Please cite as:

Flett, G. L., Mara, C. A., Hewitt, P. L., Sirois, F., & Molnar, D. S. (2016). How Should Discrepancy Be Assessed in Perfectionism Research? A Psychometric Analysis and Proposed Refinement of the Almost Perfect Scale–Revised. *Journal of Psychoeducational Assessment*. doi:10.1177/0734282916651382

How Should Discrepancy be Assessed in Perfectionism Research? A Psychometric Analysis and Proposed Refinement of the Almost Perfect Scale - Revised

Gordon L. Flett
York University

Constance A. Mara
Cincinnati Children's Hospital Medical Center

Paul L. Hewitt
University of British Columbia

Fuschia Sirois
The University of Sheffield

Danielle S. Molnar
Brock University

Abstract

Research on perfectionism with the Almost Perfect Scale-Revised (APS-R) distinguishes adaptive perfectionists versus maladaptive perfectionists based primarily on their responses to the 12-item unidimensional APS-R discrepancy subscale, which assesses the sense of falling short of standards. People described as adaptive perfectionists have high standards but low levels of discrepancy (i.e., relatively close to attaining these standards). Maladaptive perfectionists have perfectionistic high standards and high levels of discrepancy. In the current work, we re-examine the psychometric properties of the APS-R discrepancy subscale and illustrate that this supposedly unidimensional discrepancy measure may actually consists of more than one factor. Psychometric analyses of data from student and community samples distinguished a pure five-item discrepancy factor and a second four-item factor measuring dissatisfaction. The five-item factor is recommended as a brief measure of discrepancy from perfection and the four-item factor is recommended as a measure of dissatisfaction with being imperfect. Overall, our results confirm past suggestions that most people with maladaptive perfectionism are characterized jointly by chronic dissatisfaction as well as a sense of being discrepant due to having fallen short of expectations. These findings are discussed in terms of their implications for the assessment of perfectionism, as well as the implications for research and practice.

How Should Discrepancy be Assessed in Perfectionism Research? A Psychometric Analysis and Proposed Refinement of the Almost Perfect Scale - Revised

The current article re-examines the psychometric characteristics of the Almost Perfect Scale – Revised (APS-R; Slaney, Rice, Mobley, Trippi, & Ashby, 2001), which is a measure with an ironic name given the likelihood that no measure is perfect. While this paper is primarily focused on assessment issues, it stands as an example of how assessment and conceptual issues are often inextricably linked. Our particular focus is the APS-R discrepancy subscale, which is a unique, important, and influential subscale in theory and research on perfectionism (for a discussion, see Flett & Hewitt, 2014). Perfectionism is defined within the framework espoused by Slaney and colleagues as having both positive and negative aspects. Discrepancy largely accounts for the negative element. According to Slaney et al. (2001), discrepancy is "… the central and defining negative aspect of perfectionism" (p. 133). They defined discrepancy as "the perceived discrepancy or difference between the standards one has for oneself and one's actual performance" (p. 133). In a subsequent chapter, Slaney, Rice, and Ashby (2002) defined discrepancy as "… the perception that one consistently fails to meet the high standards one has set for themselves." When perfectionists judge themselves with discrepancy as their key context, perfectionists are distressed primarily because they tend to fall short of extant standards and expectations.

The APS-R discrepancy subscale developed by Slaney and colleagues has 12 items. Why is this discrepancy subscale worth investigating? First, extensive correlational research with this subscale has shown that discrepancy is almost invariably associated with various forms of maladjustment across various types of respondents, including adolescents (Gilman & Ashby,

2003), university students (Paulson & Rutledge, 2014), and clinical patients (Levinson et al., 2015). Second, the discrepancy dimension has been examined both as a vulnerability factor and a key mediator of the association that perfectionism and various negative outcomes (Sherry, MacKinnon, Fossum, Antony, Stewart, Sherry, Nealis, & Mushquash, 2013; Sherry, MacKinnon, Macneil, & Fitzpatrick, 2013). Finally, a key distinction in the literature is the distinction between adaptive perfectionists and maladaptive perfectionists. Adaptive perfectionists are described and conceptualized as people who typically have high standards but low levels of discrepancy on the APS-R discrepancy subscale (i.e., relatively close to attaining these standards); maladaptive perfectionists have both high standards and high levels of discrepancy on the APS-R discrepancy subscale; in other words, maladaptive perfectionists are people who desperately want to be perfect but see themselves as being far from perfect (for a discussion, see Slaney, Rice, & Ashby, 2002). In short, maladaptive perfectionists fall short of expectations and standards, while adaptive perfectionists meet and often exceed standards and expectations.

To our knowledge, the APS-R discrepancy subscale has not been carefully scrutinized in terms of testing the central assumption that it is unifactorial. Our evaluation of the discrepancy subscale suggests that this measure is complex and problematic in some key respects. First, three items among the 12 items include referents to chronic negative affect. These items are "I often feel frustrated because I can't meet my goals," "I often worry about not measuring up to my own expectations," and "I often feel disappointment after completing a task because I know I could have done better." These items with the references to frequent frustration, worry, and disappointment should inflate to some extent the magnitude of the association between discrepancy and measures of chronic or state negative affect. Historically, Nicholls, Licht, and

Pearl (1982) have been credited with drawing attention to the key issue of item content overlap in questionnaires. Subsequent concerns were raised due to personality vulnerability measures with item content that overlaps with adjustment outcome measures (for a discussion, see Flett, Hewitt, Endler, & Bagby, 1993).

Second, we propose on the basis of our inspection of items that it is possible to empirically differentiate a subset of items that clearly tap dissatisfaction (e.g., I am not satisfied even when I know I have done my best) and another subset of items that tap discrepancy in a much purer sense (e.g., I rarely live up to my high standards). Our work is based on the premise that because several APS-R discrepancy items refer openly to dissatisfaction after falling short, it is actually possible to identify a dissatisfaction factor after item analyses.

Why is important to ascertain whether the APS-R has a distinguishable dissatisfaction factor? This is a key issue for both pragmatic and conceptual reasons. The main pragmatic concern was outlined above – because a focus on dissatisfaction is built into certain items on the APS-R discrepancy subscale, the magnitude of obtained associations involving the discrepancy are likely impacted and inflated in many instances. The conceptual issue here is just as important. As we discuss below, an overarching question in the perfectionism literature is the extent to which dissatisfaction is a central element of the perfectionism construct. Here we believe it is important to acknowledge that a tendency to treat dissatisfaction as part of the discrepancy concept is not entirely supportable because it then becomes difficult to account for or allow for some of the most intriguing perfectionists in existence – that is, those relatively successful and highly driven people (e.g., elite athletes and superstar performers) who are never satisfied with their high accomplishments. These people should have low scores on the discrepancy dimension because they usually come close to the elusive standard they are pursing

but they still have profound dissatisfaction with their performance level and with themselves. Previously, we encountered some of these people in a study of elite professional performers (i.e., actors, dancers, and classical musicians). Higher self-oriented perfectionism among these people was associated significantly with *less goal satisfaction* in terms of how they judged their careers despite objective indicators of being highly accomplished (see Mor, Day, Flett, & Hewitt, 1995). Another relevant investigation by Enns, Cox, Sareen, and Freeman (2001) found that perfectionism (both maladaptive and adaptive) was associated with performance dissatisfaction in a sample of medical students and this association was still evident in regression analyses even after controlling for individual differences in actual performance!

It is conceivable that the APS-R scale creators (i.e., Slaney and his colleagues) or other scholars in the field could reasonably contend from a conceptual perspective driven by construct validity concerns that dissatisfaction should actually be part of the discrepancy dimension on the grounds that it is a key element of the perfectionism construct. This claim could be supported at a conceptual level by referring to Hamachek's (1978) suggestion that a central distinction between normal perfectionism and a more pathological form of neurotic perfectionism is the tendency for normal perfectionists to be satisfied with their outcomes and achievement while neurotic perfectionists seldom experience satisfaction and instead experience profound dissatisfaction. Hamachek (1978) had expanded upon prior observations by Missildine (1963) who described at length the chronic dissatisfaction inherent in extreme perfectionism. Missildine (1963) emphasized that many perfectionists are actually quite successful in an objective sense, but at the subjective level, they perceive successes as failures due to their abiding dissatisfaction and sense that they could and should do better.

The alternative perspective is that discrepancy domain should not include a dissatisfaction component for theoretical and applied reasons. At the theoretical level, a case can be made for not including dissatisfaction as part of discrepancy because dissatisfaction has a strong affective, evaluative component, and this affective focus here goes well beyond a sense of failing short of standards. Indeed, dissatisfaction has been interpreted as a negative emotional reaction to imperfection (see Stoeber & Rambow, 2007). It can also be suggested from a conceptual perspective that there is a small but identifiable subset of perfectionists from a person-centered perspective who are highly successful and they are not at all discrepant in the usual sense because they are at or near the standard, yet their drive and determination is fuelled by a sense that they must keep striving and that they can never be satisfied, not even for a moment. Hamachek (1978) observed that these perfectionists operate according to their abiding perception that no performance or effort is ever quite good enough despite typically achieving at a very high level.

Our chief concern about the APS-R discrepancy subscale item content tapping dissatisfaction is that various authors have used the discrepancy subscale as a predictor of related individual differences in various measures of satisfaction or dissatisfaction. It is not very surprising when findings are reported indicating that perfectionists with high levels of discrepancy on the APS-R also have elevations on measures that explicitly tap some component of dissatisfaction such as measures of life satisfaction (see Gilman & Ashby, 2003; Gnilka, Ashby, & Noble, 2013; Grzegorek, Slaney, Franze, & Rice, 2004; Pearson & Gleaves, 2006; Wang, Yuen, & Slaney, 2009).

In the current research, we evaluated our observations by assessing the factor structure of the APS-R across two large samples. Our goal was to make some recommendations about how

the APS-R should be scored and interpreted. Additionally, we provide evidence that the original factor structure proposed by Slaney et al. is less than optimal. We examine four alternative models that were hypothesized *a priori* based on item content and the definitions of perfectionism previously discussed. It was anticipated that the original model utilized by Slaney et al. (see Figure 1) would be the poorest fitting among the models tested. We hypothesized that a model that either removes or accounts for items that measure "negative affect" or "dissatisfaction" would better explain the relations among the items in the APS-R (*i.e.,* the hypothesized models will have a better fit to the data). In Model 1, we exclude items we believe are measuring "negative affect" (see Figure 2). In Model 2, we isolate a factor called "dissatisfaction", composed of four items currently in the "discrepancy" subscale (see Figure 3). For Model 3, we exclude the four "dissatisfaction" items to remain consistent with the original three factors proposed by Slaney et al (see Figure 4).

Method and Analytical Plan

To account for the ordered categorical item responses on the APS-R, confirmatory factor analysis models were fitted to polychoric correlations using the mean- and variance-adjusted weighted least squares (WLSMV) estimator implemented with Mplus 6.1 and its theta parameterization (Muthén & Muthén, 1998-2011). Model fit was evaluated using the root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker–Lewis index (TLI) fit indices. To demonstrate good fit to the data, an estimated model should have a RMSEA of near .06 or less along with either CFI or TLI near or above .95 (Hu & Bentler, 1999). In order to identify each of the models, the variance of each factor was constrained equal to one which set the scale for each of these factors. Factors in all models were permitted to correlate with one another. We ran all the hypothesized models on a second, separate, validation sample in order to

assess the replicability of our initial findings. Once a final factor structure was obtained, reliability analyses were performed to determine the internal consistency (Cronbach's alpha) of the obtained factors.

Sample 1

Descriptive Statistics

Our first sample consisted of 670 undergraduate university students (209 men, 449 women, and 12 undeclared). The average age was 21.09 years (*SD* = 3.55). Participants were recruits from a participant pool at a university located in Western Canada for a broader study that included multiple measures of perfectionism and well-being.

Sample 2

Our second sample consisted of 977 community-dwelling adults (354 women, 622 men, one undeclared) from either the United States or Canada who completed an anonymous online survey. The APS-R item responses were gathered as part of a broader investigation on personality and health by Sirois and Molnar (2014), and Sirois (2015). Overall, 57% of this sample were characterized as being "normal controls" as determined by the lack of a chronic health condition, while the rest of the sample had people that indicated the presence of one or more chronic health conditions on a checklist of 12 health conditions, including chronic fatigue syndrome, arthritis, and fibromyalgia (see Sirois & Molnar, 2014). Their mean age was 32.60 years (*SD* = 9.94), with an age range of 16 to 70 years old.

Measure

Almost Perfect Scale-Revised. The APS-R is a self-report inventory derived from an earlier version (see Johnson & Slaney, 1996; Slaney & Johnson, 1992). Respondents make 7-point Likert-type ratings with options ranging from "strongly disagree" to "strongly agree." The

initial APS-R item pool had 39 items. Item and factor analyses of responses from various samples of participants involving several stages reduced the final version to 23 items (see Slaney et al., 2002). The final APS-R has seven items assessing standards, four items assessing order, and 12 items assessing discrepancy.

Results

Sample 1

Original Model

First we tested the original model on first sample ($N = 669$) to ensure replicability of Slaney et al.'s (2001) results. For this initial model, we did not account for the ordered categorical item responses (*i.e.,* used maximum likelihood estimation in Mplus, not WLSMV) to be directly comparable to previous work on this scale. In the original paper, Slaney et al. reported the following results in their final CFA: χ^2 (227) = 459.22, $p < .05$, RMSEA = .07, CFI = .90. The standardized factor loadings ranged from .56 to .87 for the Discrepancy factor, .42 to .84 for the Standards factor and .58 to .82 for the Order factor. Our results of this initial model were comparable to the results described in the original paper: $\chi^2(227) = 1204.04$, $p < .001$, CFI = .888, RMSEA = .081. Standardized factor loadings ranged from .420 to .843 for the Standards factor, .610 to .863 for the order factor, and .632 to .837 for the discrepancy factor. However, when we re-ran the original model, accounting for the ordered categorical item responses, the fit for the original model was worse: $\chi^2(227) = 2123.99$, $p < .001$, CFI = .919, TFI = .909, RMSEA = .112 (90% CI = .107, .116). Standardized factor loadings ranged from .493 to .892 for the Standards factor, .642 to .869 for the Order factor, and .658 to .877 for the Discrepancy factor (see Table 2).

Model 1

We believe that items 3, 15, and 23 are measuring negative affect as well as discrepancy based on a critical review of the item content. Thus, the first model we hypothesized tests the original APS-R scale factor structure, excluding these three items. The Standards and Order factors remained unaltered from the original model. See Figure 2.

The results of Model 1 were better than the original model: χ^2 (167) = 1162.67, $p < .001$, TLI = .948, CFI = .954, RMSEA = .094 (90% CI = .089, .099). Standardized factor loadings for Standards ranged from .470 to .892, from .640 to .868 for the Order factor and from .730 to .887 for the Discrepancy factor. Given the results of Model 1 and our theoretical rationale, we continue to exclude items 3, 15, and 23 for all further analyses (see Table 2).

Model 2

As mentioned above, we hypothesized that several items intended to measure discrepancy are best interpreted as measuring dissatisfaction. These items are 13, 17, 20, and 21. These items tend to reflect degree of feeling dissatisfaction with performance and accomplishments. Item 21 could arguably have been included on either the dissatisfaction or discrepancy factor but was deemed to best reflect dissatisfaction because of the reference to a feeling (i.e., I hardly ever feel that what I've done is good enough). Thus, we propose an alternate model with a fourth factor called "Dissatisfaction" composed of these four items (see Figure 3). The Standards and Order factors remained unaltered from the original model.

The fit of Model 2 was a slight improvement over Model 1: χ^2 (164) = 1110.17, $p < .001$, TLI = .949, CFI = .956, RMSEA = .093 (90% CI = .088, .098). Standardized factor loadings ranged from .470 to .892 for Standards, .641 to .868 for Order, .741 to .859 for Discrepancy, and .748 to .906 for Dissatisfaction. See Table 2.

Model 3

Finally, in an effort to maintain the three factors reported by Slaney et al. (2001), we hypothesized a third model that excluded all the items that do not measure pure discrepancy (*i.e.,* those included as a "dissatisfaction" factor in Model 2), and ran a model with only items 6, 9, 11, 16, and 19 as indicators of Discrepancy, and the Standards and Order factors remaining unaltered from the original model (see Figure 4).

The fit of Model 3 was not as good as the fit for Models 1 and 2, but still an improvement over the original model: χ^2 (101) = 856.19, p < .001, TLI = .938, CFI = .948, RMSEA = .106 (90% CI = .099, .112). Standardized factor loadings ranged from .462 to .893 for the Standards factor, .640 to .869 for the Order factor, and .762 to .854 for the Discrepancy factor. See Table 2.

Sample 2

Model 1

Model 1 is depicted in Figure 2. As with our previous sample, this model was a decent fit to the data, χ^2 (167) = 1886.66, p < .001, TLI = .963, CFI = .968, RMSEA = .103. Standardized factor loadings were much higher in this sample, ranging from .73 to .93 for the Discrepancy factor, .71 to .87 for the Standards factor, and .73 to .88 for the Order factor.

Model 2

Model 2 is depicted in Figure 3. As with our original sample, this model was the best fit to the data (compared to the other models tested) from the second sample: χ^2 (164) = 1654.49, p < .001, TLI = .968, CFI = .972, RMSEA = .096. Standardized factor loadings ranged from .74 to .90 for the Discrepancy factor, .61 to .87 for the Standards factor, .73 to .88 for the Order factor, and .83 to .94 for the Dissatisfaction factor.

Model 3

Model 3 is diagrammed in Figure 4. This model had the worst fit of our three hypothesized models, congruent with our findings from our first sample: χ^2 (101) = 1188.32, p < .001, TLI = .954, CFI = .962, RMSEA = .105. Standardized factor loadings ranged from .76 to .88 for the Discrepancy factor, .61 to .87 for the Standards factor, and .73 to .88 for the Order factor.

Summary

Given the results from both samples, we conclude that Model 2 (see Figure 3) is the best fitting model for the data among the models tested. The common variance among the APR-S scale items seems best explained by 4 factors - discrepancy, dissatisfaction, standards, and order. Correlations among the subscales from Model 2 for both samples can be found in Table 3.

Reliability

Finally, as a last step, we assessed the internal consistency of each the subscales from Model 2 using Cronbach's alpha (Cronbach, 1951). A general rule of thumb is that the alpha values should be above .70 for good reliability within a scale. In both samples, we found high Cronbach alphas, ranging from .84 to .88 in Sample 1 and ranging from .86 to .92 in Sample 2 (see Table 4). It is worth noting that the brief five-item discrepancy factor had estimated alphas of .88 or greater despite now having fewer items than the original 12-item discrepancy subscale; reducing the number of items in a scale typically has the effect of reducing the calculated level of internal consistency according to the Cronbach formula.

Discussion

The current study evaluated the psychometric properties of the APS-R discrepancy subscales and uniquely tested the complexity of this subscale. Our findings point to the potential benefits of further scrutinizing the psychometric characteristics of well-known and widely used

measures of perfectionism. We tested and found support across two samples for our contention that the discrepancy subscale has two clearly identifiable and replicable factors after removing three items that had item wording that blend discrepancy and negative affectivity. Our analyses established that it is possible and advisable to distinguish two discrepancy factors - a pure discrepancy factor and a dissatisfaction factor. These alterations in the structure of the measure yielded replicable findings and a substantially improved overall fit according to several criteria when evaluated along with two other factors representing standards and order. Our evidence of improved overall fit is noteworthy given that past confirmatory factor analyses have tended to show that analyses of APS-R yield an overall level of fit that is adequate but certainly not ideal (e.g., Mobley, Slaney, & Rice, 2005; Slaney et al., 2001; Wang, Slaney, & Rice, 2007).

While our findings can be regarded as a challenge to the validity of the APS-R, another potentially useful way of interpreting these findings is that we have identified a brief five-item measure of pure discrepancy that seems suitable for use in future research; importantly, this brief measure does not have item content that could overlap with the items comprising key outcome measures. On a similar note, researchers who are interested in distinguishing dissatisfied perfectionists and less dissatisfied perfectionists in future research could be justified in using the four-item dissatisfaction factor detected in the current research. The value of considering individual differences in dissatisfaction in future research is supported by past research findings. For instance, Mor et al. (1994) showed quite clearly that both self-oriented and socially prescribed perfectionism were correlated significantly with performance dissatisfaction in a sample of highly accomplished professional performers. Their dissatisfaction and unhappiness while performing painted a portrait of driven people who are unable to enjoy their successes and accomplishments. This inability to be satisfied means that even excellent performances by

extreme perfectionists will be shrouded in negative affect. This tendency becomes especially problematic when it becomes generalized and incorporated into a chronic sense of self-dissatisfaction.

To our knowledge, our current results represent the first empirical test of the possibility that the APS-R discrepancy dimension actually has more than one factor. These different facets were likely not identified earlier because Slaney et al. (2001) used a statistical approach that constrained the various APS-R items to the one factor they were intended for – the discrepancy factor. The key question at this point involves how best to interpret our current results. Do our findings suggest that the APS-R is problematic in ways that impact its future use? As mentioned earlier, one reasonable response to our findings is to point to past theoretical views of the nature of discrepancy among perfectionists (e.g., Hamachek, 1978; Missildine, 1963) and suggest any failure to include item content that taps dissatisfaction would have actually posed a significant problem in terms of content validity-- that is, dissatisfaction is central to a sense of falling short. Another reasonable objection to our results would be to note that the two factors found within the discrepancy factor were indeed detected but these factors were very highly correlated. However, as we noted earlier, the failure to acknowledge the emphasis on dissatisfaction that is built into a distinguishable subset of the discrepancy items is important at least to the extent that past research has not considered item overlap issues and possible inflation in the correlations between the discrepancy subscale and the various well-being and adjustment measures that tap low satisfaction. And perhaps most importantly, when we examined the pattern of subscale scores for the individual participants in our sample after completing our psychometric analyses, we were able to identify at the level of persons a small but unique subset of seemingly adaptive perfectionists (i.e., high standards but low discrepancy) who actually had *high levels of*

dissatisfaction rather than low levels of dissatisfaction. These people, while not abundant, were clearly evident. Presumably, if we had focused on a reasonably large sample of elite performers who must strive to be perfect, we would have identified a much higher proportion of perfectionists who are non-discrepant because they are exceptionally accomplished and close to or at the standard but these same people are also quite dissatisfied. In some instances, these people far exceed the standards in place. In all likelihood, these individuals likely are highly frustrating to colleagues who cannot fathom why such high performance does not meet with the perfectionist`s approval and satisfaction. For instance, it is well-known in the music field that this sentiment has often been expressed by members of the E Street Band who have sometimes been exasperated and exhausted by the extreme perfectionism displayed by Bruce Springsteen.

We see clear merit in conceptually distinguishing perfectionistic standards from evaluations and responses retained to the attainment or nonattainment of these standards, and this distinction may help distinguish neurotic perfectionists from more narcissistic perfectionists who strive for perfection and are relatively satisfied with themselves and their standing relative to other people. However, how this situation is regarded by others researchers in the perfectionism field likely varies substantially across researchers. For example, Stairs et al. (2012) included a dissatisfaction dimension as one of their nine key perfectionism dimensions in developing their new measure, and then they found a strong correlation (r = .78) between this dissatisfaction factor and the APS-R discrepancy subscale. Their approach treats dissatisfaction as part of the perfectionism construct. Similarly, the earlier influential work by Slade and associates also viewed dissatisfaction as part of the perfectionism construct (see Slade & Owens, 1998). While this is a justifiable approach, it is worth noting that the factor that Stairs et al. (2012) have labelled as being dissatisfaction also seems problematic in ways that seemingly plague the APS-

R discrepancy scale; that is, it seems to blend items tapping dissatisfaction with items that tap a pure form of discrepancy with no reference to dissatisfaction.

While our focus has been on psychometric issues, it is important to also consider the implications of these findings from clinical and counseling perspectives. Many people who have high levels of perfectionism discrepancy tend to chronically experience these discrepancies, and our work suggests that underscoring this discrepancy is a palpable sense of dissatisfaction that likely extends to a pervasive sense of dissatisfaction with the self. An intervention focused on this sense of never being satisfied and associated forms of self-evaluation is potentially illuminating because it suggests that certain vulnerable perfectionists would benefit from learning to engage in more self-reinforcement as well as having more frequent positive self-talk and a greater degree of self-compassion or they run the risk of being miserable due to clinical anhedonia. These observations are in keeping with data showing that maladaptive perfectionists with elevated APS-R discrepancy scores are indeed low in self-compassion (Neff, 2003).

The current findings also have possible implications in terms of evaluating measures derived from the APS-R. Impressive work by Rice and associates has resulted in the development of a briefer version of the APS-R (see Rice, Richardson, & Tueller, 2014). Unfortunately, the four items that comprise the brief discrepancy measure seem to have inherited some of the problems and concerns identified in our current work. The items that comprise this brief discrepancy factor consists of two items that tap the purer form of discrepancy, one dissatisfaction item, and one item that we actually opted to remove in our work due to its focus on the emotion of disappointment. There is merit in future research creating a brief measure of pure discrepancy that is similar to the five-item factor described in our current paper.

While we evaluated the replicability of our findings in a second sample and found quite comparable results, it is clear that the generalizability of these findings needs to be established in future research with other samples, including samples of children and adolescents. The need for further APS-R scale development and evaluation is suggested by existing research with adolescents. Mobley, Slaney, and Rice (2005) performed analyses of the APS-R item responses from 342 academically talented middle school students and they found that three factors emerged via exploratory factor analyses, but goodness-of-fit criteria fell just below the criteria for acceptability. Similarly, Wang et al. (2009) explored the APS-R item responses of over 500 Chinese high school students and found via confirmatory factor analyses that the APS-R yielded only "qualified support" (p. 266) for the adequacy of the factor structure due to goodness of fit statistics indicating that the factor structure was less than ideal.

In summary, the current research found that the APS-R discrepancy subscale is complex and actually has items that tapped discrepancy per se, or dissatisfaction, and there is a third subset of items that tap discrepancy but also tapped negative emotional reactions. We found in large samples of students and adults from the general community that it was possible to identify two replicable factors representing discrepancy and dissatisfaction, and our suggested adjustments to the APS-R yielded an overall fit that was substantially better than the fit found in past research. Our work also identified a brief five-item measure of pure discrepancy that we recommend for use for investigators seeking a brief uncontaminated measure. Collectively, our findings illustrate the potential usefulness of conducting further psychometric analyses of well-known perfectionism inventories and then using the information that is gained to guide the construction of measures that are certainly not perfect or even almost perfect but are definitely improved.

References

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.

Enns, M. W., Cox, B. J., Sareen, J., & Freeman, P. (2001). Adaptive and maladaptive perfectionism in medical students: A longitudinal investigation. *Medical Education*, *35*, 1034-1042.

Flett, G. L., & Hewitt, P. L. (2014). Measures of perfectionism. In G. J. Boyle, D. H. Saklofske, & G. Matthews (Eds.), *Measures of personality and social psychological constructs (Vol. 2, pp. 595-618)*. London: Elsevier.

Flett, G. L., Hewitt, P. L., Endler, N. S., & Bagby, R. M. (1995). Conceptualization and assessment of personality factors in depression. *European Journal of Personality*, *9*, 309-350.

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: 10.1177/027243160302300005

Gnilka, P. B., Ashby, J. S., & Noble, C. M. (2013). Adaptive and maladaptive perfectionism as mediators of adult attachmen styles and depression, hopelessness, and life satisfaction. *Journal of Counseling and Development*, *91*, 78-86.

Grzegorek, J. L., Slaney, R. B., Franze, S., & Rice, K. R. (2004). Self-criticism, dependency, self-esteem, and grade-point average satisfaction among clusters of perfectionists and nonperfectionists. *Journal of Counseling Psychology*, *51*, 192-200.

Hamachek, D. E. (1978). Psychodynamics of normal and neurotic perfectionism. *Psychology*, *15*, 27-34

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

Johnson, D. P., & Slaney, R. B. (1996). Perfectionism: Scale development and a study of perfectionists in counselling. *Journal of College Student Development*, *37*, 29-41.

Levinson, C. A., Rodebaugh, T. L., Shumaker, E. A., Menatti, A. R., Weeks, J. W., White, E. K., Heimberg, R. G., Warren, C. S., Blanco, C., Schneier, F., & Leibowitz, M. R. (2015). Perception matters for clinical perfectionism and social anxiety. *Journal of Anxiety Disorders*, *29*, 61-71.

Missildine, W. H. (1963). *Your inner child of the past*. New York: Pocket Books.

Mobley, M., Slaney, R. B., & Rice, K. G. (2005). Cultural validity of the Almost Perfect Scale-Revised for African American college students. *Journal of Counseling Psychology*, *52*, 629-639. DOI: 10.1037/0022-0167.52.4.629

Mor, S., Day, H. I., Flett, G. L., & Hewitt, P. L. (1995). Perfectionism, control, and components of performance anxiety in professional performers. *Cognitive Therapy and Research*, *19*, 207-225.

Muthén, L. K., & Muthén, B. O. (1998-2011). *Mplus User's Guide. Sixth Edition*. Los Angeles, CA: Muthén & Muthén.

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, *2*, 223–250. doi:10.1080/15298860309027

Nicholls, J. G., Licht, B. G., & Pearl, R. A. (1982). Some dangers of using personality questionnaires to measure personality. *Psychological Bulletin*, *92*, 572-580.

Paulson, L. R., & Rutledge, P. C. (2014). Effects of perfectionism and exercise on disordered eating in college students. *Eating Behaviors*, *15*, 116-119.

Pearson, C. A., & Gleaves, D. H. (2006). The multiple dimensions of perfectionism and their relation with eating disorder features. *Personality and Individual Differences*, *41*, 225-235.

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.

Sherry, S.B., MacKinnon, A.L., Fossum, K-L., Antony, M. M., Stewart, S. H., Sherry, D. L., Nealis, L. J., & Mushquash, A. R. (2013). Perfectionism, discrepancies, and depression: Testing the perfectionism social disconnection model in a short-term, four-wave longitudinal study. *Personality and Individual Differences*, *54*, 692-697.

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.

Sirois, F. M. (2015). Is procrastination a vulnerability factor for hypertension and cardiovascular disease? Testing an extension of the procrastination-health model. *Journal of Behavioral Medicine*, *38*, 578-589

Sirois, F. M., & Molnar, D. S. (2014). Perfectionism and maladaptive coping styles in chronic fatigue syndrome, irritable bowel syndrome, fibromyalgia/arthritis, and healthy controls. *Psychotherapy and Psychosomatics*, *83*, 384-385.

Slade, P. D., & Owens, R. G. (1998). A dual process model of perfectionism based on reinforcement theory. *Behavior Modification*, *22*, 372-390.

Slaney, R. B., & Johnson, D. G. (1992). *The Almost Perfect Scale*. Unpublished manuscript, Pennsylvania State University, State College.

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.

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

Stairs, A. M., Smith, G. T., Zapolski, T. C. B., Combs, J. L., & Settles, R. E. (2012). Clarifying the construct of perfectionism. *Assessment*, *19*, 146-166.

Stoeber, J., & Rambow, A. (2007). Perfectionism in adolescent school students: Relations with motivation, achievement, and well-being. *Personality and Individual Differences*, *42*, 1379-1389.

Wang, K. T., Yuen, M., & Slaney, R. B. (2009). Perfectionism, depression, loneliness, and life satisfaction: A study of high school students in Hong Kong. *The Counseling Psychologist*, *37*, 249-274. DOI: 10.1177/0011000008315975

Table 1. *Confirmatory Factor Analysis model fit for the 5 models for each sample.*

Model	χ^2	df	p-value	CFI	TLI	RMSEA
			Sample 1			
Original	2123.99	227	< .001	.919	.909	.112
Model 1	1162.67	167	< .001	.954	.948	.094
Model 2	1110.17	164	< .001	.956	.949	.093
Model 3	856.19	101	< .001	.948	.938	.106
			Sample 2			
Original	2672.57	227	< .001	.958	.953	.105
Model 1	1886.66	167	< .001	.968	.963	.103
Model 2	1654.50	164	< .001	.972	.968	.096
Model 3	1188.32	101	< .001	.962	.954	.105

Table 2. *Factor Loadings and Standard Errors from the CFA results in both samples.*

			Sample 1 (N = 671)		Sample 2 (N = 977)	
Model	Factor	Item	Standardized Estimate	Std. Error	Standardized Estimate	Std. Error
Original	Discrepancy	3	.688	.019	.596	.018
		6	.798	.013	.827	.009
		9	.720	.017	.730	.013
		11	.834	.011	.845	.008
		13	.803	.013	.823	.009
		15	.689	.019	.708	.014
		16	.805	.013	.864	.008
		17	.722	.017	.839	.008
		19	.805	.013	.892	.006
		20	.877	.009	.928	.004
		21	.812	.012	.873	.008
		23	.658	.020	.793	.010
	Standards	1	.781	.016	.782	.013
		5	.493	.032	.604	.020
		8	.892	.011	.870	.009
		12	.874	.011	.863	.009
		14	.796	.015	.855	.010
		18	.620	.024	.703	.016
		22	.788	.016	.833	.011
	Order	2	.779	.018	.730	.016
		4	.869	.014	.823	.012
		7	.642	.023	.777	.014
		10	.842	.017	.880	.010
Model 1	Discrepancy	6	.799	.013	.826	.009
		9	.730	.017	.730	.013
		11	.843	.011	.845	.008
		13	.813	.012	.824	.009
		16	.799	.013	.859	.008
		17	.731	.017	.842	.008
		19	.805	.013	.893	.006
		20	.887	.009	.933	.004
		21	.821	.012	.874	.008
	Standards	1	.791	.015	.788	.013
		5	.470	.032	.605	.020
		8	.892	.011	.870	.009
		12	.867	.011	.859	.009
		14	.797	.015	.854	.010

		18	.640	.023	.709	.016
		22	.784	.016	.830	.011
	Order	2	.787	.018	.733	.016
		4	.868	.014	.822	.012
		7	.640	.023	.776	.014
		10	.837	.016	.879	.010
Model 2	Discrepancy	6	.810	.013	.833	.009
		9	.741	.017	.736	.013
		11	.859	.011	.851	.008
		16	.814	.013	.868	.007
		19	.824	.012	.904	.006
	Standards	1	.791	.015	.789	.013
		5	.470	.032	.606	.020
		8	.892	.011	.871	.009
		12	.867	.011	.860	.009
		14	.797	.015	.854	.010
		18	.640	.023	.708	.016
		22	.784	.016	.829	.011
	Order	2	.786	.018	.733	.016
		4	.868	.014	.823	.012
		7	.641	.023	.776	.014
		10	.837	.016	.879	.010
	Dissatisfaction	13	.835	.012	.832	.009
		17	.748	.017	.849	.008
		20	.906	.008	.943	.004
		21	.835	.011	.881	.008
Model 3	Discrepancy	6	.828	.012	.855	.008
		9	.762	.016	.758	.012
		11	.854	.012	.877	.007
		16	.804	.013	.849	.008
		19	.800	.014	.862	.008
	Standards	1	.797	.015	.787	.013
		5	.462	.032	.606	.020
		8	.893	.011	.870	.009
		12	.862	.011	.861	.009
		14	.797	.015	.854	.010
		18	.652	.022	.707	.016
		22	.779	.016	.830	.011
	Order	2	.789	.018	.727	.016
		4	.869	.013	.824	.012
		7	.640	.023	.777	.013
		10	.834	.016	.881	.010

Table 3. *Correlations among the subscales in Model 2.*

	Discrepancy	Dissatisfaction	Standards	Order
		Sample 1		
Discrepancy	1.00	--	--	--
Dissatisfaction	.820**	1.00	--	--
Standards	.166**	.182**	1.00	--
Order	.113**	.074	.437**	1.00
		Sample 2		
Discrepancy	1.00	--	--	--
Dissatisfaction	.872**	1.00	--	--
Standards	.066*	-.016	1.00	--
Order	.024	-.073*	.520*	1.00

* $p < .05$
** $p < .01$

Table 4. *Cronbach Alphas (internal consistency) and descriptive statistics for the four subscales of Model 2.*

Subscale	# of items	Cronbach Alpha	Mean	Standard Deviation
		Sample 1		
Discrepancy	5	.883	19.78	6.51
Dissatisfaction	4	.845	14.16	5.55
Standards	7	.872	36.49	7.17
Order	4	.841	19.67	4.65
		Sample 2		
Discrepancy	5	.915	20.18	7.30
Dissatisfaction	4	.923	14.77	6.20
Standards	7	.891	38.24	7.31
Order	4	.860	20.00	5.12

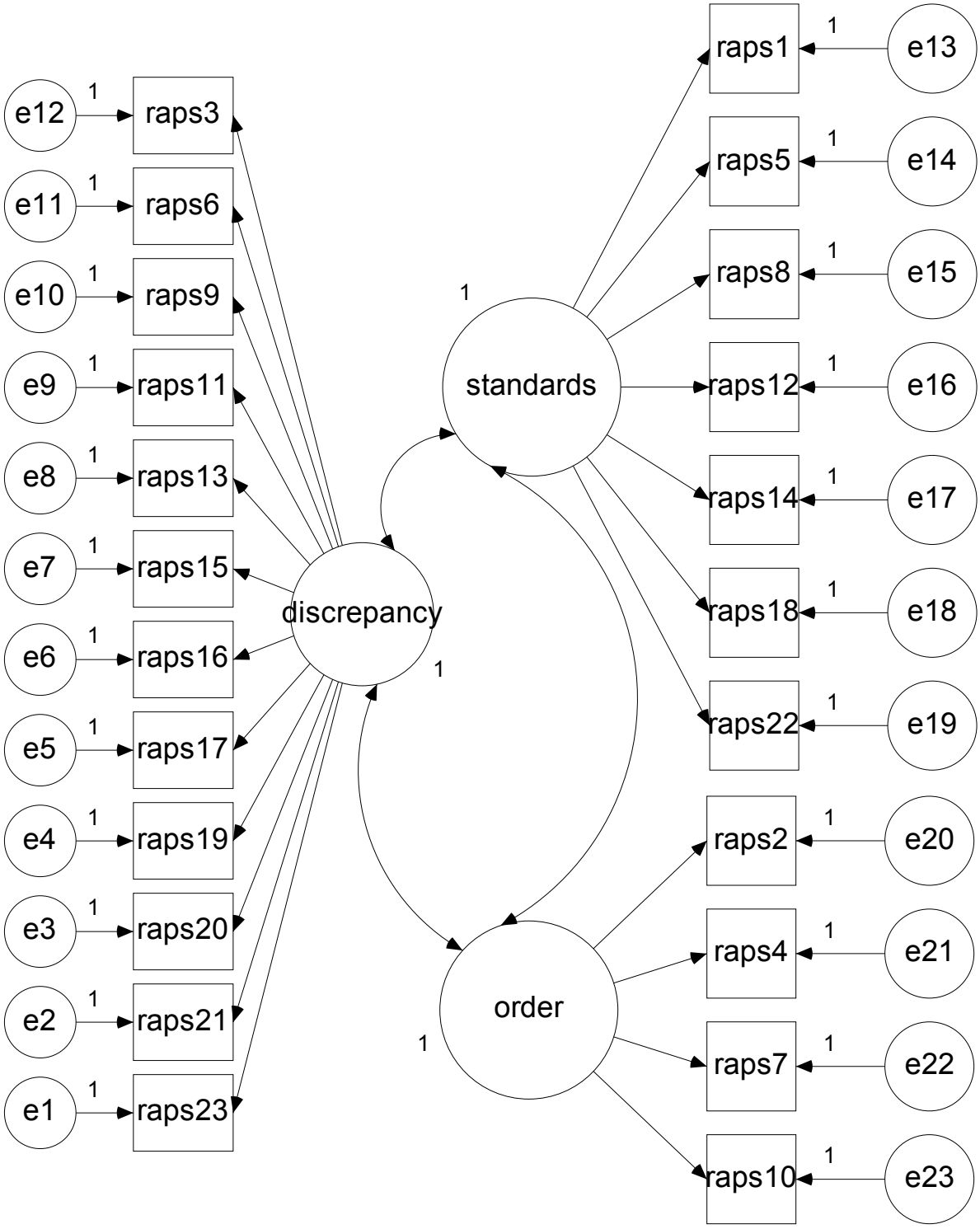

Figure 1. Original Model outlined in Slaney et al., 2001

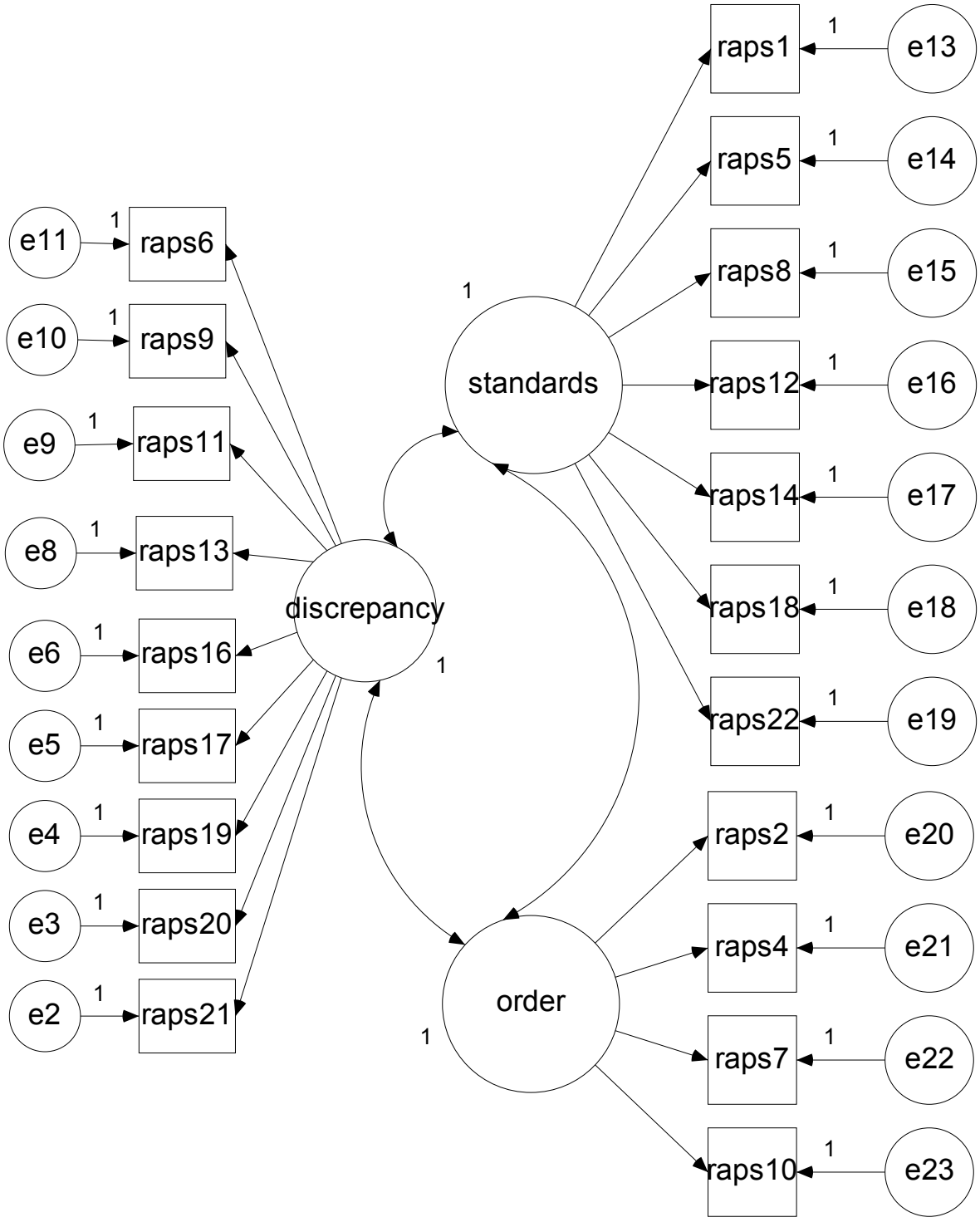

Figure 2. Model 1, removing items 3, 15, and 23.

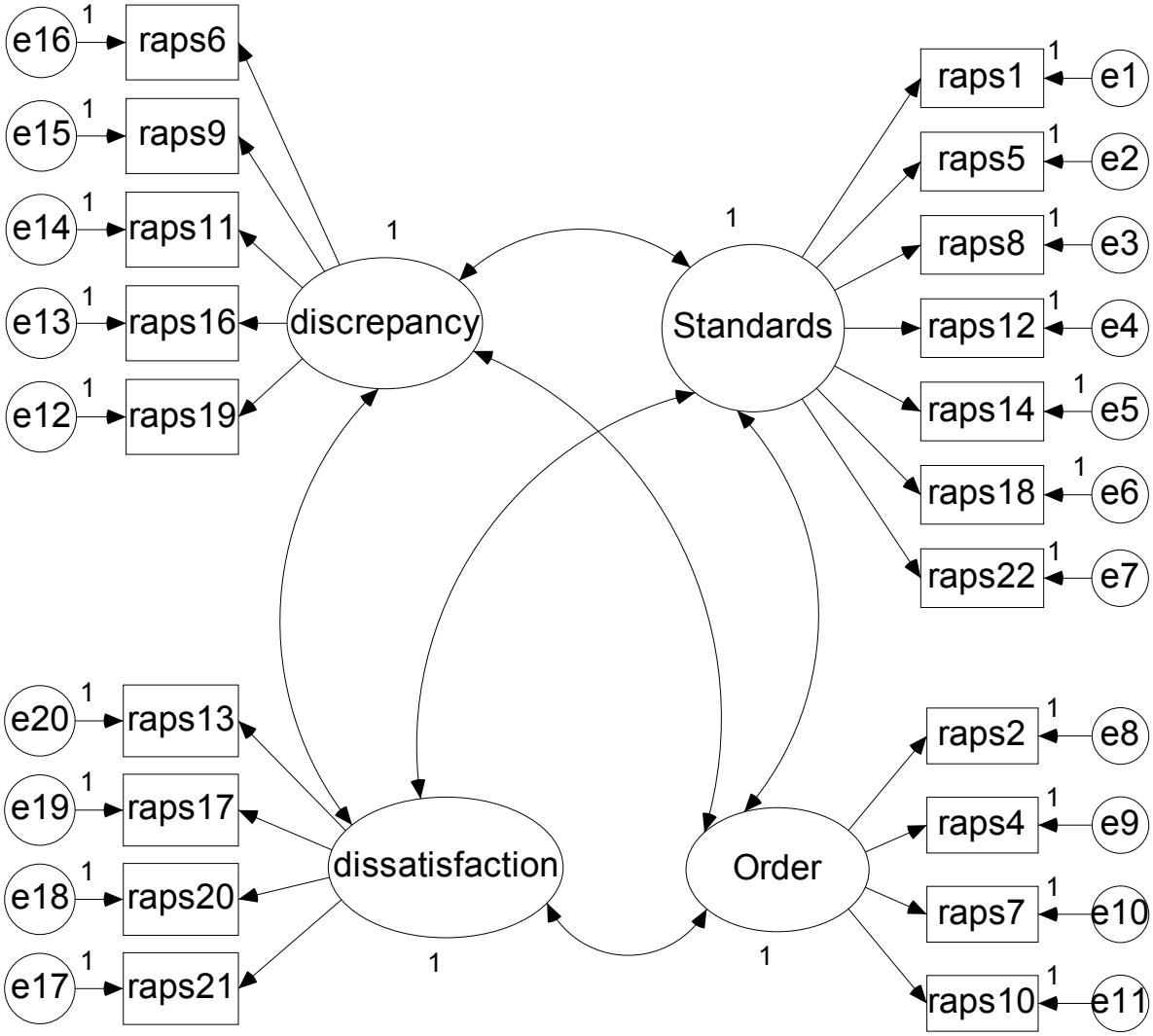

Figure 3. Model 2, including "Dissatisfaction" factor with 4 indicators.

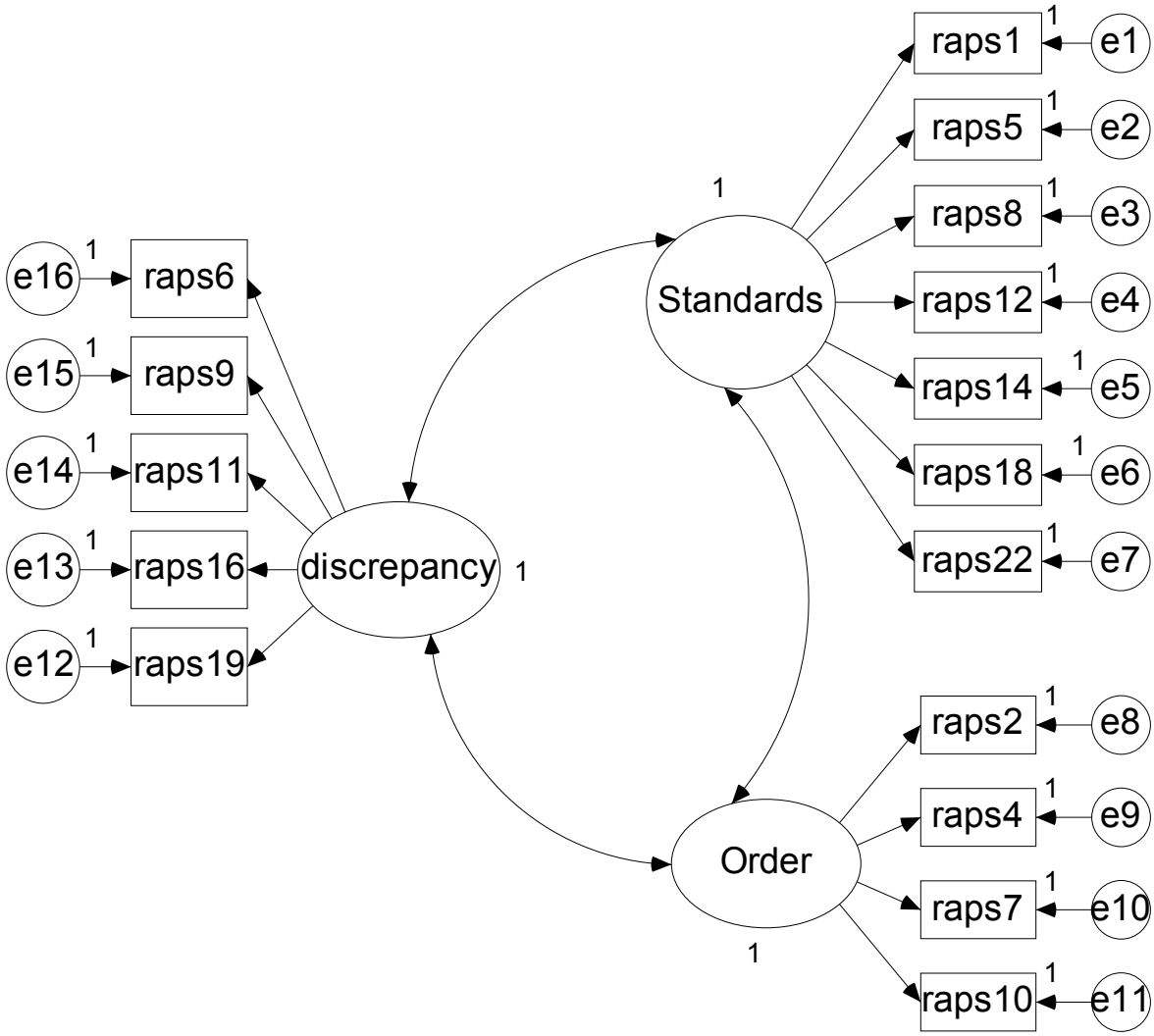

Figure 4. Model 3, removing "Dissatisfaction" factor

Appendix 1

Dissatisfaction Factor

13. I am never satisfied with my accomplishments.
17. I am not satisfied even when I know I have done my best.
20. I am hardly ever satisfied with my performance. *
21. I hardly ever feel that what I've done is good enough.

Discrepancy Factor

6. My best just never seems to be good enough for me.
9. I rarely live up to my high standards.
11. Doing my best never seems to be enough. *
16. My performance rarely measures up to my standards. *
19. I am seldom able to meet my own high standards for performance.

Items Overlapping With Negative Emotion/Affect

3. I often feel frustrated because I can't meet my goals.
15. I often worry about not measuring up to my own expectations.
23. I often feel disappointment after completing a task because I know I could have done better. *

Note. Asterisk denotes items in the short form of the Almost Perfect Scale--Revised