Journal of Counseling Psychology
2014, Vol. 61, No. 4, 641–648

0022-0167/14/$12.00 http://dx.doi.org/10.1037/cou0000040

BRIEF REPORT

Classification Challenges in Perfectionism

Kenneth G. Rice
Georgia State University

Clarissa M. E. Richardson
University of Florida

High performance expectations are central to perfectionism, but because most participants endorse high standards, it becomes difficult for practitioners and researchers to accurately screen for perfectionists. We addressed problems linked to the measurement and classification of perfectionism by testing various strategies aimed at broadening the range and skew of scores on the Standards subscale from the Almost Perfect Scale—Revised (APS–R; Slaney, Mobley, Trippi, Ashby, & Johnson, 1996). Randomly assigned participants ($N = 506$) completed the APS–R following standard instructions or 1 of 2 variations, one prompting participants to consider their responses in light of a normal distribution of scores and another in which participants used a visual analog (slider) scale. The visual analog scale produced more differentiated scores, but range restrictions and skewed distributions remained for all 3 variations. Statistical transformations improved skew. Factor mixture modeling was conducted using transformed and nontransformed perfectionism scores along with criterion indicators of emotion regulation (reappraisal or suppression), perceived stress, and depression. Results supported a 3-class model, although more balanced distributions of classes emerged than were previously reported. Perfectionists were differentiated from nonperfectionists by their higher standards scores. Maladaptive perfectionists scored highest among the classes on most self-critical perfectionism indicators, suppression, perceived stress, and depression. Adaptive perfectionists had the lowest levels of perceived stress and depression and scored highest on reappraisal. Both perfectionist classes had generally comparable concerns about mistakes, but criterion indicators suggested those were more problematic for maladaptive perfectionists. Results supported the value of incorporating adaptive and maladaptive criterion indicators in classification models.

Keywords: perfectionism, factor mixture modeling, emotion regulation, stress, depression

Several models of perfectionism emphasize standards or standard setting as not only a key element of perfectionism but a key focus or target of counseling interventions (Lo & Abbott, 2013). Ostensibly, the argument is that standards are unrealistically high among (maladaptive) perfectionists, so a reasonable course of action would be to help such perfectionists lower their standards (Egan, Piek, Dyck, Rees, & Hagger, 2013). Others for whom high standards appears psychologically healthy (e.g., those with low self-criticism) might be encouraged to maintain their standards. Given the centrality of standards to conceptualization and intervention, it is surprising that fundamental concerns with assessing standards have not been addressed. Our efforts in this article are aimed at addressing that gap, with specific attention paid to the scores obtained from the Almost Perfect Scale—Revised (APS–R; Slaney et al., 1996).

The APS–R measures self-performance expectations (Standards subscale) and self-critical evaluation of one's ability to meet expected standards (Discrepancy subscale). When reported, scores on the APS–R Standards subscale have been substantially range restricted and negatively skewed (e.g., Rice & Ashby, 2007; Rice, Richardson, & Tueller, 2014). Strikingly few respondents have low Standards scores. The rounded average item response score on the APS–R Standards subscale is 6 on a 7-point agreement scale (Rice & Ashby, 2007; Rice et al., 2014); approximately 2% of respondents have low Standards scores (average item responses on the *disagree* end of the scale) and 88%–91% have high Standards scores (in the *agree* range; K. G. Rice, personal communication, May 20, 2014). Although range restrictions and skew may not be unusual in psychology, when about 90% of a sample scores in the high range, classification approaches may be logically less compelling and empirically limited in detecting bona fide latent classes (Bauer & Curran, 2004).

In the current study, we used an experimental design to evaluate whether response ranges could be expanded and skew attenuated through the use of different scaling options and instructional prompts. We also implemented different visual (i.e., normal curve

This article was published Online First August 11, 2014.

Kenneth G. Rice, Department of Counseling and Psychological Services, Georgia State University; Clarissa M. E. Richardson, Department of Psychology, University of Florida.

We are grateful to Angela Montfort and Marieke van Nuenen for their help with this article.

Correspondence concerning this article should be addressed to Kenneth G. Rice, Department of Counseling and Psychological Services, P.O. Box 3980, Georgia State University, Atlanta, GA 30302-3980. E-mail: kgr1@gsu.edu

distribution) and response format (i.e., visual analog scale) conditions. We then examined whether perfectionism profiles observed in recent studies (e.g., Herman, Trotter, Reinke, & Ialongo, 2011; Rice, Lopez, & Richardson, 2013) can be replicated if the distribution and range of Standards scores can be expanded. If reliable profiles could be obtained, we also sought to evaluate associations between class structure outcomes of interest to counseling psychologists, such as emotion regulation (Richardson, Rice, & Devine, 2014), perceived stress (Chang, Watkins, & Banks, 2004), and depression (Sherry, Mackinnon, Macneil, & Fitzpatrick, 2013) while covarying effects of Conscientiousness and Neuroticism (Dunkley, Blankstein, & Berg, 2012; Rice, Ashby, & Slaney, 2007). Furthermore, testing differences among classes on these other variables would provide further information about the nature of the classes.

Method

Participants

Participants were 506 undergraduate students (386 women, 106 men, 14 missing gender data) attending a southeastern U.S. university. Ages ranged from 18 to 32 years (M = 19.49 years, SD = 1.67). The largest ethnic and racial groups were non-Hispanic Whites (60%), followed by Hispanic Whites (11.3%), non-Hispanic Blacks or African Americans (10.7%), and non-Hispanic Asians (7.8%), with the remaining groups comprising 3.5% or less of the sample.

Measures

Two subscales, Standards and Discrepancy, were used from the APS–R (Slaney et al., 1996, 2001). Items on the APS–R use a 7-point scale anchored by 1 = *strongly disagree* and 7 = *strongly agree*. Score reliability, convergent validity, and criterion-related validity have been supported in several studies (e.g., Rice & Ashby, 2007; Slaney et al., 2001).

The Frost Multidimensional Perfectionism Scale (FMPS; Frost et al., 1990) is a self-report measure that uses a 5-point scale (1 = *strongly disagree* to 5 = *strongly agree*). Score reliability and convergent and criterion validity have been supported (Frost et al., 1990). Consistent with other studies (Boone et al., 2010; Lee & Park, 2011), we limited analyses to the Personal Standards, Concerns Over Mistakes, and Doubts About Actions subscales.

The Positive and Negative Self-Oriented Performance Perfectionism subscales from the Performance Perfectionism Scale (PPS; Chang, 2006) were used. Items responses use a 5-point scale (1 = *extremely untrue of me* through 5 = *extremely true of me*). Chang (2006) reported good score reliability, along with support for convergent and criterion-related validity.

The Emotional Regulation Questionnaire (ERQ; Gross & John, 2003) contains two subscales, Reappraisal (changing a situation's meaning to change its emotional effects) and Suppression (inhibiting negative thoughts and feelings). Items are responded to using a 7-point scale (1 = *strongly disagree* through 7 = *strongly agree*). Adequate internal consistency and score validity estimates have been reported (Gross & John, 2003).

The Perceived Stress Scale (PSS; Cohen, Kamarck, Mermelstein, 1983) uses a 5-point response scale (0 = *never* to 4 = *very often*) and reflects frequency of stress over the past month. Good score reliability estimates and evidence of validity have been reported (Cohen et al., 1983).

On the Center for Epidemiologic Studies Depression scale (CES–D; Radloff, 1977), respondents rate the frequency with which they experienced depressive symptoms during the previous week (0 = *rarely or none of the time* through 3 = *most or all of the time*). There is good evidence supporting score reliability and validity for CES–D scores (Radloff, 1991).

Conscientiousness and Neuroticism from the Mini-International Personality Item Pool (Mini-IPIP; Donnellan, Oswald, Baird, & Lucas, 2006) were used as covariates. Mini-IPIP items use a 5-point scale (1 = *very inaccurate* through 5 = *very accurate*). Donnellan et al. (2006) reported adequate reliability and criterion-related validity estimates for both scores.

Procedure

The study was approved by the university's institutional review board, and all participants provided informed consent to participate. Using Qualtrics survey software, participants were randomly assigned to APS–R completion procedures and one of three different prompts: typical instructions, normal curve (consider responses while viewing a normal distribution of typical scores), or visual analog (slider) for item responses. Randomization resulted in 170 participants receiving typical APS–R instructions, 170 receiving the normal curve condition, and 166 receiving the visual analog condition. After completing the APS–R, participants completed the FMPS, PPS, ERQ, PSS, CESD, and Mini-IPIP in randomized order. To detect careless responding, we embedded three items within the questionnaires in which respondents were directed to endorse a specific response option, and 506 participants passed all three items.

Results

Different Administrations of the APS–R

Mean differences. Table 1 displays means, standard deviations, and reliability estimates for scores obtained from the different conditions. Most of the score internal consistency estimates were in the .80–.90 range. There was a significant main effect in the analysis of Discrepancy scores ($p < .0001$) but not for Standards scores ($p = .920$). Post hoc tests (Games–Howell) indicated that participants in the visual analog condition had significantly lower Discrepancy scores compared with participants in the other two conditions (ds = 0.48 and 0.32). Neuroticism scores were also lower ($p = .037$) in the visual analog condition compared with typical instructions ($d = 0.29$).

Score ranges and distributions. Standards scores ranged from 2.86 to 7.00 (typical instructions), 2.14 to 7.00 (normal curve), and 2.29 to 7.00 (visual analog). Approximately 2% of participants receiving typical instructions and 4% of those in the normal curve or visual analog conditions had Standards scores that were less than the scale midpoint of 4, and approximately 90% in each condition had scores greater than 4. Range restriction was similar across conditions, although finer grained item responses in the visual analog condition yielded 4 times as many different

Table 1
Descriptive Statistics for Three Conditions Involving Different Administrations of the Almost Perfect Scale—Revised

Subscale	Typical instructions	Normal curve	Visual analog	F	Partial η^2
Standards					
M	6.07	6.03	6.04	0.08	<.0001
SD	0.73	0.82	0.88		
α	.87	.87	.86		
Discrepancy					
M	3.67_a	3.46_a	3.03_b	9.82^{**}	.04
SD	1.28	1.26	1.40		
α	.94	.94	.95		
Concern Over Mistakes					
M	2.72	2.53	2.61	2.10	.008
SD	0.82	0.85	0.83		
α	.89	.90	.89		
Personal Standards					
M	3.64	3.66	3.65	0.03	<.0001
SD	0.66	0.69	0.66		
α	.81	.80	.79		
Doubt About Actions					
M	2.69	2.56	2.49	2.55	.01
SD	0.84	0.89	0.82		
α	.79	.81	.75		
Positive Self-Oriented Perfectionism					
M	3.80	3.85	3.85	0.21	.001
SD	0.78	0.89	0.86		
α	.76	.87	.88		
Negative Self-Oriented Perfectionism					
M	1.93	1.89	1.88	0.19	.001
SD	0.84	0.80	0.85		
α	.86	.83	.85		
Conscientiousness					
M	3.52	3.61	3.53	0.67	.003
SD	0.75	0.84	0.79		
α	.70	.78	.73		
Neuroticism					
M	2.87_a	2.77_{ab}	2.63_b	3.33^{*}	.013
SD	0.80	0.88	0.84		
α	.67	.74	.72		
Reappraisal					
M	4.81	4.89	4.89	0.36	.001
SD	1.04	1.06	1.07		
α	.86	.88	.85		
Suppression					
M	3.66	3.37	3.58	2.49	.01
SD	1.23	1.28	1.26		
α	.80	.80	.80		
Perceived Stress					
M	27.65	26.77	26.26	1.26	.005
SD	8.30	8.28	7.81		
α	.88	.86	.85		
Depression					
M	17.93	16.27	15.73	1.82	.007
SD	12.11	10.86	10.08		
α	.93	.92	.90		

Note. Rows not sharing the same subscript significantly differ at $p < .05$.
* $p < .05$. ** $p < .0001$.

scores (109 different scores) as the other two conditions (24 and 25 different scores).

Negative skew was extreme and significant ($p < .0001$, two-tailed test) for Standards scores obtained in the typical instructions, normal curve, and visual analog conditions, zs $= -6.20$, -7.68, and -8.25, respectively. Skew for Standards was substantial in each condition and was worse with the nontypical administrations. Skew was not substantial for the Discrepancy scores obtained in the typical instructions and normal curve conditions (zs = 1.06 and 1.89, respectively; $p > .05$, two-tailed test) but was substantial for the visual analog condition ($z = 4.25$, $p < .0001$). The Kolmogorov–Smirnov test revealed significant nonnormality for Standards scores in all three conditions (K–S = .117, .126, and .137, ps < .0001) and in the typical instructions and visual analog conditions for Discrepancy scores (K–S = .093, $p = .001$, and .112, $p < .0001$, respectively) but not for Discrepancy scores in the normal curve condition (K–S = .064, $p = .082$).

Although not part of the manipulation, scores on the FMPS Personal Standards scale were also negatively skewed ($z = -3.61$, $p = .0003$), as were scores for PPS Positive Self-Oriented Perfectionism ($z = -7.94$, $p < .0001$); Negative Self-Oriented Perfectionism scores were positively skewed ($z = 9.49$, $p < .0001$), as were Concern Over Mistakes scores ($z = 3.13$, $p = .0018$). These score distributions were also significantly nonnormal: K–S ranged from .063 to .176 (ps < .0001).

Reflecting and logarithmic transformations (Tabachnick & Fidell, 2007) improved skew for Standards scores in the typical instructions and normal curve administrations, $z = -1.03$, $p = .302$, and $z = -1.76$, $p = .078$, respectively; after transformation, some skew remained in visual analog scores, $z = -2.67$, $p = .008$. Log-transformed Standards (lgSTD) scores were multiplied by -1.0 to facilitate interpretation; the correlation between lgSTD and nontransformed Standards score was $r = .97$ in each condition.

Because the visual analog condition transformed scores were still substantially skewed, subsequent analyses were based on participants from the typical instructions and normal curve conditions ($N = 340$). Log transformations, with or without reflection, were also used to transform substantially skewed scores from the other perfectionism subscales. Discrepancy, Concern Over Mistakes, and Doubts About Actions scores were not transformed; z values for skew for those scores were 2.19, 2.39, 0.83, respectively, ps > .015.

Factor Mixture Modeling

We used factor mixture modeling (FMM) to simultaneously analyze two dimensions of perfectionism (Perfectionistic Standards and Self-Critical Perfectionism) and potentially multiple categories or latent classes of perfectionists and nonperfectionists. Conscientiousness and Neuroticism were covariates to improve accuracy (Lubke & Muthén, 2007). Fit was evaluated with the Bayesian information criterion (BIC), the sample-size adjusted BIC (aBIC), the Lo–Mendell–Rubin (LMR) likelihood ratio test, and the bootstrap likelihood ratio test (BLRT). Relative entropy was used to evaluate classification adequacy.

We initially fitted a two-factor, single-class baseline (purely dimensional) model. Standardized factor loadings for the three log-transformed indicators of Perfectionistic Standards ranged from .79 to .80. For Self-Critical Perfectionism, loadings ranged from .51 to .83. The correlation between the two perfectionism factors was .11, $p = .257$. Next, two-, three-, four-, and five-class models were fitted to each of the comparison models (Lubke & Muthén, 2005): (a) nonvariant with freely estimated intercepts across classes; (b) Perfectionistic Standards partially invariant, that is, class-invariant intercepts for Self-Critical Perfectionism and class-specific (freely estimated) intercepts for Perfectionistic Standards; (c) Self-Critical Perfectionism partially invariant, that is, class-invariant Perfectionistic Standards intercepts and class-specific Self-Critical Perfectionism intercepts; and (d) fully invariant Perfectionistic Standards and Self-Critical Perfectionism intercepts. To ensure the same constructs were evaluated across the classes, we constrained factor loadings to be class-invariant in all models, as were residuals. Table 2 displays fit statistics for these models.

Fit indices and entropy were generally supportive of Model 3 (freely estimated Self-Critical Perfectionism indicator intercepts) and the three-class model. Transition matrices revealed reasonable stability of classes until the four-class model. About 80% of Class 3 members in the three-class solution had been grouped into Class 1 in the two-class results. Similarly, about 76% of Class 1 members in the three-class results had been in Class 2 in the two-class model. Although 84% of those who had been in Class 2 in the three-class model were grouped in Class 4 in the four-class model, the other two classes from the three-class model were substantially more dispersed in the four-class arrangement. Class 1 from the three-class model was spread to Class 1 (28%), Class 2 (48%), and Class 4 (23%). Likewise, 47% of those who had been in Class 3 (three-class model) were distributed to Class 2 and 44% were clustered into Class 3 in the four-class model. We also analyzed Model 3 using nontransformed scores and found substantial differences. Although one class size proportion was comparable in size (19%) between results based on transformed or nontransformed scores, the largest class proportion was 67% when identified with nontransformed scores versus 44% when based on transformed scores. Likewise, one class represented 11% when based on nontransformed scores, but a third class when based on transformed scores represented 37%. Model 3's BIC (5,016.45) and aBIC (4,845.15) values were substantially higher than the BIC and aBIC values of the model based on transformed scores. The LMR and BLRT results were less clear on supporting the three-class solution, with p values of .0680 (LMR) and < .0001 (BLRT). Entropy was lower (.748) compared to the transformed-scores results (.929). In short, these indicators seemed meaningfully worse than those obtained with transformed scores.

The Auxiliary (*e*) option in Mplus revealed several mean differences between the classes on the criterion indicators for emotion regulation (Reappraisal, Suppression), Perceived Stress, and depression (CES–D). Thus, these criterion indicators were modeled as outcomes, resulting in the scores being regressed on class. This approach of directly incorporating adaptive and maladaptive criteria into class structure can change structure (Muthén, 2004), but a benefit is that results can help refine or clarify classes and their proper labeling.

Final model fit was good: BIC = 7,470.42, aBIC = 7,248.37, and LMR ($p = .014$) and BLRT ($p < .0001$) again indicated a significant effect for the three-class structure over the two-class structure (a four-class version was less supported: BIC = 7,484.02, aBIC = 7,204.86, a nonsignificant LMR [$p = .199$], and a significant BLRT [$p < .001$]). Table 3 displays class counts and proportions, means for the Perfectionistic Standards factor and Self-Critical Perfectionism indicator intercepts, and descriptive statistics. On the basis of posterior probabilities, the first (estimated $n = 92$), second (estimated $n = 119$), and third (estimated $n = 130$) classes accounted for an estimated 27%, 35%, and 38% of the population, respectively. The Perfectionistic Standards factor mean for one class (labeled *nonperfectionists*) was lower than the means for the

Table 2
Fit Indices and Entropy for One- to Five-Class Factor Mixture Models

Model	BIC	aBIC	LMR *p*	BLRT *p*	Entropy
Baseline single class					
0. Unstructured means	779.13	696.66			
Two classes					
1. Noninvariant	745.74	612.50	.018	<.0001	0.699
2. Partially invariant (PS intercepts)	765.39	641.67	.150	<.0001	0.693
3. Partially invariant (SCP intercepts)	772.72	645.83	.158	<.0001	0.618
4. Fully invariant	811.34	700.31	.530	.375	0.775
Three classes					
1. Noninvariant	730.60	546.61	.136	<.0001	0.742
2. Partially invariant (PS intercepts)	810.27	645.32	.351	.667	0.625
3. Partially invariant (SCP intercepts)	745.58	574.28	.000	<.0001	0.929
4. Fully invariant	848.68	709.10	.462	1.000	0.846
Four classes					
1. Noninvariant	775.02	540.27	.633	.217	0.723
2. Partially invariant (PS intercepts)	843.56	637.37	.805	.600	0.737
3. Partially invariant (SCP intercepts)[a]	774.51	558.80	.170	.200	0.892
4. Fully invariant	862.92	694.79	.678	.030	0.747
Five classes					
1. Noninvariant	802.62	517.13	.749	.082	0.771
2. Partially invariant (PS intercepts)[b]					
3. Partially invariant (SCP intercepts)	823.68	563.56	.880	.667	0.899
4. Fully invariant	904.16	707.49	.759	1.000	0.693

Note. PS intercepts = noninvariant intercepts for Perfectionistic Standards indicators; SCP intercepts = noninvariant intercepts for Self-Critical Perfectionism indicators; BIC = Bayesian information criterion; aBIC = Sample-adjusted BIC; LMR = Lo–Mendell–Rubin Test; BLRT = Bootstrap likelihood ratio test.
[a] Despite several increases in random starts, the best log-likelihood value for the generated data was not replicated in most of the bootstrap draws. [b] Failed to converge.

other two classes, and there was not a statistically significant difference between those two classes. Several Self-Critical Perfectionism indicators were higher in one class than in the others, as were scores on Perceived Stress and CES–D, suggesting that class could be labeled *maladaptive perfectionists*. Lower Self-Critical Perfectionism indicators and the lowest scores on Perceived Stress and CES–D were evident for the remaining class, labeled *adaptive perfectionists*. It is interesting that Concern Over Mistakes scores were comparable for both maladaptive and adaptive perfectionists. Reappraisal scores were highest for adaptive perfectionists, followed by nonperfectionists, then maladaptive perfectionists. Suppression scores were comparable for the maladaptive and nonperfectionists, and lowest for adaptive perfectionists. Compared with adaptive perfectionists, the maladaptive and nonperfectionists were likely to have lower Conscientiousness scores ($B = -1.00$, $SE = 0.24$, and $B = -0.82$, $SE = 0.24$, respectively) and higher Neuroticism scores ($B = 3.11$, $SE = 0.48$, and $B = 1.51$, $SE = 0.40$, respectively; $ps < .001$). In an alternative parameterization for comparison, the maladaptive and nonperfectionists were also unlikely to differ on Conscientiousness ($B = 0.18$, $SE = 0.27$), $p = .501$), but the nonperfectionists were likely to have lower Neuroticism scores compared with the maladaptive perfectionists ($B = -1.61$, $SE = 0.29$, $p < .001$).

Discussion

Endorsing high personal performance standards or expectations is a core feature of perfectionism (Lo & Abbott, 2013; Slaney et al., 2002), but because of item response range restrictions and skewed score distributions, endorsing high standards represents a potentially problematic dimension to assess and use in practice, especially if standards scores are intended to identify perfectionists, reveal patterns of association with other variables, or serve as a target for intervention (Egan et al., 2012, 2013; Lo & Abbott, 2013; Stoeber & Hotham, 2013). The present study raised additional concerns about the impenetrability of self-reported standards, which appeared immune to attempts to shift response styles. One implication is that future research might consider less reliance on self-report and instead incorporate other reports or implicit measures to gauge perfectionism (De Cuyper, Pieters, Claes, Vandromme, & Hermans, 2013).

Another possibility is that restricted response tendencies for some indicators may be the result of developmental or learned progressions in self-ratings. In one of the few studies of latent class structure among early adolescents, Herman et al. (2011) reported no substantial skew in their perfectionism scores in a study of sixth graders. Unfortunately, perfectionism was not measured in their follow-up when the children were in 12th grade (Herman, Wang, Trotter, Reinke, & Ialongo, 2013), therefore, whether score distributions changed or played a role in their growth mixture modeling results could not be addressed (Bauer & Curran, 2003). In other samples, score range and skew have not typically been addressed, but descriptive statistics for early to midadolescents and high school–age adolescents suggest lower overall average scores on APS–R Standards compared with older college student samples

Table 3
Final Three-Class Mixture Model Incorporating Criterion Outcome Indicators

Class	Class proportion	Reappraisal		Suppression		Perceived stress		Depression		Factors	
		M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	PS	SCP
1. Maladaptive	.27	4.39	0.14	3.81	0.18	36.72	0.65	32.06	1.53	−0.16	DS: 4.06 CM: 2.46 DA: 2.99 NS: 0.31
2. Nonperfectionists	.35	3.85	0.12	3.65	0.13	28.29	1.16	15.71	1.48	−0.31*	DS: 3.08 CM: 1.61 DA: 2.20 NS: 0.25
3. Adaptive	.38	5.17	0.09	3.17	0.13	19.57	0.86	7.78	0.70	0	DS: 2.97 CM: 2.76 DA: 2.28 NS: 0.15

Note. Results are based on factor mixture models that assumed weak measurement invariance (invariant factor loadings) across classes. PS = Perfectionistic Standards; SCP = Self-Critical Perfectionism; DS = Discrepancy; CM = Concern Over Mistakes; DA = Doubts About Actions; NS = log-transformed Negative Self-Oriented Perfectionism. Intercepts for SCP were based on indicators with different scale ranges. DS range = 1–7; CM and DA range = 1–5; NS range = 0–.70.
* p = .023.

(e.g., Gilman & Ashby, 2003; Rice, Ashby, & Gilman, 2011). Perhaps standards scores in late adolescent and young adult U.S. samples represent a developmental pattern of increasing skew from childhood or early adolescence through adulthood, one that corresponds with increasing academic pressures to perform at a high level or more general societal expectations to be the best or endorse an enhanced self-image (Twenge, Konrath, Foster, Campbell, & Bushman, 2008). Thus, future studies should address whether there is a selection effect involving higher performing students in college versus non-college-attending young adults, whether a developmental trend is affecting score distributions, and whether skew is more likely with homogenous indicators of pure personal standards (e.g., APS–R Standards) than it is with more heterogenous items tapping a similar dimension (e.g., Self-Oriented Perfectionism; see Cox, Enns, & Clara, 2002).

Using transformed scores in mixture models, we were able to replicate classes reported in several other studies (e.g., Richardson et al., 2014), although structure and interpretation of those classes were different from past research. Incorporating associations with criterion indicators produced a more balanced distribution of three latent classes than previously had been observed; in recent studies, nearly 8 out of 10 participants were classified as perfectionists (Rice et al., 2013; Richardson et al., 2014). Samples of college students very likely contain a high proportion of high-achieving individuals and could therefore be expected to net higher proportions of perfectionists as a result of selection effects. Although plausible, it also seems likely that the high numbers of perfectionists in those samples might have been the result of misclassification attributable to the distribution and range restriction of standards scores (Bauer & Curran, 2004). Alternatively, we may have underestimated the number of perfectionists in the current study as a result of a conservative or novel approach to class identification. Applying the approach taken here to past and future perfectionism data sets might help reconcile this issue.

Maladaptive perfectionists in this study had very high levels of stress and depression, consistent with other research (Sherry et al., 2013). Conversely, adaptive perfectionists were the least stressed and depressed group and, for emotion regulation, least likely to use suppression and most likely to use reappraisal. It is interesting that adaptive and maladaptive perfectionists were more comparable than different on Concern Over Mistakes scores. Other results suggested that the maladaptive perfectionists may be more adversely affected by their standards and concerns about mistakes than are adaptive perfectionists. That is, both kinds of perfectionists are more concerned about making mistakes than are nonperfectionists, but adaptive perfectionists may use those concerns to their advantage rather than disadvantage, as suggested by their low levels of psychological distress. This distinction has not emerged in cluster-analytic studies involving perfectionism measures, perhaps because in those studies, neither mixture modeling nor other variables were part of the statistical formation of the groups (e.g., Lee & Park, 2011; Rice & Mirzadeh, 2000). Reactions to mistakes may be another mechanism for reconsidering the relative adaptiveness and maladaptiveness of perfectionism (Gotwals, Stoeber, Dunn, & Stoll, 2012).

The current study has several limitations. Results were based on a large sample and experimental design, but the sample was also restricted by geographic region and university setting. Some of the major findings were consistent with other studies using the same instruments, so it seems less likely that limited breadth of representation in the current study substantially impaired generalizability. Nevertheless, future research should determine whether similar latent profiles are upheld with data derived from diverse samples in the United States and other countries.

Another limitation in the current study involves the manipulations used to alter score characteristics. After the typical administration condition, the other conditions represented changes in more than one potential variable that could affect score ranges, such as combinations of response scale modifications and prompts for decentering and perspective-taking. One component might have worsened score patterns, whereas another might have improved them. Future work might examine such manipulations using a focused, dismantling approach that tests one revision to the measure at a time. Additional work could also test variations implemented with different measures of standards to evaluate the scale specificity of the current findings.

Future research might help resolve the inconsistencies between the current study and prior work by incorporating different criterion variables, such as performance indicators (e.g., Gotwals et al., 2012; Stoeber & Eysenck, 2008) and measures of resilience and problem solving (Dunkley, Ma, Lee, Preacher, & Zuroff, 2014), to determine which categorical approach best accounts for adaptive or maladaptive responses in challenging, coping-oriented situations. A practical issue also needs to be addressed. Logarithmic score transformations make scoring an otherwise simple scale into a more complicated and less clinically friendly procedure. Score transformations may also have traded one problem (score distributions) for another problem in that interval-level data for log-transformed scores would be less tenable. Other future work might examine alternative presentations or scaling of the APS–R and other perfectionism measures to refine a user-friendly approach to assessment and classification of perfectionism. Future work can also reexamine cutoff scores used to classify perfectionists and nonperfectionists (e.g., Rice & Ashby, 2007) because skewed distributions may have resulted in equations yielding overclassification of perfectionists. A synthesis of user friendly approaches to scoring and classification with empirically defensible scores subjected to mixture analyses seems warranted. Finally, and perhaps even more fundamentally, our results could be interpreted as a call for the selection or development of indicators that better differentiate the majority of respondents who score at the upper ends of standards measures. Item response theory could be used to help identify such indicators.

Our results suggest that arguments regarding the relative adaptiveness of perfectionism (Gotwals et al., 2012) might be adjudicated by directly incorporating criterion indicators within latent profile modeling approaches. Whether perfectionism is adaptive or not likely depends on directly linking the classification to criterion indicators when conceptually or empirically feasible to do so. Cut scores can be useful for clinical, research, or other classification purposes, but cutoff scores for perfectionism do not appear to have incorporated the idea that personality characteristics have relative value or cost given the cultural, clinical, or empirical context within which they are examined (McCrae, 2013). The differentiation of adaptive from maladaptive perfectionism may require more direct consideration of such contexts. The individual perfectionism subscales or items themselves cannot be expected to make such determinations.

To summarize, our findings point to the importance of addressing limitations in measures of personal standards and the use of rigorous and integrative statistical approaches for the classification of perfectionists. Future studies could combine several of the ideas and findings from the current study to evaluate stability of latent classes as well as distal performance, behavioral, or interpersonal outcomes predicted by proximal perfectionism class structure built from concurrent indicators of perfectionism, emotion regulation, and stress reactivity.

References

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods, 8,* 338–363. doi:10.1037/1082-989X.8.3.338

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods, 9,* 3–29. doi:10.1037/1082-989X.9.1.3

Boone, L., Soenens, B., Braet, C., & Goossens, L. (2010). An empirical typology of perfectionism in early-to-mid adolescents and its relation with eating disorder symptoms. *Behaviour Research and Therapy, 48,* 686–691. doi:10.1016/j.brat.2010.03.022

Chang, E. (2006). Conceptualization and measurement of adaptive and maladaptive aspects of performance perfectionism: Relations to personality, psychological functioning, and academic achievement. *Cognitive Therapy and Research, 30,* 677–697. doi:10.1007/s10608-006-9060-7

Chang, E. C., Watkins, A. F., & Banks, K. (2004). How adaptive and maladaptive perfectionism relate to positive and negative psychological functioning: Testing a stress-mediation model in Black and White female college students. *Journal of Counseling Psychology, 51,* 93–102. doi:10.1037/0022-0167.51.1.93

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24,* 385–396. doi:10.2307/2136404

Cox, B. J., Enns, M. W., & Clara, I. P. (2002). The multidimensional structure of perfectionism in clinically distressed and college student samples. *Psychological Assessment, 14,* 365–373. doi:10.1037/1040-3590.14.3.365

De Cuyper, K., Pieters, G., Claes, L., Vandromme, H., & Hermans, D. (2013). Indirect measurement of perfectionism: Construct and predictive validity. *Journal of Social and Clinical Psychology, 32,* 844–858. doi:10.1521/jscp.2013.32.8.844

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP Scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18,* 192–203. doi:10.1037/1040-3590.18.2.192

Dunkley, D. M., Blankstein, K. R., & Berg, J. (2012). Perfectionism dimensions and the five-factor model of personality. *European Journal of Personality, 26,* 233–244. doi:10.1002/per.829

Dunkley, D. M., Ma, D., Lee, I. A., Preacher, K. J., & Zuroff, D. C. (2014). Advancing complex explanatory conceptualizations of daily negative and positive affect: Trigger and maintenance coping action patterns. *Journal of Counseling Psychology, 61,* 93–109. doi:10.1037/a0034673

Egan, S. J., Dick, M., & Allen, P. J. (2012). An experimental investigation of standard setting in clinical perfectionism. *Behaviour Change, 29,* 183–195. doi:10.1017/bec.2012.16

Egan, S. J., Piek, J. P., Dyck, M. J., Rees, C. S., & Hagger, M. S. (2013). A clinical investigation of motivation to change standards and cognitions about failure in perfectionism. *Behavioural and Cognitive Psychotherapy, 41,* 565–578. doi:10.1017/S135246581200063X

Frost, R. O., Marten, P. A., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Therapy and Research, 14,* 449–468. doi:10.1007/BF01172967

Gilman, R., & Ashby, J. S. (2003). A first study of perfectionism and multidimensional life satisfaction among adolescents. *The Journal of Early Adolescence*, *23*, 218–235. doi:10.1177/0272431603023002005

Gotwals, J. K., Stoeber, J., Dunn, J. H., & Stoll, O. (2012). Are perfectionistic strivings in sport adaptive? A systematic review of confirmatory, contradictory, and mixed evidence. *Canadian Psychology/Psychologie Canadienne*, *53,* 263–279. doi:10.1037/a0030288

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85,* 348–362. doi:10.1037/0022-3514.85.2.348

Herman, K. C., Trotter, R., Reinke, W. M., & Ialongo, N. (2011). Developmental origins of perfectionism among African American youth. *Journal of Counseling Psychology, 58,* 321–334. doi:10.1037/a0023108

Herman, K. C., Wang, K., Trotter, R., Reinke, W. M., & Ialongo, N. (2013). Developmental trajectories of maladaptive perfectionism among African American adolescents. *Child Development, 84,* 1633–1650. doi:10.1111/cdev.12078

Lee, D., & Park, H. (2011). Cross-cultural validity of the Frost Multidimensional Perfectionism Scale in Korea. *The Counseling Psychologist, 39,* 320–345. doi:10.1177/0011000010365910

Lo, A., & Abbott, M. (2013). The impact of manipulating expected standards of performance for adaptive, maladaptive, and non-perfectionists. *Cognitive Therapy and Research, 37,* 762–778. doi:10.1007/s10608-013-9528-1

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10,* 21–39. doi:10.1037/1082-989X.10.1.21

Lubke, G., & Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling, 14,* 26–47. doi:10.1207/s15328007sem1401_2

McCrae, R. R. (2013). Exploring trait assessment of samples, persons, and cultures. *Journal of Personality Assessment, 95,* 556–570. doi:10.1080/00223891.2013.821075

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *The SAGE handbook of quantitative methodology for the social sciences* (pp. 346–370). Newbury Park, CA: Sage doi:10.4135/9781412986311.n19

Radloff, L. S. (1977). The CES–D scale: A self-report depression scale for research in a general population. *Applied Psychological Measurement, 1,* 385–401. doi:10.1177/014662167700100306

Radloff, L. S. (1991). The use of the Center for Epidemiologic Studies Depression Scale in adolescents and young adults. *Journal of Youth and Adolescence, 20,* 149–166. doi:10.1007/BF01537606

Rice, K. G., & Ashby, J. S. (2007). An efficient method for classifying perfectionists. *Journal of Counseling Psychology, 54,* 72–85. doi:10.1037/0022-0167.54.1.72

Rice, K. G., Ashby, J. S., & Gilman, R. (2011). Classifying adolescent perfectionists. *Psychological Assessment, 23,* 563–577. doi:10.1037/a0022482

Rice, K. G., Ashby, J. S., & Slaney, R. B. (2007). Perfectionism and the five factor model of personality. *Assessment, 14,* 385–398. doi:10.1177/1073191107303217

Rice, K. G., Lopez, F. G., & Richardson, C. E. (2013). Perfectionism and performance among STEM students. *Journal of Vocational Behavior, 82,* 124–134. doi:10.1016/j.jvb.2012.12.002

Rice, K. G., & Mirzadeh, S. A. (2000). Perfectionism, attachment, and adjustment. *Journal of Counseling Psychology, 47,* 238–250. doi:10.1037/0022-0167.47.2.238

Rice, K. G., Richardson, C. M. E., & Tueller, S. (2014). The short form of the revised Almost Perfect Scale. *Journal of Personality Assessment, 96,* 368–379. doi:10.1080/00223891.2013.838172

Richardson, C. M. E., Rice, K. G., & Devine, D. P. (2014). Perfectionism, emotion regulation, and the cortisol stress response. *Journal of Counseling Psychology, 61,* 110–118. doi:10.1037/a0034446

Sherry, S. B., Mackinnon, S. P., Macneil, M. A., & Fitzpatrick, S. (2013). Discrepancies confer vulnerability to depressive symptoms: A three-wave longitudinal study. *Journal of Counseling Psychology, 60,* 112–126. doi:10.1037/a0030439

Slaney, R. B., Mobley, M., Trippi, J., Ashby, J., & Johnson, D. G. (1996). *The Almost Perfect Scale—Revised*. Unpublished manuscript, Pennsylvania State University, University Park, PA.

Slaney, R. B., Rice, K. G., & Ashby, J. S. (2002). A programmatic approach to measuring perfectionism: The Almost Perfect Scales. In G. L. Flett & P. L. Hewitt (Eds.), *Perfectionism: Theory, research, and treatment* (pp. 63–88). Washington, DC: American Psychological Association. doi:10.1037/10458-003

Slaney, R. B., Rice, K. G., Mobley, M., Trippi, J., & Ashby, J. S. (2001). The revised Almost Perfect Scale. *Measurement and Evaluation in Counseling and Development, 34,* 130–145.

Stoeber, J., & Eysenck, M. W. (2008). Perfectionism and efficiency: Accuracy, response bias, and invested time in proof-reading performance. *Journal of Research in Personality, 42,* 1673–1678. doi:10.1016/j.jrp.2008.08.001

Stoeber, J., & Hotham, S. (2013). Perfectionism and social desirability: Students report increased perfectionism to create a positive impression. *Personality and Individual Differences, 55,* 626–629. doi:10.1016/j.paid.2013.04.023

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston, MA: Allyn & Bacon.

Twenge, J. M., Konrath, S., Foster, J. D., Campbell, W., & Bushman, B. J. (2008). Further evidence of an increase in narcissism among college students. *Journal of Personality, 76,* 919–928. doi:10.1111/j.1467-6494.2008.00509.x

Received March 22, 2014
Revision received May 25, 2014
Accepted May 27, 2014 ■