–5	1.96 (1.08)	1–5
6. (13) People who were close have let me down.	2.42 (1.21)	1–5	3.04 (1.22)	1–5	2.77 (1.25)	1–5
7. (14) I go back and forth between feelings.	2.21 (1.18)	1–5	2.69 (1.26)	1–5	2.48 (1.25)	1–5
8. (15) I get into trouble . . . do things without thinking.	2.37 (1.16)	1–5	2.46 (1.27)	1–5	2.42 (1.23)	1–5
9. (16) I worry that people I care about will leave.	2.31 (1.28)	1–5	2.83 (1.41)	1–5	2.60 (1.38)	1–5
10. (18) How I feel about myself changes a lot.	2.17 (1.13)	1–5	2.50 (1.17)	1–5	2.35 (1.17)	1–5
11. (20) My friends and I are mean to each other.	1.98 (1.06)	1–5	1.81 (.90)	1–5	1.89 (.98)	1–5

Note. BPFS-C = Borderline Personality Features Scale for Children. Item content is abbreviated for the purpose of space; full item content is available from the corresponding author by request. Numbers in parentheses are the item numbers corresponding to the full 24-item BPFS-C and is how they are referred to in the article.

Table 2
Measurement Invariance of the BPFS-C-11 Across Gender

Models	χ^2	*df*	CFI	RMSEA	ΔCFI	ΔRMSEA	$\Delta\chi^2$	Δ*df*	*p*
Year 1									
1. Configural	263.83	88	.923	.074	—	—	—	—	—
2. Metric	281.38	98	.920	.072	.003	.002	8.78	10	.55
3. Scalar	381.38	108	.880	.083	.040	.011	50.00	10	<.001
3a. Partial scalar (20. *My friends and I are mean*)	346.48	107	.895	.078	.025	.006	65.10	9	<.001
3b. Partial scalar (20, 6. *Let people know*)	326.40	106	.903	.075	.017	.003	45.02	8	<.01
3c. Partial scalar (20, 6, 13. *People let me down*)	311.13	105	.910	.073	.010	.001	29.75	7	.04
Year 2									
1. Configural	277.45	88	.911	.080	—	—	—	—	—
2. Metric	285.31	98	.912	.075	.001	.005	3.93	10	.95
3. Scalar	418.29	108	.854	.092	.058	.017	66.49	10	<.001
3a. Partial scalar (20. *My friends and I are mean*)	386.01	107	.868	.088	.044	.013	100.70	9	<.001
3b. Partial scalar (20, 11. *I'm careless with things*)	361.77	106	.879	.085	.033	.010	76.46	8	<.001
3c. Partial scalar (20, 11, 15. *Do things without thinking*)	335.57	105	.891	.081	.021	.006	50.26	7	<.001
3d. Partial scalar (20, 11, 15, 9. *Something missing*)	314.71	104	.901	.078	.011	.003	29.40	6	.02
3e. Partial scalar (20, 11, 15, 9, 6. *Let people know*)	303.11	103	.906	.076	.006	.001	17.80	5	.11
Year 3									
1. Configural	252.01	88	.922	.077	—	—	—	—	—
2. Metric	262.18	98	.922	.073	.000	.004	5.09	10	.89
3. Scalar	347.01	108	.887	.084	.035	.011	42.42	10	<.001
3a. Partial scalar (20. *My friends and I are mean*)	325.90	107	.897	.081	.025	.008	63.72	9	<.001
3b. Partial scalar (20, 13. *People let me down*)	307.59	106	.905	.078	.017	.005	45.41	8	<.01
3c. Partial scalar (20, 13, 6. *Let people know*)	294.16	105	.911	.076	.011	.003	31.98	7	.03
3d. Partial scalar (30, 13, 6, 11. *I'm careless with things*)	283.29	104	.915	.074	.007	.001	21.11	6	.10
Year 4									
1. Configural	313.82	88	.914	.101	—	—	—	—	—
2. Metric	325.92	98	.913	.096	.001	.005	6.05	10	.81
3. Scalar	428.12	108	.878	.108	.035	.012	51.10	10	<.001
3a. Partial scalar (20. *My friends and I are mean*)	389.53	107	.893	.102	.020	.006	63.62	9	<.001
3b. Partial scalar (20, 15. *Do things without thinking*)	368.21	106	.900	.099	.013	.003	42.29	8	<.01
3c. Partial scalar (20, 13, 11. *I'm careless with things*)	350.03	105	.907	.096	.006	.000	24.11	7	.10

Note. BPFS-C = Borderline Personality Features Scale for Children; *df* = degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root-mean-square residual; ΔCFI = change in comparative fit index from the less restrictive model (change statistics reported for both Models 3a, 3b, 3c, 3d, 3e are compared with Model 2); ΔRMSEA = change in root mean square error of approximation from the less restrictive model (change statistics reported for both Models 3a, 3b, 3c, 3d, 3e are compared with Model 2); $\Delta\chi^2$ = change in χ^2 from previous model the less restrictive model (change statistics reported for both Models 3a, 3b, 3c, 3d, 3e are compared with Model 2); Δ*df* = change in degrees of freedom from previous model the less restrictive model (change statistics reported for both Models 3a, 3b, 3c, 3d, 3e are compared with Model 2).

Year 2, five items were freed; in Year 3, four items were freed; in Year 4, three items were freed. Items 11 and 5 were freed across all 4 years and Items 2, 6, and 8 were each freed at two different years. Thus, the BPFS-C-11 structure demonstrated partial measurement invariance after freeing five of the 11 item intercepts that were largely related to relationship functioning and impulsivity as hypothesized.

Considering that so few items demonstrated measurement invariance, and those that did differed across years, we concluded that it would be appropriate to evaluate longitudinal invariance for girls and boys separately. Indeed, to examine them as one group would introduce unnecessary bias into the findings (Vandenberg & Lance, 2000). Results of longitudinal invariance, separated by gender, are displayed in Table 3. Among females, the single factor structure demonstrated configural invariance over time. Next, the factor structure demonstrated metric invariance at the item level. When evaluating scalar invariance of BPFS-C-11 items, results of the χ^2 difference tests were significant; however, change in CFI and RMSEA were below stated cut points. Given the limitations of χ^2 in a sample size as large as the current sample, we concluded that scores demonstrated scalar invariance over time. Thus, the BPFS-C-11 structure demonstrated full measurement invariance over time among females. Among males, the single factor structure of the BPFS-C-11 demonstrated configural, metric, and scalar invariance of the items across time. Thus, the BPFS-C-11 structure demonstrated full measurement invariance over time among males. Lastly, findings were replicated in the full sample because despite partial measurement invariance across gender, we concluded that examining measurement invariance within the full sample may provide clinically useful data that could be used to support extant findings of the BPFS-C-11 observed over time. Specifically, the single factor structure demonstrated configural, metric, and scalar invariance over time. Similar to the longitudinal invariance models tested among females, analyses of scalar invariance in the full sample yielded a significant χ^2 difference test; however, change in CFI and RMSEA were below stated cut points, and was therefore concluded to indicate full scalar invariance.

Discussion

The objective of the current study was to examine the invariance of the BPFS-C-11 across gender and over time from late adolescence to young adulthood in a community sample. Results demonstrated partial invariance across genders at each time point and

Table 3
Measurement Invariance of the BPFS-C-11 Across Age

Models	χ^2	*df*	CFI	RMSEA	ΔCFI	ΔRMSEA	$\Delta\chi^2$	Δ*df*	*p*
Females									
1. Configural	1328.22	830	.927	.038	—	—	—	—	—
2. Metric	1357.52	860	.927	.037	.000	.001	14.65	30	.99
3. Scalar	1449.75	890	.918	.039	.009	.002	46.11	30	.03
Males									
1. Configural	1338.64	830	.912	.043	—	—	—	—	—
2. Metric	1369.65	860	.912	.042	.000	.001	15.00	30	.99
3. Scalar	1437.82	890	.905	.043	.007	.001	34.09	30	.28
Full sample									
1. Configural	1660.42	830	.934	.036	—	—	—	—	—
2. Metric	1688.03	860	.935	.036	.001	.000	13.81	30	.99
3. Scalar	1809.42	890	.927	.037	.008	.001	60.70	30	.001

Note. BPFS-C = Borderline Personality Features Scale for Children; *df* = degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root-mean-square residual; ΔCFI = change in comparative fit index from previous model; ΔRMSEA = change in root mean square error of approximation from previous model; $\Delta\chi^2$ = change in χ^2 from previous model; Δ*df* = change in degrees of freedom from previous model.

full longitudinal invariance. The current findings add to a growing body of evidence for the reliability of BPFS-C-11 scores. Further, in support of the construct validity of measurement of BP in adolescence, the current study suggests that BP, as operationalized with the BPFS-C-11 retains its internal structure through late adolescence. We can conclude that the BPFS-C-11 can be used to compare levels of BPD features across this developmental phase; however, examinations of mean differences between genders should be tempered by knowledge of differential likelihood of endorsement of items relating to relationships and impulsivity despite levels of underlying BP.

As hypothesized, instability in relationships were more likely to be endorsed by females regardless of levels of BP whereas items relating to impulsivity were more likely to be endorsed by males. Interestingly, males were more likely to endorse an item relating to interpersonal aggression (*Lots of times, my friends and I are really mean to each other*) than females with the same underlying level of BP. The pattern of items that demonstrated variance across genders was largely consistent across the four years of the current study. These findings are in line with research of gender differences in impulsivity (Cross, Copping, & Campbell, 2011) and relationship processes (Rose & Rudolph, 2006). Implications of these findings are important for assessing BP and suggest that measurement bias may have, in part, contributed to gender differences observed in the literature.

The value of evaluating longitudinal measurement invariance is twofold. First, demonstrated measurement invariance over time provides robust evidence to support that mean changes in some construct over time are due to actual changes in an individual's level of a latent construct rather than an artifact of the measurement tool (Millsap, 2011). Second, it provides information about whether the construct can be measured in the same way at different ages. Previous research has found invariance of an adult measure of BP: The International Personality Disorders Examination in a high-risk sample of girls between the ages of 14–17 (Wright et al., 2016) and the full BPFS-C-24 has demonstrated partial measurement variance across early adolescence (Haltigan & Vaillancourt, 2016). Therefore, these studies largely support the prospective use of the IPDE and BPFS-C-24 in evaluating BP. Based on the present evidence, the BPFS-C-11 can now similarly be used to evaluate change in BP as adolescents age into adulthood. Together, this research is important against the background of research finding that levels of borderline features increase until mid-adolescence and then level out through adulthood (Wright et al., 2016). The current study (alongside other invariance studies) suggests that these trajectories are reflective of true (mean level) changes pointing to possible homotopic continuity of *DSM*-based borderline symptoms through adolescence.

The current study has several limitations. First, it was found that adolescents with missing data on various years tended to be male, older, and report lower BPFS-C-11 scores. Altogether, given some suggestions of response bias, these findings must be interpreted with caution and further replication is necessary. Second, the way we conceptualized BP in the current study was based on *DSM–IV* and *DSM–5* Section II formulations. Taking a maladaptive trait perspective in examining invariance beyond the confines of a single PD is an important avenue for future research. Third, the BPFS-C-11 does not include any items evaluating self-harm or suicidality. This is a limitation of this measure given the centrality of self-harm and suicidal behaviors and ideation in the presentation of BP (Soloff, Lis, Kelly, Cornelius, & Ulrich, 1994). Finally, the findings of the current study do not overlap completely with other studies of gender invariance for the measurement of borderline features in adolescence. Specifically, one study found gender invariance of BPFS-C-11 scores among community adolescents in Italy. However, the authors did note that a model in which equality of factor loadings and thresholds across genders were relaxed fit the data significantly better, fit was not markedly different, which is why gender invariance was concluded (Fossati et al., 2016). It is possible that cultural or language differences contributed to these differences; however, findings must be replicated.

Despite limitations, these findings offer important considerations for the assessment of BP in youth. Based on results, it is advisable that clinicians and scientists should consider gender bias present in certain items related to BP. Additional work may also investigate whether modeling the invariance of items showing

gender bias may increase reliability of the borderline diagnosis across time. Considering that much of the work using the BPFS-C-11 has included males and females together, future work should include separate recommendations (e.g., clinical cutoffs) for each gender. Altogether, the current study extends previous work on the BPFS-C-11 by providing evidence for psychometric properties. Additionally, the current study adds to existing support of the validity of the measurement of borderline construct in adolescence.

References

Aggen, S. H., Neale, M. C., Røysamb, E., Reichborn-Kjennerud, T., & Kendler, K. S. (2009). A psychometric evaluation of the *DSM–IV* borderline personality disorder criteria: Age and sex moderation of criterion functioning. *Psychological Medicine, 39,* 1967–1978. http://dx.doi.org/10.1017/S0033291709005807

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association Publishing.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 411–423. http://dx.doi.org/10.1037/0033-2909.103.3.411

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246. http://dx.doi.org/10.1037/0033-2909.107.2.238

Blakemore, S.-J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience, 9,* 267–277. http://dx.doi.org/10.1038/nrn2353

Blakemore, S.-J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology, 65,* 187–207. http://dx.doi.org/10.1146/annurev-psych-010213-115202

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Chanen, A. M., Jovev, M., Djaja, D., McDougall, E., Yuen, H. P., Rawlings, D., & Jackson, H. J. (2008). Screening for borderline personality disorder in outpatient youth. *Journal of Personality Disorders, 22,* 353–364. http://dx.doi.org/10.1521/pedi.2008.22.4.353

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464–504. http://dx.doi.org/10.1080/10705510701301834

Crick, N. R., Murray–Close, D., & Woods, K. (2005). Borderline personality features in childhood: A short-term longitudinal study. *Development and Psychopathology, 17,* 1051–1070. http://dx.doi.org/10.1017/S0954579405050492

Cross, C. P., Copping, L. T., & Campbell, A. (2011). Sex differences in impulsivity: A meta-analysis. *Psychological Bulletin, 137,* 97–130. http://dx.doi.org/10.1037/a0021591

Fossati, A., Sharp, C., Borroni, S., & Somma, A. (2016). Psychometric properties of the Borderline Personality Features Scale for Children-11 (BPFSC-11) in a sample of community dwelling Italian adolescents. *European Journal of Psychological Assessment*. Advance online publication. http://dx.doi.org/10.1027/1015-5759/a000377

Haltigan, J. D., & Vaillancourt, T. (2016). The Borderline Personality Features Scale for Children (BPFS-C): Factor structure and measurement invariance across time and sex in a community-based sample. *Journal of Psychopathology and Behavioral Assessment, 38,* 600–614. http://dx.doi.org/10.1007/s10862-016-9550-1

Hoertel, N., Peyre, H., Wall, M. M., Limosin, F., & Blanco, C. (2014). Examining sex differences in *DSM–IV* borderline personality disorder symptom expression using item response theory (IRT). *Journal of Psychiatric Research, 59,* 213–219. http://dx.doi.org/10.1016/j.jpsychires.2014.08.019

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. 2011 (3rd ed.). New York, NY: Guilford Press.

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY: Routledge.

Morey, L. C. (2007). *The personality assessment inventory—Adolescent professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin, 132,* 98–131. http://dx.doi.org/10.1037/0033-2909.132.1.98

Rutter, M. (2008). Developing concepts in developmental psychopathology. In J. Hudziak (Ed.) *Developmental psychopathology and wellness: Genetic and environmental influences* (pp. 3–22). Washington, DC: American Psychiatric Association Publishing.

Sharp, C., Michonski, J., Steinberg, L., Fowler, J. C., Frueh, B. C., & Oldham, J. M. (2014). An investigation of differential item functioning across gender of BPD criteria. *Journal of Abnormal Psychology, 123,* 231–236. http://dx.doi.org/10.1037/a0035637

Sharp, C., Steinberg, L., Temple, J., & Newlin, E. (2014). An 11-item measure to assess borderline traits in adolescents: Refinement of the BPFSC using IRT. *Personality Disorders, 5,* 70–78. http://dx.doi.org/10.1037/per0000057

Sharp, C., & Tackett, J. L. (2014). Introduction: An idea whose time has come. In C. Sharp & J. L. Tackett (Eds.), *Handbook of borderline personality disorder in children and adolescents* (pp. 3–8). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4939-0591-1_1

Soloff, P. H., Lis, J. A., Kelly, T., Cornelius, J., & Ulrich, R. (1994). Self-mutilation and suicidal behavior in borderline personality disorder. *Journal of Personality Disorders, 8,* 257–267. http://dx.doi.org/10.1521/pedi.1994.8.4.257

Steenkamp, J.-B. E. M., & Baumgartner, H. (2000). On the use of structural equation models for marketing modeling. *International Journal of Research in Marketing, 17,* 195–202. http://dx.doi.org/10.1016/S0167-8116(00)00016-1

Temple, J. R., Shorey, R. C., Tortolero, S. R., Wolfe, D. A., & Stuart, G. L. (2013). Importance of gender and attitudes about violence in the relationship between exposure to interparental violence and the perpetration of teen dating violence. *Child Abuse & Neglect, 37,* 343–352. http://dx.doi.org/10.1016/j.chiabu.2013.02.001

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3,* 4–70. http://dx.doi.org/10.1177/109442810031002

Wright, A. G. C., Zalewski, M., Hallquist, M. N., Hipwell, A. E., & Stepp, S. D. (2016). Developmental trajectories of borderline personality disorder symptoms and psychosocial functioning in adolescence. *Journal of Personality Disorders, 30,* 351–372. http://dx.doi.org/10.1521/pedi_2015_29_200

Received December 6, 2017
Revision received May 2, 2018
Accepted June 14, 2018 ■