Toward an index of adaptive personality regulation.

OPEN ACCESS

Empirical Research Paper

Toward an Index of Adaptive Personality Regulation

Paul Irwing[1] (iD), **Clare Cook**[2] (iD), and **David J. Hughes**[1] (iD)

Abstract
The idea that matching personality expression with situational demands is adaptive is implicit in many accounts of personality. Numerous constructs and measures have been posited to address this or similar phenomena. Few have proven adequate. In response, we proposed and tested a novel measurement approach (the APR index) assessing real-time behavior to rate participants' success in matching personality expression with situational demands, which we denote adaptive personality regulation. An experimental study ($N = 88$) and an observational study of comedians ($N = 203$) provided tests of whether the APR index constituted a useful metric of adaptive personality regulation. In both studies, the APR index showed robust psychometric properties; was statistically unique from mean-level personality, self-monitoring, and the general factor of personality expression; and provided incremental concurrent prediction of task/job performance. The results suggest that the APR index provides a useful metric for studying the phenomenon of successfully matching personality expression to situational demands.

Mean or trait levels of personality provide useful descriptions of human behavior, so too does the considerable intra-individual variability in personality expression (aka state personality; Fleeson, 2001, 2004; Fleeson & Gallagher, 2009). Experience sampling studies that assess personality expression over short timeframes have estimated that mean levels explain around 35% of the variance in momentary behavior with around 65% explained by intraindividual variability in personality expression (Fleeson & Gallagher, 2009; Sherman et al., 2015).

Subsequent studies have demonstrated that intraindividual variability in personality expression correlates with situational characteristics (e.g., Fleeson & Law, 2015; Sherman et al., 2015; Wright & Mischel, 1987) and current goals (e.g., Bleidorn, 2009; Heller et al., 2007; McCabe & Fleeson, 2016; Perunovic et al., 2011; Sauerberger & Funder, 2017). Individual differences in intraindividual variability in personality expression also show impressive levels of reliability (Baird et al., 2006; Fleeson, 2001). Thus, it appears that "not only does behavior change across contexts, but it changes meaningfully with the demands of the situation at hand" (Sauerberger & Funder, 2017, p. 270). Perhaps the situation-contingent personality expression, observed within the literature, represents, in part, attempts to regulate personality expression to meet situational requirements and goals.

Many theories are concerned with the relationship between expressed personality and situational variables/ appraisals, including whole trait theory (Fleeson & Jayawickreme, 2015), cybernetic Big 5 theory (DeYoung, 2015), and trait-activation theory (Tett & Burnett, 2003). Whole trait theory notes that an empirical relationship exists between personality expression and situational variables, but is mute as to a possible explanation. Cybernetic Big 5 Theory explicitly discusses this relationship in the form of characteristic adaptations. Trait-activation theory considers that different situations elicit the expression of different personality traits. Each of these theories, at least implicitly, suggests that a meaningful proportion of personality state expressions represent attempts to express situation relevant personality characteristics. Put another way, "To survive and function well as a member of society—or at all—people must be able to adapt their behavior to the situation that confronts them" (Sauerberger & Funder, 2017, p. 264). Thus, finding an index or measure of how people express their personality in response to situational requirements would greatly facilitate

[1]The University of Manchester, UK
[2]Northumbria University, Newcastle upon Tyne, UK

Corresponding Author:
Paul Irwing, Alliance Manchester Business School, The University of Manchester, Room 7.005, Booth Street West, Manchester M15 6PB, UK.
Email: paul.irwing@manchester.ac.uk

current research efforts to understand personality dynamics (e.g., Cervone & Little, 2019; Sauerberger & Funder, 2017, see also special issues in the *European Journal of Personality*, 2015; *Frontiers in Psychology*, 2017; *Journal of Research in Personality*, 2015, 2017, *Personality and Individual Differences*, 2019).

Accordingly, we propose a novel measurement index of the phenomenon outlined above, which we term, adaptive personality regulation, tentatively defined as, *the ability or propensity to express personality to meet situational requirements.* Based on this definition, an index of adaptive personality regulation must assess (a) the expression of personality required within a situation and (b) the extent to which the appropriate levels of personality are expressed.

Numerous constructs address similar phenomena, including self-monitoring (Snyder, 1974); functional flexibility (Kashdan & Rottenberg, 2010; Paulhus & Martin, 1988); behavioral signatures (Mischel & Shoda, 1995); flux, pulse, and spin measures (Côté et al., 2012; Moskowitz & Zuroff, 2004, 2005); intraindividual personality variability across situations (Lievens et al., 2018); and situation-contingent personality expression (Bleidorn, 2009; Judge et al., 2014; Minbashian et al., 2010; Wood et al., 2019). Table 1 displays prototypical summaries of these constructs, including definitions, approach to personality assessment, approach to assessing situational requirements, and whether they differentiate between adaptive and maladaptive regulation. Creating constructs and measures for the sake of it stifles rather than strengthens scientific progress (Hughes, 2018; Hughes & Evans, 2018); thus, we have taken great care to review existing work to ensure that adaptive personality regulation and the associated measurement index are indeed substantively novel. As evident from Table 1, and the discussion that will follow, these constructs and their assessment procedures differ from that we propose.

Regarding definition, functional flexibility is "the ability to adjust one's behavior to the interpersonal demands of a wide range of situations" (Paulhus & Martin, 1988, p. 91)—and one aspect of psychological flexibility—"how a person (1) adapts to fluctuating situational demands" (Kashdan & Rottenberg, 2010, p. 866)—perhaps comes closest to what we think of as adaptive personality regulation. However, both of these conceptions are in different ways circumscribed. Psychological flexibility because it remains theoretical and does not have an associated measure, and functional flexibility because it focuses exclusively on interpersonal situations, whereas the need to regulate personality expression should include all classes of situations, such as when working alone.

Regarding personality assessment, most of the existing approaches utilize static and global self-assessments akin to normal self-report personality inventories (see Table 1). Although such global self-reports can be informative, they are prone to many well-established limitations (e.g., socially desirable responding) and do not provide dynamic assessments of personality. Put simply, we agree with Erickson et al. (2009) who concluded that the Battery of

Interpersonal Capabilities (BIC; the assessment tool for functional flexibility) does not measure flexibility and that "this particular personality construct may not be amenable to standard self-report measurement formats" (p. 908). Beyond the use of global self-reports, one other reason that many previous measures have proven suboptimal is that they do not assess whether variation in behavior is adaptive/functional (Lievens et al., 2018; Minbashian et al., 2010) or maladaptive/dysfunctional (Clifton & Kuper, 2011; Côté et al., 2012; Fournier et al., 2009; Zeigler-Hill et al., 2013). To do so, one would need to take account of whether the expressed personality states match situational requirements.

Regarding the assessment of personality-situation match, the most common approach has been to identify "psychologically significant" situational characteristics and to measure their contingent relationships with personality. Early work based on the Cognitive-Affective System Theory of Personality typically took an ad hoc ideographic approach to identifying both relevant situational characteristics and behaviors using retrospective qualitative reports (Mischel & Shoda, 1995; Wright & Mischel, 1987). This work yields rich information but does not yield a generalizable index of adaptive personality regulation that would meet our criteria.

Other approaches have used standardized and generalizable descriptive frameworks such as job characteristics or broader taxonomies of situations (e.g., DIAMONDS, CAPTION, Parrigon et al., 2017; Rauthmann et al., 2014; Riverside Situational Q-Sort, Sauerberger & Funder, 2020) and worked backward to identify relevant personality traits (e.g., Minbashian et al., 2010), somewhat similarly to trait-activation theory (e.g., Holman & Hughes, 2021). One difficulty with using these frameworks to assess adaptive personality regulation is that the line of theoretical inference from situational characteristics to required personality characteristics is in its infancy (Sherman et al., 2015) and requires a convoluted line of inference. For example, using the CAPTIONS model, we might conclude that a situation is high in (a) "Typicality" because it is characterized as "typical, regular, standard" and (b) "Adversity" because it is characterized as "stressful, fatiguing, frustrating." Next, we would work backward to understand which traits such situations trigger. In the case of typicality, this seems fairly straightforward because the defining characteristics describe conscientiousness. However, in terms of "Adversity," this is less straightforward because the characteristics arguably describe a mix of neuroticism, disagreeableness, and low extraversion. Similar difficulties are evident with other situational models.

A simpler approach, which we favor, is to assess situational requirements directly in terms of expert ratings of the personality states needed to facilitate successful goal attainment. For example, successful networking (i.e., meeting and attempting to befriend strangers) is likely to require the expression of a range of extraverted behaviors and would be rated accordingly. Interestingly, this was exactly the approach adopted in a study of transactions between personality and

Table 1. Prototypical Descriptions of the Definitions and Assessment Protocols of Constructs Related to Adaptive Personality Regulation.

Construct label and definition	Personality assessment	Personality-situation match	Differentiates adaptive and maladaptive regulation
Self-monitoring: "the extent to which individuals can and do monitor their self-presentation, expressive behaviour, and non-verbal affective display" (Snyder, 1974, pp. 526–527).	Global self-reported tendencies	No measure	No
Psychological flexibility: "dynamic processes that unfold over time . . . reflected by how a person: (1) adapts to fluctuating situational demands, (2) reconfigures mental resources, (3) shifts perspective, and (4) balances competing desires, needs, and life domains." (Kashdan & Rottenburg, 2010, p. 866).	No measure	No measure	Yes
Functional flexibility: "The ability to adjust one's behaviour to the interpersonal demands of a wide range of situations" (Paulhus & Martin, 1988, p. 91).	Global self-reported tendencies to perform behaviors, to feel anxious when doing so, and avoid triggering situations	No measure	Yes
Cognitive-Affective System Theory of Personality: "the personality system refers to the cognitive-affective mediating units organized in a distinctive network of relations. . . . This system interacts with relevant psychological features of situations, generating the distinctive patterns of variability in social cognition, affect, and action across situations—the individual's personality signature" (Mischel & Shoda, 1995, p. 257).	Many varied assessment procedures, often using qualitative, retrospective descriptions	Tend to be situation-specific and ideographic, thus nongeneralizable	Yes
Situation-contingent personality expression (e.g., Minbashian et al., 2010): No specific definition, draws on Cognitive-Affective System Theory of Personality (Mischel & Shoda, 1995).	Self-report assessment of momentary personality expression	Uses established frameworks (e.g., task characteristics) to describe situational triggers of personality traits	Yes
Flux, pulse, and spin: "flux, variability about an individual's mean score on an interpersonal behavior sampled from the interpersonal circumplex, such as dominance, submissiveness, agreeableness, and quarrelsomeness; pulse, variability about an individual's mean extremity of behavior scores on the interpersonal circumplex; and spin, variability about an individual's mean angular coordinate on the interpersonal circumplex." (Moskowitz & Zuroff, 2004, p. 880).	Global self-reported tendencies	Personality rated following "significant social interactions," but the features of these interactions are not assessed	No
Intraindividual personality variability across situations (e.g., Lievens et al., 2018): No specific definition, draws on Whole Trait Theory to describe personality states (Fleeson & Jayawickreme, 2015).	Self-report responses to situational judgment test scenarios	Yes, embedded within the test scenarios, so not generalizable	Yes

occupational roles (Denissen et al., 2014), which found that ratings of optimal personality profiles of jobs (i.e., the traits or trait expressions required to perform successfully) are very reliable whether made by experts or amateurs.

The Current Study[1]

Our goal is to identify an index of adaptive personality regulation, which we define as *the ability or propensity to express personality to meet situational requirements*. The literature reviewed above supports the contention that, although many conceptually similar constructs and measures exist, none meet our criteria as a measure of adaptive regulation. We now propose a general framework.

First, we avoid global self-reports of typical tendencies and instead directly assess behavior. Second, to assess whether variation in personality expression is adaptive or maladaptive, we must take account of whether the expressed personality states match situational requirements. Third, we should adopt standardized and generalizable frameworks. Fourth, these frameworks should be intuitive and preclude convoluted processes of inference when matching situations and required personality expression. Accordingly, we propose to measure adaptive personality regulation by observing individuals in situ and assessing the extent to which (a) their expressed personality (b) approximates expert ratings of the personality expression required within the situation. Henceforth, we will refer to this approach to measuring adaptive personality regulation as the APR index. For clarity, we provide a general expression for the calculation of the APR index:

$$\frac{\sum_{1}^{Sit_s} \left(\frac{\sum_{1}^{Trait_s} \left| \overline{OpP} - \overline{ExP} \right|}{Trait_s} \right)}{Sit_s}, \tag{1}$$

where $\overline{OpP}$ = mean ratings of the optimal level of expressed personality, $\overline{ExP}$ = mean ratings of expressed personality, $Trait_s$ = the number of traits relevant to performance in the situation, Sit_s = the number of situations, and $\|$ = take the modulus of the score such that the signs of the difference are eliminated[2]. This formula will require adaptation, for example if $Trait_s$ or Sit_s equals 1, then the associated integral disappears. Formula 1 produces an index in which higher scores indicate less adaptive personality regulation (i.e., personality expression that is further away from optimal levels). To reverse code the scores, subtract the APR index scores from the maximum possible score. Furthermore, the formula as written applies to observed variables only. When using latent variables within a confirmatory factor analysis (CFA) model, one should calculate $|\overline{ExP} - OpP|$ for each relevant

trait and use each score as an indicator of a single general factor of adaptive personality regulation. If there are multiple situations, then the expression for each situation should be loaded on a general factor. See Figure S1, Supplemental Material for an illustration.

The question now arises as to what form of evidence would support this approach to conceptualizing adaptive personality regulation. First, and most fundamentally, is that individuals shift their expression of personality according to situational requirements. There is already a substantial body of evidence supporting this contention (see above). Nevertheless, we must establish this in the current framework:

> **Hypothesis 1 (H1):** Individuals shift their expression of personality according to situational requirements.

Second, although many systems might be involved in regulating personality expression, we posit adaptive personality regulation as an index of a general ability to regulate personality across situations. Thus, we predict that those who are able to regulate one trait in one situation will be able to regulate a second trait in a second situation, which will give rise to a common or general factor. This prediction is in line with studies of the structure of intraindividual variability that also support a general factor (e.g., Baird et al., 2006):

> **Hypothesis 2 (H2):** The APR index should conform to a single-factor model.

Third, as we discussed earlier, introducing novel concepts, terms, or labels without good cause hinders rather than helps psychological science. Thus, "discriminant [validity] evidence is particularly important for showing that a new measure is actually new and preventing construct proliferation . . . which occurs when ostensibly new constructs are proposed that are theoretically and/or empirically indistinguishable from existing constructs" (Hughes, 2018, p. 771). Thus, we provide robust tests of what is traditionally termed "discriminant validity" by examining whether adaptive personality regulation differs from conceptually similar constructs (Le et al., 2010; Shaffer et al., 2016). Specifically, we examine whether adaptive personality is discriminant from mean-level personality and self-monitoring (Snyder, 1974), with the latter providing one measure of goal-directed regulation of personality expression:

> **Hypothesis 3 (H3):** The APR index should show discriminant validity with respect to mean-level personality and self-monitoring.

In addition to capturing unique variance (i.e., being discriminant) in describing behavior, for the APR index to be of value, it must also explain unique variance in relevant outcomes (Hughes, 2018). The expectation is that the APR index will predict task/job performance.[2] The Big 5 broad

personality factors predict job performance, but weakly (coefficients from .07 to .24; Barrick et al., 2001). However, the "predictive validity" of personality can be improved using other ratings, facets rather than broad factors, and by measuring specific rather than broad performance criteria (see Hughes & Batey, 2017, for review). One reason that mean-level personality provides underwhelming prediction likely stems from the fact that most jobs require performance in multiple domains/situations, thus requiring adaptive personality regulation:

> **Hypothesis 4 (H4):** The APR index will show concurrent and incremental prediction of task performance compared with mean-level personality and self-monitoring.

We now present two studies in which we examined these hypotheses and the potential suitability of the APR index as a measure of adaptive personality regulation. In both, we observed people in situations that differed in terms of the personality expression required and assessed the degree to which they expressed the relevant personality. Study 1 was a laboratory experiment and Study 2 was a real-world, high-stakes, job performance setting, namely, observing comedians perform live at one of the premier venues in the United Kingdom.

Study 1 Method

Design

The key feature of adaptive personality regulation is that people shift their personality to meet situational requirements. Thus, we used a repeated measures design in which participants would have to express different levels of personality to succeed. We chose tasks that required large shifts in the levels of expressed extraversion. Specifically, we observed participants in two experimental conditions, that represented "strong situations", and were unambiguous regarding the requirements for success (Judge & Zapata, 2015): first, a networking condition in which participants were tasked with being the "most memorable person in the room", and second, an independent working condition in which participants were tasked with scoring a number of test responses.

These tasks likely require regulation of multiple traits. However, we focused on extraversion because it is included in the majority of personality models (Funder, 2004), and according to the realistic accuracy model (Funder, 1995, 2012), expressions of extraverted traits are readily observable and well suited to observational ratings. By contrast, facets of conscientiousness (e.g., detail-consciousness), which might help in the scoring task, are difficult to observe (Funder, 1995, 2012).

Sample

We used G*Power 3.1.9.4 to conduct a series of a priori power analyses for each study hypothesis. Following current recommendations (Funder & Ozer, 2019; Lakens, 2013), we set alpha to a level of .05, power to .80, and the effect size to $f = .204$ (i.e., $r = .2$), representing a medium effect. H1 concerns a test of two dependent means (i.e., repeated measures or matched pairs t test) for which G*Power recommended a minimum required sample of $N = 27$. H3 concerns tests of discriminant validity through the assessment of bivariate correlations, for which G*Power recommended a minimum required sample of $N = 118$. H4 concerns the incremental validity of the APR index above six other variables (i.e., mean-level big five and self-monitoring) within a linear multiple regression, for which G*Power recommended a minimum required sample of $N = 78$. For H2, which concerned a test of the fit of a single-factor CFA, we used the power calculation procedures of Preacher and Coffman (2006). MacCallum et al. (2006) recommend the calculation of power based on a range of plausible models. In this instance, we did not have strong a priori predictions concerning the magnitude of the factor loadings and so estimated required power based on simulation models that adopt moderate loadings (e.g., Hu & Bentler, 1998, 1999) and large loadings (McNeish et al., 2017), respectively. With moderate loadings, models are considered to fit with root mean square error of approximation (RMSEA) values of approximately .05 with values greater than .1 indicating model misfit (Hu & Bentler, 1998, 1999). When loadings are large, models are considered to fit with RMSEA values of approximately .1 with values greater than .2 indicating model misfit (McNeish et al., 2017). According to these power calculations, minimum samples sizes of 245 for a model with moderate loadings and 62 for large loadings were required.

In practice, the initial sample comprised 99 participants, which was reduced to 88 due to failures in video-recording equipment. Of the final sample, 32 (36.4%) were male, 55 (62.5%) female, with one missing value. The mean age was 27.90 ($SD = 6.53$). With respect to educational level, 63.6% had a postgraduate university education, 20.5% had undergraduate university education, 4.5% had nonuniversity higher education, and 10.2% had secondary school education to age 18. Thus, the sample surpassed the minimum required N in most cases, but not all, meaning that we generally had adequate power to detect the specific effects hypothesized in this study. However, it must be noted that all power calculations are estimates and the power analyses conducted here pertain to the detection of significant medium-sized effects (i.e., $r = .2$ or greater) rather than precise estimates of the true population parameter. To estimate the true population parameter, much larger sample sizes are required (Lakens, 2013). However, the current sample is close to adequate for the focal hypotheses tested.

Procedure

Participants were recruited via student email contact lists from the University of Manchester, Department of Humanities and via Facebook advertisements. Those who indicated their

willingness to participate were contacted by telephone or message and invited to one of the prescheduled laboratory sessions at the Alliance Manchester Business School. Prior to the experimental session, participants completed online self-report measures of Self-Monitoring (Snyder, 1974), the Big Five Inventory (BFI; John et al., 1991), and demographic items. Participants received a £10 Amazon voucher.

During the experimental session, participants were seated at a table with five people (three participants and two confederates recruited from the University drama society). Each participant first completed a consent form approved by the university ethics committee.

The participants then took part in the memorability task. The memorability task, which required participants to be the most memorable person in the room and lasted 5 min. Participants were instructed that

> Your task is to be the most memorable person in the room. You need to get people to pay attention to you and remember as many [of the ten facts] about you as possible. If you do not know the answer to any of these questions you can make them up for the purpose of this exercise. (See Supplementary Materials, "Study 1 task materials" for the full details)

The scoring task involved scoring response sheets for a putative 40-item questionnaire and lasted 9 min. Participants were instructed that

> . . .you will be taking part in a "speed and accuracy" exercise. You have been provided with some completed response sheets to a questionnaire (you do not have the question sheets). Your task is to score these as quickly and accurately as possible.

During this time, the confederates attempted to engage participants in conversation, thus encouraging extraverted behavior, which was dysfunctional with respect to high performance on this task. Although the confederates had no active role in the first task, they took part in both tasks to preserve their appearance as genuine participants. The behavior of each participant was videotaped for the full duration of both sessions, using two cameras, each positioned face-on (see Supplementary Materials for full experimental protocols and instructions).

Measures

The following measures were derived from the pre-experiment online questionnaire.

Self-Monitoring. We used Snyder and Gangestad's (1986) 18-item scale. Each item was rated as true or false. A meta-analysis has reported a sample weighted reliability of .73 for this scale (Day et al., 2002), and in an extensive review, Gangestad and Snyder (2000) argued that the general factor conforms most closely to the construct of self-monitoring; thus, the scale score equated to the mean item score.

Self-Reported Personality. Self-ratings of mean-level personality were gathered using the BFI (John et al., 1991). The BFI comprises 44 items assessing extraversion, neuroticism, openness, conscientiousness, and agreeableness. Participants respond using a 5-point Likert-type scale (1 = *strongly disagree*, 5 = *strongly agree*). The BFI demonstrates strong internal consistency ($\alpha \approx .83$), a clear factor structure, and convergence with other Big 5 measures (John & Srivastava, 1999). In the current Study, Cronbach alpha reliabilities ranged from .76 to .85 with an average of .81.

The following measures were derived from the experimental tasks and subsequent observations.

Task Performance

Memorability Criterion. Eight raters were recruited to assess the memorability of participants. All raters were educated to undergraduate degree or above and all were working in a professional role that required team working within consequential and high-pressure situations. Thus, they were well attuned to processing real-time verbal information. Each rater received one-to-one training conducted by the authors, which involved briefing them on the procedure (i.e., what they would see, notes could not be taken, and how to complete the form after watching the video) but not the study goals. Raters were shown an example video (not one that was to be included in the subset they were rating) and completed their ratings as a form of direct training. There was also the opportunity for them to ask questions before the rating began. After training, the eight raters were organized into pairs and each pair viewed approximately one quarter of the recordings. Immediately after viewing each recording, the raters recorded the information they could recall about each participant. Each participant had been instructed to communicate 10 facts about themselves (e.g., birthday month, favorite drink, nature of their last night out; see Supplemental Materials for full list), so the maximum score was 10. Interrater reliabilities within the pairs, estimated using Spearman's rho, ranged from $r_s = .59$ to .71 with an average of .66.

Test-Scoring Criterion. Performance in Task 2 was assessed by summing the total number of subscales scored by the participant.

Adaptive Personality Regulation. Two experts in personality rated the videotapes of all 88 participants for level of expressed extraversion using seven of the corresponding BFI items (John et al., 1991, 2008). The items were rated on a 5-point scale ranging from *strongly disagree* to *strongly agree* (for details, see Supplemental Materials—Study 1 rating form), which were scored from 0 to 4. We used these ratings to calculate the APR index following Equation 1. Specifically, we subtracted the mean ratings of expressed personality from ratings of the optimal level of expressed

Table 2. Fit Indices for Invariance Tests.

Model	χ^2	df	p	RMSEA	CFI	TLI	SRMR
Full configural invariance	132.21	76	.00	.092	.964	.957	.049
Full metric invariance	158.950	82	.00	.103	.951	.945	.099
Partial metric invariance	151.50	81	.00	.099	.955	.949	.085
Partial scalar invariance	171.94	87	.00	.105	.946	.943	.095

Note. N = 88. RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker–Lewis index; SRMR = standardized root mean square residual.

personality (i.e., maximum extraversion in Task 1 and minimum extraversion in Task 2), for each personality item. This generated an item-level APR index. We then took the modulus (to remove the signs) of the item-level APR index scores and averaged them to create a total APR index. In this study, the APR index had a mean of 2.73 and a standard deviation of 0.45, on the 0 to 4 scale.

Analysis

Scale Scoring. All analyses were carried out using Mplus version 8.2. We used confirmatory factor models to test Hypotheses 1 (participants should shift in their personality expression) and 2 (the APR index should conform to a single factor), given that they provide stringent tests of the psychometric properties of constructs (Booth & Hughes, 2014). However, to maintain consistency with conventional scoring of self-monitoring and the BFI, all predictor variables in the multiple regressions were represented by observed scores calculated as the average of the item scores.

Estimators. When analyzing ordinal data, we used the Weighted Least Squares Means and Variances (WLSMV) estimator because simulations have supported the consistency, efficiency and unbiasedness of its parameter estimates (Flora & Curran, 2004; Wirth & Edwards, 2007). For similar reasons, we preferred the maximum likelihood estimator with non-normality corrections (MLR), when using scale scores, as these provide an approximation to continuous measurement (Yuan & Bentler, 2000). For the cross-classified mixed effects models, we used a Bayes estimator (Nestler & Back, 2017).

Goodness of Fit. We considered acceptable model fit to be indicated by values of ≤.08 for the RMSEA and standardized root mean square residual (SRMR), and ≥.90 for the Tucker–Lewis index (TLI) and comparative fit index (CFI; Hu & Bentler, 1999; Schermelleh-Engel, et al., 2003).

Results

Correlations and descriptive statistics for all Study 1 variables are shown in Supplemental Materials, Table S1.

Table 3. Standardized Factor Loadings for Task 1 and Task 2 Extraversion: Scalar Invariant.

	Factor loading	
Factor/items	1	2
Factor 1: Task 1 extraversion		
Is talkative	.90	–
Has an assertive personality	.66	–
Is full of energy	.84	–
Is outgoing, sociable	.88	–
Is reserved	.93	–
Tends to be quiet	.92	–
Is sometimes shy, inhibited	.90	–
Factor 2: Task 2 extraversion		
Is talkative	–	.92
Has an assertive personality	–	.80
Is full of energy	–	.89
Is outgoing, sociable	–	.96
Is reserved	–	.91
Tends to be quiet	–	.95
Is sometimes shy, inhibited	–	.91

Note. N = 88.

H1: Individuals Shift Their Expression of Personality to Conform to Situational Requirements. The memorability task required high levels of extraversion to achieve the goal of being "the most memorable person," whereas the scoring task required low levels of extraversion to achieve the goal of scoring the maximum number of tests. To examine shift, we applied an intraindividual CFA to the averaged extraversion ratings from Tasks 1 and 2 (Adolf et al., 2014). First, we tested for measurement invariance because for comparison of latent mean scores to be meaningful; configural, metric, and scalar invariance must hold (Chen, 2007; Millsap & Tein, 2004). Following Millsap and Tein's (2004) procedure for ordered categorical variables and applying Chen's (2007) fit criteria, configural invariance, partial metric invariance, and partial scalar invariance were supported (see Table 2).

The standardized factor loadings of the partially scalar invariant solution are shown in Table 3. The item measuring assertiveness was noninvariant, probably because in Task 2, it was unobservable, as participants worked largely in silence. Thus, it was removed from the estimate of mean shift.

Table 4. Shift in Extraversion Scores From the Memorability to the Scoring Task.

Range	Number of participants	Percentage of sample
4 to 0.01	5.0	5.68
–0 to –0.99	31.0	35.2
–1 to –1.99	32.0	36.4
–2 to –2.99	21.0	23.9
–3 to –4	0.0	0.0

Note. N = 88.

Table 5. Standardized Factor Loadings on Adaptive Personality Regulation of Average Ratings of Expressed Extraversion.

Item	Factor loading
1. Is talkative	.86 [.79, 92]
2. Has an assertive personality	.68 [.54, .81]
3. Is full of energy	.81 [.74, .89]
4. Is outgoing, sociable	.86 [.81, .91]
5. Is reserved (r)	.90 [.86, .94]
6. Tends to be quiet (r)	.93 [.90, .96]
7. Is sometimes shy, inhibited (r)	.89 [.85, .93]
Average variance extracted	66%
McDonald's ω	.95

Note. N = 88. r = reverse scored. The scores are difference scores between required and expressed personality, on a scale from 0 to 4, averaged across raters and situations.

The mean shift in extraversion scores from Task 1 to 2 was estimated at $d = -1.43$ [–1.74, –1.18]. Cohen (1988) describes a d score of ≥.8 as large, so the observed shift from extraversion in Task 1 to introversion in Task 2 is very large. Table 4 provides descriptive statistics for the mean shift in observed extraversion scores on a scale from 0 to 4. It may be concluded that the observed shift is highly variable, and very large in some cases, but never achieves the maximum possible.

H2. The APR Index Should Conform to a Single-Factor Model. As explained above, we calculated "match scores" (i.e., the difference between "ideal personality expression" and "rated personality expression") for each of the BFI items and averaged the "match" scores across both raters and tasks. These average match scores constituted the indicators of adaptive personality regulation and were entered into a single-factor CFA. The model showed close fit with respect to three of the fit statistics, with the RMSEA marginally above the cut-off indicative of close fit ($\chi^2 = 24.17$, $df = 14$, $p = .04$, CFI = .98, TLI = .97, RMSEA = .091 [.015–.150], SRMR = .027). However, the RMSEA tends to overreject models when factor loadings are large (McNeish et al., 2017), which is the case here (see Table 5). Thus, the APR index conforms to a single factor and the large loadings, high average variance

extracted (AVE), and reliability all support its psychometric robustness (Table 5).

H3. The APR index Should Show Discriminant Validity With Respect to Mean-Level Personality, and Self-Monitoring. Discriminant validity evidence was obtained when two criteria were met: first, when the APR index correlated with the other variables at less than $r = .85$ (Shaffer et al., 2016) and second, when the correlation between the two variables was less than the square root of the average variance extracted of each variable (Henseler et al., 2015). Both tests are convincingly satisfied; indeed, all correlations are not significantly different from zero (Table 6).

H4. The APR Index Will Show Concurrent and Incremental Prediction of Task Performance Compared With Mean-Level Personality and Self-Monitoring

   Memorability Task. We used cross-classified mixed effects models (CC-MM; Nestler & Back, 2017) to estimate the regression models because the memorability scores are nested in two sources of nonindependence: participants and raters. First, we examined the proportion of variance (i.e., intraclass correlation) in the memorability scores, due to participants (.67), raters (.05), and the residual (.28), showing that a substantial proportion of variance in scores was due to participants, which is required for the subsequent regressions to be meaningful. Second, we regressed the memorability criterion (i.e., the number of facts recalled by the raters) from the memorability task onto mean-level personality and self-monitoring (Model 1.1, Table 7). Second, we added in the APR index as a predictor (Model 1.2, Table 7). In the first model, only mean-level extraversion was a significant predictor and the overall model accounted for 20.6% of the variance. When added to the regression equation, the APR index was the largest predictor and explained an additional 4.9% variance in performance.

   Scoring Task. We regressed the test-scoring criterion (i.e., the number of tests scored in the allotted time) onto mean-level personality and self-monitoring in one model, and then added the APR index in a second. From the model that comprised mean-level personality and self-monitoring (Model 1.3, Table 7), Neuroticism, Extraversion, and Openness were significant and negative predictors of performance. Again, when added to the regression equation, the APR index was the largest predictor and explained an additional 8.1% variance in performance.

Study 1: Discussion

To a substantial degree, the results of Study 1 support the four hypotheses. First, as predicted, and has been shown in many other studies (e.g., Fleeson & Gallagher, 2009; Sauerberger &

Table 6. Correlations, and Square Root of AVE, for the APR Index With Self-Monitoring and Mean-Level Personality.

Correlated construct	r_I (95% CI)	$\sqrt{AVE_1}$	$\sqrt{AVE_2}$
Self-monitoring	−.078 [−.205, .156]	.86	.28
Mean-level neuroticism	−.146 [−.362, .071]	.86	.51
Mean-level extraversion	−.033 [−.268, .203]	.86	.59
Mean-level openness	.004 [−.205, .213]	.86	.46
Mean-level agreeableness	−.089 [−.325, .146]	.86	.49
Mean-level conscientiousness	.176 [−.041, .393]	.86	.57

Note. AVE = average variance extracted; CI = confidence interval.

Table 7. Regression Models of Memory Score for 10 Facts From the Most Memorable Person Task 1, and Test Scores From Task 2, on Self-Monitoring, Mean-Level Personality, and Adaptive Personality Regulation.

	Most memorable person		Test scoring	
	Model 1.1	Model 1.2	Model 1.3	Model 1. 4
Variables	β [95% CI]	β [95% CI]	β [95% CI]	β [95% CI]
Self-monitoring	−.060 [−.317, .214]	−.019 [−.241, .264]	.140 [−.100, .380]	.134 [−.093, .362]
Mean-level neuroticism	−.037 [−.268, .184]	.002 [−.212, .245]	**−.236 [−.438, −.035]**	**−.207 [−.405, −.008]**
Mean-level extraversion	**.194 [.035, .428]**	**.223 [.026, .440]**	**−.184 [−.420, −.051]**	**−.184 [−.399, −.031]**
Mean-level openness	−.164[−.362, .063]	−.128 [−.352, .097]	**−.285 [−.487, −.083]**	**−.280 [−.469, −.111]**
Mean-level agreeableness	−.131 [−.434, .106]	−.082 [−.313, .162]	−.033 [−.235, .168]	.001 [−.186, .188]
Mean-level conscientiousness	.226 [−.040, .498]	.192 [−.038, .431]	.093 [−.171, .357]	.039 [−.215, .293]
APR index		**.245 [.086, .499]**		**.291 [.127, .455]**
R^2	**.206 [.077, .354]**	**.255 [.090, .444]**	**.194 [.061, .328]**	**.275 [.123, .426]**
ΔR^2		.049		.081

Note. N = 87, β = beta coefficient; CI = 95% confidence interval. Effects significant at the .05 level are in bold.

Funder, 2017), individuals' expression of personality shifted across situations in line with situational requirements, and on average, they showed a very large degree of shift, in terms of Cohen's (1988) effect size criteria. Here, we observed high and low levels of extraversion. An additional and novel observation was that when the required expression of personality is diametrically opposite (i.e., at both extremes of the trait spectrum), that individuals vary greatly in their degree of shift from none at all, to large.

Second, the APR index, which assessed the degree to which expressed personality matched with expert ratings of the optimal levels of personality expression, conformed to a single-factor CFA model with excellent psychometric properties. Calculation of the APR index revealed a range of individual differences whereby for some, personality expression was close to the optimal levels, whereas for others, it was not.

Third, discriminant validity evidence demonstrated that the APR index was distinct from self-monitoring and mean-level personality. Indeed, the correlations were near-zero meaning that the APR index captures novel variance.

Fourth, the unique and novel variance captured by the APR index proved useful, explaining nontrivial levels of incremental variance in task performance compared with mean-level personality and self-monitoring.

Overall, the psychometric properties, correlations, and predictive capacity of the APR index provide initial support for its accuracy and appropriateness as a measure of adaptive personality regulation.

Study 2

Study 2 was devised to investigate a number of the same hypotheses as Study 1, but within a real-world, high-stakes scenario, namely, by studying comedians performing stand-up comedy. Specifically, we examined whether the APR index conforms to a single-factor structure (H2), is discriminant from mean-level personality and self-monitoring (H3), and offers concurrent and incremental prediction of task performance (H4). The methodology was also similar, but represented an advance in three ways that we expected would enhance ecological validity, relative to Study 1.

First, Study 1 examined adaptive personality regulation using a convenience sample within a controlled but relatively contrived environment. In contrast, Study 2 examined adaptive personality regulation using adult comedians in a

real-world, high-stakes, job performance setting, namely, watching them perform live at one of the two pre-eminent U.K. comedy establishments (Irwing et al., 2020). For professional comedians, stage performances comprise a key part of their job that can "make or break" their career. Thus, the comedians are exhibiting a practiced skill and their motivation is high.

Second, the video recordings we used to assess adaptive personality regulation and level of performance were routinely recorded by the comedy venue, so it seems likely that there were minimal study-specific demand characteristics associated with gathering these data (Orne, 1962). Although effects of demand characteristics, evaluation apprehension, and other artifacts induced by taking part in research can never be fully eliminated, the high-stakes nature of the comedians' performance should have acted as a strong counterbalance to minimize such effects (Orne, 1962).

Third, the ratings of task performance (i.e., success of the stand-up performance) were taken from subject matter experts, namely, comedy professionals of considerable experience and stature within the industry.

In addition, in Study 1 we were able to show that the APR index, averaged across scenarios that required high and low extraversion, conformed to a single factor, thus suggesting that adaptive regulation of personality generalizes across situations. An important advance in Study 2 was that we were able to investigate adaptive personality regulation with regard to all dimensions of the Big 5, not just extraversion, because facets from each domain were judged to be important to comedic performance, and observable (Funder, 2012). Accordingly, we were able to test the extent to which adaptive personality regulation generalized across the Big 5 dimensions and to determine whether the resultant APR index still conformed to a single-factor model.

Method

Participants. An a priori power analysis using G*Power 3.1.9.4, testing for R^2 increase in a linear multiple regression fixed model, with seven predictor variables, with alpha = .05, power = .80, and effect size of $f = .204$ (i.e., $r = .2$) (Funder & Ozer, 2019; Lakens, 2013) suggested a minimum sample of $N = 78$. In practice, the recruited sample comprised 206 comedians, 77 professional comedians (67 males, 10 females, $M_{age} = 42.5$, $SD = 9.4$) and 126 amateur comedians (107 males, 18 females, one missing, $M_{age} = 35.4$, $SD = 8.57$). Thus, our sample exceeded the minimum required N. In some cases, there were small amounts of missing data, and in some cases, we had additional data, which are noted in the tables. Professional comedians were employed to perform at a "new material" show or were the master of ceremonies at an amateur comedian night. Amateur comedians worked, unpaid, at a professional "gong show." Both groups were operating in a high-stakes situation, the professionals to further their career and the amateurs to establish one.

Procedure. Comedians were contacted via email by the club's general manager and/or in person on the night of their performance. Participants consented to have their performance recorded by the comedy club and completed the BFI and self-monitoring scales. All, except gong show participants, completed the questionnaire on the night of their performance in the comedy club dressing room. Gong show participants were emailed the questionnaire, which they completed within 2 weeks of their performance. A test of whether completing the BFI on the night or at home biased responses found a nonsignificant Hotelling's T ($V = .051$, $F = 1.20$, $df1 = 5$, $df2 = 112$, $p = .31$), with follow-up tests similarly nonsignificant, suggesting little-no bias resulted from the location in which the questionnaires were completed.

With the comedians' agreement, two experts viewed and rated a 5-min sample of their videoed performance. However, for gong show comedians, the sample only lasted until they were "gonged." The experts were an internationally successful comedian and the comedy club's Technical Director.

Measures

Self-Reported Personality. We used the BFI (John & Srivastava, 1999). In this study, Cronbach's alpha reliabilities ranged from .74 to .84 with an average of .80.

Self-Monitoring. We used Snyder and Gangestad's (1986) 18-item scale. In the current study, the Cronbach's alpha was .67.

Comedic Performance Scale. The measurement of job performance represents a key problem in organizational psychology that arguably has never been fully resolved (Campbell & Wiernik, 2015) and none of the major attempts to develop universal frameworks (e.g., Bartram, 2005) applies very well to the stage performance of comedians. For this reason, we developed a dedicated measure.

Twelve subject matter experts, who had multiple roles in the comedy industry, including reviewer, headline comedian, agent, television comedy producer, promoter, and festival director were asked to list "as many criteria as you can think of to assess a comedian's stage performance." The resultant list conformed to five categories: material, delivery, stagecraft, off-stage professionalism, and audience response (see Supplementary Materials). The recordings of the comedians' stage performance did not allow us to assess off-stage professionalism or the audience response, so we asked the expert raters (the same two industry experts who rated other aspects of the videoed performance of comedians) to rate six items, two each with regard to material, delivery, and stagecraft, on a 7-point scale ranging from *strong disagreement* (1) to *strong agreement* (7). A CFA supported this three-factor structure, and a known-groups design provided evidence of construct validity (see Supplementary Materials).

Adaptive Personality Regulation. Ten comedy-relevant facets of personality formed the basis of the observations that were used to develop the APR index. To identify these, we interviewed the 12 subject matter experts noted above. The 30 NEO-PI-R (Revised NEO Personality Inventory; Costa & McCrae, 1992) facets were described to interviewees. In response to each, interviewees were asked to indicate whether the facet (a) is observable when comedians perform and (b) impacts comedians' effectiveness. Ten facets met both criteria: Angry Hostility, Straightforwardness, Self-consciousness, Assertiveness, Ideas, Compliance, Self-discipline, Anxiety, Deliberation, and Impulsiveness.

For the observations that formed the basis of the APR index, the poles of each facet were defined and combined with a 10-point scale where 1 denoted the low end of the facet and 10 denoted the high end. Each facet together with its response scale was presented twice. In response to the first, the two experts were asked to indicate the level that was required for the comedians' performance, to achieve as high a level of success as possible. In response to the second, they rated the comedians' expressed behavior in relation to the facet. The reliability of the mean ratings of situational requirements with respect to the 10 facets was 0.86 across the two sets of expert ratings. To control for the possibility that, requiring the two experts to rate both required and observed behavior, may have created a method artifact, we obtained a further sample of 11 professional comedians who provided independent ratings of required behavior. The means of the two sets of ratings correlated at .99. As the original ratings were recorded immediately after watching the videos of stage performances (about 180 times), and were, therefore, not subject to memory biases, we used these ratings.

The expert ratings were used to calculate the APR index, in line with Equation 1, essentially adopting the same scoring approach as described in Study 1, but with one additional step because we assessed adaptive personality regulation across each of the Big 5. First, we subtracted the mean ratings of expressed personality from ratings of the optimal level of expressed personality, for each personality item. This generated item-level APR indices. We then took the modulus (to remove the signs) of the item-level APR indices and then the mean to create domain-level APR indices for each of the big five. Finally, we averaged the APR indices from each of the Big 5 to create a total APR index.

The optimal personality expression for the comedy relevant traits identified by our comedy experts, on a scale of 1 to 10, were: angry hostility 3.67, self-consciousness 2.58, anxiety 3.00, and impulsivity 3.41, for neuroticism facets; assertive 7.58, for extraversion faects; intellectual curiosity 7.54, for openness facets; straightforwardness 4.55 and compliance 4.65, for agreeableness facets; and self-discipline 6.95, deliberation 4.57, for conscientiousness facets. In short, successful stage performances are generally low in expression of neurotic traits, high in assertiveness, high in self-discipline, and about average in agreeableness. These traits are

Table 8. Standardized Factor Loadings on Adaptive Personality Regulation.

Factor/facet	Factor loading
Neuroticism	.690 [.575, .805]
Extraversion	.664 [.545, .783]
Openness-to experience	.340 [.182, 499]
Agreeableness	.751 [.656, .846]
Conscientiousness	.574 [.462, .686]

Note. N = 206.

similar to those of successful actors (Nettle, 2006) and differ markedly from comedians' mean-level personalities (Irwing et al., 2020). In addition, successful comedic performances were judged to exhibit comparatively high levels of intellectual curiosity, which no doubt contributes to the surprise and novelty characteristic of successful comedy. In Study 1, optimal performance in each scenario required maximal expression of the personality trait. In Study 2, the comedians could express either too little or too much of the relevant trait. In consequence, the match scores that comprised the overall APR index diverged substantially from simple linear measures of trait expression.

Analysis Strategy. Given the similarity in hypotheses, the analysis strategy essentially duplicated that of Study 1.

Results

Correlations and descriptive statistics for all Study 2 variables are shown in Supplemental Materials, Table S2.

H2. The APR Index Should Conform to a Single-Factor Model. First, we tested whether the APR index conformed to a single, higher-order factor model. In contrast to Study 1 (which examined adaptive personality regulation exclusively in regard to Extraversion), we assessed adaptive personality regulation across each of the Big 5, because the expert judges expected that facets from each would be important for performance. We entered all five averaged match scores into a single-factor CFA. The SRMR and CFI suggested excellent fit and the TLI and RMSEA were indicative of acceptable fit (χ^2 = 19.16, *df* = 5, *p* = .001, CFI = .94, TLI = .87, RMSEA = .117 [.065–.175], SRMR = .040). Table 8 includes the factor loadings from this model, which range from .34 to .75. The McDonald's Omega of .70 and the average variance extracted of 33% suggest the model is reliable (Widaman et al., 2011). As in Study 1, these findings are broadly supportive of the construct, its indicators, and the presence of a higher-order factor of adaptive personality regulation.

H3. The APR Index Should Show Discriminant Validity With Respect to Mean-Level Personality, and Self-Monitoring. To examine the degree to which the APR index was distinct from

Table 9. Correlations, and Square Root of AVE, for Adaptive Personality Regulation With Self-Monitoring, and Mean-Level Personality.

Correlated construct	r_I (95% CI)	$\sqrt{AVE_1}$	$\sqrt{AVE_2}$
Self-monitoring	.010 [−.23, 25]	.58	.34
Mean-level openness-to-experience	.01 [−.03, .06]	.58	..47
Mean-level conscientiousness	−.02 [−.08, .05]	.58	.56
Mean-level extraversion	.00 [−.07, .07]	.58	.62
Mean-level agreeableness	−.01 [−.07, .05]	.58	.49
Mean-level neuroticism	.08 [−.00, .15]	.58	.65

Note. N = 202–232 for the r_I, and 226–230 for the AVE. AVE = average variance extracted; CI = confidence interval.

self-monitoring and each of the Big 5, we again assessed the magnitude of the correlations, the 95% confidence interval for the correlations, and the magnitude of the correlations relative to the square root of the average variance extracted. The discriminant validity evidence convincingly supported the uniqueness discriminant validity of the APR index (see Table 9).

H4. The APR Index Will Show Concurrent and Incremental Prediction of Task Performance Compared With Mean-Level Personality and Self-Monitoring. We examined the extent to which the APR index offered concurrent and incremental prediction of three dimensions of job performance (i.e., quality of delivery, material, and stagecraft) after controlling for the effects of self-monitoring and mean-level Big Five scores.

First, we tested whether the five mean-level personality dimensions and self-monitoring each showed concurrent prediction with respect to the three measures of stage performance, which were modeled as latent variables. All but two effects (Neuroticism as a predictor of Material and Stagecraft) were negligible, and collectively, all six predictors explained between 4.7% and 8.3% of the variance in comedians' performance (see Table 10, Models 2.1–2.3). When the APR Index was added to these equations, its beta coefficients were substantial and significant, adding between 14% and 35.5% incremental variance (Models 2.4–2.6).

Additional Exploratory Analysis: The Relative Magnitude of the Concurrent Effects of Separate Big 5 Components of the APR Index. Our a priori hypotheses regarded the APR index as a general factor. However, during the review process, we were asked to conduct some interesting exploratory analyses to examine whether regulation in some trait domains was more important than others. Accordingly, we regressed each of the three comedic performance criteria onto APR indices calculated for each Big 5 domain using both simple and multiple regression. The multiple regression analyses showed that the APR indices of openness and neuroticism had unique effects on quality of material, APR extraversion had a unique effect on delivery, and the APR indices of openness, extraversion, and neuroticism had unique effects on stagecraft (see Table 11). The simple regressions showed that all but two slope

coefficients were significant with the largest average effect sizes for extraversion followed by openness, neuroticism, conscientiousness and agreeableness, in descending order of magnitude. Thus, although adaptive regulation does appear to generalize across traits (i.e., conform to a general factor), the predictive properties of the trait domains vary and prediction is enhanced when we examine trait-specific regulation. These results resemble well-established effects in mean-level personality research, namely, that predictive power increases as personality assessment becomes more nuanced (Hughes & Batey, 2017).

Additional Exploratory Analysis: The APR Index Versus the General Factor of Personality Expression. When presenting related work at the 2019 World Conference of Personality, we were asked whether the APR index was meaningfully distinct from the general factor of personality expression, which emerges from assessment center data (see Breil & Back, 2019), a question that we had not thought of a priori. Now aware of this possibility, we conducted a post hoc exploratory analysis to examine the relationship between the APR index and general factor of personality expression.

First, we examined the nature of the general factor of personality expression using CFA. In contrast to the APR index, which showed acceptable to good levels of model fit, the general factor of personality expression failed to fit by a substantial margin ($\chi^2 = 142.38$, $df = 5$, $p = .001$, CFI = .385, TLI = −.230, RMSEA = .359 [.310, .411], SRMR = .171). Given this remarkable lack of fit, it is questionable whether the general factor of personality expression represents much more than a statistical artifact. Indeed, it is difficult to imagine what construct is represented by the common variance of personality expression within any given situation. Nevertheless, we continued to estimate a correlation between the APR index and the general factor of personality expression. The two correlated at $r = .466$ [.361, .572], sharing approximately 21.7% of variance. Thus, although related because they are calculated using the same data, they are by no means identical. It is clear from the fact that summed scores of personality expression correlate only weakly with the APR index that in the current study there is an unambiguous empirical distinction, and theoretical distinction,

Table 10. Regression of Quality of Delivery on Self-Monitoring, Mean-Level Personality, and Adaptive Personality Regulation.

	Comedian performance											
	Model 2.1 Material		Model 2.2 Delivery		Model 2.3 Stagecraft		Model 2.4 Material		Model 2.5 Delivery		Model 2.6 Stagecraft	
Variables	β	95% CI	β	95% CI	β	95% CI	β	95% CI	β	95% CI	β	95% CI
Self-monitoring	.125	[-.025, .275]	-.125	[-.315, .065]	.104	[-.082, .228]	.110	[-.031, .252]	-.127	[-.298, .044]	.124	[-.015, .263]
Mean-level openness	.005	[-.137, .146]	.075	[-.103, .252]	-.104	[-.218, .062]	-.001	[-.127, .125]	.053	[-.106, .212]	-.053	[-.184, .078]
Mean-level conscientiousness	.023	[-.133, .179]	.070	[-.119, .259]	.058	[-.053, .209]	.046	[-.091, .182]	.099	[-.067, .266]	.014	[-.119, .147]
Mean-level extraversion	.039	[-.155, .232]	.134	[-.062, .331]	-.174	[-.297, .072]	-.067	[-.240, .107]	.085	[-.108, .279]	-.054	[-.230, .122]
Mean-level agreeableness	.064	[-.067, .195]	-.019	[-.205, .166]	.072	[-.155, .143]	.061	[-.056, .178]	-.028	[-.205, .150]	.063	[-.069, .195]
Mean-level neuroticism	.208	[.033, .383]	.156	[-.045, .356]	**-.176**	**[-.316, -.009]**	.083	[-.073, .239]	.088	[-.098, .274]	-.070	[-.231, .078]
APR index							**.495**	**[.381, .609]**	**.377**	**[.237, .517]**	**.612**	**[.502, .721]**
R^2	.047		.060		.083		**.280**		**.203**		**.438**	
ΔR^2							**.233**		**.140**		**.355**	

Note. $N = 198$. Effects significant at the .05 level are highlighted in bold. β = beta coefficient; CI = confidence interval; APR = adaptive personality regulation.

13

Table 11. Simple and Multiple Regression of Quality of Performance on Big 5 Domain APRs.

| | Comedian performance | | | | | |
| | Model 2.7 Material | | Model 2.8 Delivery | | Model 2.9 Stagecraft | |
Variables	β	95% CI	β	95% CI	β	95% CI
Simple regression						
APR Openness	**.434**	**[.316, .551]**	.084	[−.079, .248]	**.302**	**[.280, .550]**
APR Conscientiousness	**.298**	**[.164, .431]**	**.127**	**[.063, .317]**	**.287**	**[.127, .447]**
APR Extraversion	**.315**	**[.193, .436]**	**.569**	**[.460, .678]**	**.587**	**[.487, .688]**
APR Agreeableness	**.180**	**[.045, .314]**	.151	[−.033, .324]	**.333**	**[.184, .482]**
APR Neuroticism	**.390**	**[.271, .510]**	**.243**	**[.085, .402]**	**.323**	**[.302, .568]**
Multiple regression						
APR Openness	**.324**	**[.200, .448]**	.182	[−.255, .085]	**.229**	**[.096, .363]**
APR Conscientiousness	.129	[−.022, .280]	−.115	[−.295, .066]	−.021	[−.174, .131]
APR Extraversion	.108	[−.038, .253]	**.667**	**[.516, .818]**	**.455**	**[.316, .593]**
APR Agreeableness	−.146	[−.294, .002]	−.192	[−.380, .004]	−.050	[−.155, .127]
APR Neuroticism	**.309**	**[.171, .447]**	.182	[−.015, .348]	**.250**	**[.105, .394]**
R^2	**.309**		**.365**		**.454**	

Note. $N = 200$. CI = confidence interval; Bold = significant at .05 level; APR = adaptive personality regulation.

between expressed personality and the index of adaptive personality regulation.

Discussion

Our major research goal was to identify an index of adaptive personality regulation, which we defined as *the ability or propensity to express personality to meet situational requirements*. In two studies, we investigated the properties of the APR index, which examined the degree to which participants expressed situation-appropriate personality (as defined by expert judges). To a substantial degree, the results of both studies support our hypotheses and suggest that the APR index has value for researchers and practitioners who wish to assess individuals' propensity to regulate their personality expression to meet situational demands. We now consider each hypothesis in turn.

H1 and H2 concerned the most fundamental features of adaptive personality regulation: that individuals would shift their personality expression according to situational requirements (H1) and that the APR Index reflects a general ability to regulate personality across different situations and/or traits and thus should fit a single-factor CFA (H2). Evidence regarding the shift of personality expression across situations was best demonstrated in Study 1. Here, participants shifted from a high extraversion (memorability) to a low extraversion (scoring) task. Consistent with past research, the overall degree of shift was very large and there were considerable individual differences, with some adapting well in both tasks, some in just one, and some in none (Sauerberger & Funder, 2017). Two pieces of evidence supported H2: (a) APR index measures estimated within separate tasks correlated substantially and (b) these two within-task measures, when averaged, conformed to a well-fitting single-factor CFA model. In Study 2, we observed comedians in just one situation, so we were unable to assess the degree of personality shift across situations. However, we have previously reported a group level analysis of some of the data from the current study (Irwing et al., 2020). Specifically, we found that, on average, comedians' on-stage personality expression differed considerably from their mean-level personality, suggesting that shift does occur when performing on stage.

Evidence that adaptive regulation of personality generalizes across traits was best demonstrated in Study 2. Here, we assessed personality expression across facets from each of the Big 5 domains. Again, the trait-specific APR index measures calculated for each Big 5 domain loaded onto a single higher-order APR index factor, within a well-fitting CFA model. Taken together, the results from the two studies suggest, that even when examined using the most robust methods (i.e., CFA, which is notoriously challenging to personality measures; Booth & Hughes, 2014), that the APR index generalizes across tasks and traits. Thus, we would tentatively suggest that initial evidence supports the notion that there is a general factor of adaptive personality regulation and that the APR index captures individuals' *ability or propensity to express personality to meet situational requirements*. However, one caveat is that, in an additional analysis, we found that APRs calculated at the domain level both varied in their predictive validity and explained more criterion variance.

To consider adaptive personality regulation (the construct) and the APR index (the assessment tool) as useful additions to the literature, a minimal condition is that they are distinct from conceptually similar phenomena (H3). Across both studies, the correlations between the APR index and mean-level personality and self-monitoring were

essentially zero. This evidence, together with the ratio between the AVEs and the correlations, strongly suggested that the APR index was distinct from these other constructs. Given the theoretical similarity between self-monitoring and adaptive personality regulation, the near-zero correlation might appear surprising. However, we believe there are two plausible explanations. First, self-monitoring is assessed by self-report scales, which have questionable psychometric properties (e.g., Briggs & Cheek, 1988; Kudret et al., 2019), and as we discussed in the introduction, it seems highly likely that dynamic constructs such as adaptive personality regulation are "not amenable to standard self-report measurement formats" (Erickson, et al., p. 908). Second, the original conceptualization of self-monitoring as, "an individual's observation, regulation, and control of his or her expressive behavior and self-presentation guided by social and situational cues (Kudret et al., 2019; Snyder, 1974; Snyder & Gangestad, 1986, p. 193), is very close to our conceptualization of adaptive personality regulation. However, the nature of self-monitoring is contested. For example, Gangestad and Snyder (2000) proposed a revised definition suggesting that self-monitoring reflects tendencies ". . . to engage in, or eschew, forms of impression management tactics that involve the construction of social appearances and cultivation of social images [that] . . . relates to status-oriented impression management motives" (Gangestad & Snyder, 2000, pp. 546–547). Defined as a measure of status-oriented impression management, self-monitoring is quite far removed from a general tendency to adaptively regulate personality.

In addition, in Study 2 the APR index was distinct from the general factor of personality expression that emerges from assessment center data (see Breil & Back, 2019). Furthermore, the APR index showed acceptable to good model fit and the general factor of personality expression did not. Collectively, this evidence calls into question the adequacy and nature of the general factor of personality expression, whereas the method for calculating the APR index is readily interpretable as the extent to which personality meets situational requirements. In support of arguments within the introduction, it seems that measuring variability in the expression of personality per se is insufficient as an index of adaptive personality regulation. Instead, it is necessary to take account of whether the expressed personality matches situational requirements to distinguish between adaptive and maladaptive variation.

A second minimal condition to support the utility of adaptive personality regulation and the APR index concerns practical value. In both studies, the APR index explained more variance in task/job performance than the big five and self-monitoring combined (H4). Mean-level personality has historically demonstrated reliable but modest prediction of task and job performance (Barrick et al., 2001). The picture here was consistent, with mean-level Big Five personality traits explaining 10% to 20% of the variance in Study 1's

experimental tasks and less than 10% of the variance in comedians' job performance. Self-monitoring fared worse, with beta coefficients often close to zero. The APR index, on the contrary, explained incremental and significant proportions of variance in both tasks in Study 1 and comedian performance in Study 2. Although preliminary, these findings might offer a partial explanation for the relatively underwhelming performance of mean-level personality as a predictor of task-related job performance (Barrick et al., 2001). That is, jobs are not simple and require more than a broad, static, personality profile to explain success (Hughes & Batey, 2017). Indeed, they likely require the ability to adapt personality to a variety of task and situational demands.

Limitations

Three notable limitations to our studies must be considered when interpreting the findings. First, the data collection methods were very intensive. Both studies required expert ratings of the personality expression needed for success, ratings of expressed personality, and standard psychometric data. In general, "direct observations of behavior have remained comparatively rare in personality research, likely because they are harder to obtain for sufficiently large samples" (Mõttus et al., 2020). The sentiments of Mõttus et al. (2020) were true of this study, particularly Study 1, for which the sample was smaller than would be ideal, and although it surpassed the minimum required by a priori power analyses for some tests (e.g., the main regression model), it did not for others (e.g., the full correlation matrix). Thus, replication and extension with much larger samples is required.

Second, our current studies were confined to "strong situations" in which the appropriate course of action was well defined (Judge & Zapata, 2015). In weak situations, individuals often pursue varied goals and employ diverse strategies to achieve them. Thus, the APR index might be less useful in weak situations because a priori expert ratings of the optimal levels of expressed personality are difficult under conditions characterized by equifinality. The nature of the generalizability of the APR index within "weak(er) situations" remains an open question that future studies should investigate.

Third, our measures of task performance were generated specifically for this study. The outcomes for Study 1 (i.e., number of memorable facts recalled and number of tests correctly scored) were relatively objective and easy to quantify. In contrast, although the measure of comedians' performance performed well psychometrically, it relied upon subjective ratings. Such ratings are common in the measurement of job performance (Campbell & Wiernik, 2015) and are widely regarded as a useful metric of performance in complex domains, especially when the ratings assess specific facets of performance (Bartram, 2005; Rojon et al., 2015). Nevertheless, a more objective metric (e.g., frequency and volume of laughter/applause) might have been preferable.

Conclusion

We proposed and tested the APR index as a metric of adaptive personality regulation: *the ability or propensity to express personality to meet situational requirements*. In both studies, the APR index showed robust psychometric properties, was distinct from related concepts (e.g., mean-level personality, self-monitoring, and the general factor of personality expression), and provided incremental concurrent prediction of task/job performance. Given the promising findings and the centrality of the concept underlying adaptive personality regulation (i.e., matching personality expression to situational requirements) to many previously proposed measures, further research is warranted.

Author Contributions

P.I. contributed to study design, data analysis, manuscript writing, theory development, and reference search. C.C. contributed to study design, data collection, theory development drafted parts of manuscript, and prepared Supplementary Materials. D.J.H. contributed to study design, data analysis, manuscript writing, theory development, and prepared Supplementary Materials.

Data Availability Statement

All unique data and code are available at: https://osf.io/bp3k2/?view_only=1e5246e0276a47b3849d29149d1 a3b8b

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

ORCID iDs

Paul Irwing https://orcid.org/0000-0002-4751-2156
Clare Cook https://orcid.org/0000-0001-7805-9182
David J. Hughes https://orcid.org/0000-0002-3807-3809

Supplemental Material

Supplemental material is available online with this article.

Notes

1. The reported studies were not preregistered. All unique data and code are currently available to reviewers at the view-only link: https://osf.io/bp3k2/?view_only=1e5246e0276a47b3849d2914 9d1a3b8b
2. Difference scores have well-known limitations, which can be overcome using polynomial regression and response surface analysis (e.g., Edwards, 1993, 2001; Edwards & Parry, 1993). It is important to understand, however, that the difference equation above gives rise to re-scaled scores, that is, the difference between a random variable and a constant, not the difference between two random variables. Re-scaled scores are not subject to many of the problems associated with difference scores.

References

Adolf, J., Schuurman, N. K., Borkenau, P., Borsboom, D., & Dolan, C. V. (2014). Measurement invariance within and between individuals: A distinct problem in testing the equivalence of intra-and inter-individual model structures. *Frontiers in Psychology*, *5*. 883–897. https://doi.org/10.3389/fpsyg.2014.00883

Baird, B. M., Le, K., & Lucas, R. E. (2006). On the nature of intraindividual personality variability: Reliability, validity, and associations with well-being. *Journal of Personality and Social Psychology*, *90*, 512–527. https://doi.org/10.1037/0022-3514.90.3.512

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *Personality and Performance*, *9*(1/2), 9–30. https://doi.org/10.1111/1468-2389.00160

Bartram, D. (2005). The great eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology*, *90*(6), 1185–1203. https://doi.org/10.1037/0021-9010.90.6.1185

Bleidorn, W. (2009). Linking personality states, current social roles and major life goals. *European Journal of Personality: Published for the European Association of Personality Psychology*, *23*(6), 509–530. https://doi.org/10.1002/per.731

Booth, T., & Hughes, D. J. (2014). Exploratory structural equation modeling of personality data. *Assessment*, *21*(3), 260–271. https://doi.org/10.1177/1073191114528029

Breil, S. M., & Back, M. D. (2019, April 4). *A deeper look at the construct validity of assessment centres: Which behaviors do actually matter*? [Paper presentation]. 3rd world conference on personality, *Hanoi, Vietnam*.

Briggs, S. R., & Cheek, J. M. (1988). On the nature of self-monitoring: Problems withassessment, problems with validity. *Journal of Personality and Social Psychology*, *54*(4), 663–678. https://doi.org/10.1037/0022-3514.54.4.663

Campbell, J. P., & Wiernik, B. M. (2015). The modeling and assessment of work performance. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*, 47–74. https://doi.org/10.1146/annurev-orgpsych-032414-111427

Cervone, D., & Little, B. R. (2019). Personality architecture and dynamics: The new agenda and what's new about it. *Personality and Individual Differences*, *136*, 12–23. https://doi.org/10.1016/j.paid.2017.07.001

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Clifton, A., & Kuper, L. E. (2011). Self-reported personality variability across the social network is associated with interpersonal dysfunction. *Journal of Personality*, *79*(2), 359–390. https://doi.org/10.1111/j.1467-6494.2010.00686.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum. https://doi.org/10.4324/9780203771587

Costa, P. T., Jr., & McCrae, R. R. (1992). Revised NEO personality inventory and NEO five-factor inventory professional manual. *Psychological Assessment Resources*. https://doi.org/10.1037/t03907-000

Côté, S., Moskowitz, D. S., & Zuroff, D. C. (2012). Social relationships and intraindividual variability in interpersonal behavior: Correlates of interpersonal spin. *Journal of Personality and Social Psychology*, *102*(3), 646–659. https://doi.org/10.1037/a0025313

Day, D. V., Shleicher, D. J., Unckless, A. L., & Hiller, N. J. (2002). Self-monitoring personality at work: A meta-analytic investigation of construct validity. *Journal of Applied Psychology*, *87*(2), 390–401. https://doi.org/10.1037/0021-9010.87.2.390

Denissen, J. J. A., Ulferts, H., Ludtke, O., Muck, P. M., & Gerstorf, D. (2014). Longitudinal transactions between personality and occupational roles: A large and heterogeneous study of job beginners, stayers, and changers. *Developmental Psychology*, *50*, 1931–1942. https://doi.org/10.1037/a0036994

DeYoung, C. G. (2015). Cybernetic Big Five theory. *Journal of Research in Personality*, *56*, 33–58. https://doi.org/10.1016/j.jrp.2014.07.0040092-6566/

Edwards, J. R. (1993). Problems with the use of profile similarity indices in the study of congruences in organizational research. *Personnell Psychology*, *46*(3), 641–665.

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods*, *4*(3), 265–287.

Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression as an alternative to differenece scores in organizational research. *The Academy of Management Journal*, *36*(6), 1577–1613.

Erickson, T. M., Newman, M. G., & Pincus, A. L. (2009). Predicting unpredictability: Do measures of interpersonal rigidity/flexibility and distress predict intraindividual variability in social perceptions and behavior? *Journal of Personality and Social Psychology*, *97*(5), 893–912. https://doi.org/10.1037/a0016515

Fleeson, W. (2001). Towards a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*(6), 1011–1027. https://doi.org/10.1037/0022-3514.80.6.1011

Fleeson, W. (2004). Moving personality beyond the person-situation debate: The challenge and the opportunity of within-person variability. *Current Directions in Psychological Science*, *13*(2), 83–87. https://doi.org/10.1111/j.0963-7214.2004.00280.x

Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology.*, *97*(6), 1097–1114 https://doi.org/10.1037/a0016786

Fleeson, W., & Jayawickreme, E. (2015). Whole trait theory. *Journal of Research in Personality*, *56*, 82–92. https://doi.org/10.1037/a0016786

Fleeson, W., & Law, M. K. (2015). Trait enactments as density distributions: The role of actors, situations, and observers in explaining stability and variability. *Journal of Personality and Social Psychology*, *109*(6), 1090–1104. https://doi.org/10.1037/a0039517

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*(4), 466–491. https://doi.org/10.1037/1082-989x.9.4.466

Fournier, M. A., Moskowitz, D. S., & Zuroff, D. C. (2009). The interpersonal signature. *Journal of Research in Personality*, *43*(2), 155–162. https://doi.org/10.1016/j.jrp.2009.01.023

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*, 652–670. https://doi.org/10.1037/0033-295X.102.4.652

Funder, D. C. (2004). The personality puzzle (3rd ed.). W. W. Norton.

Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, *21*(3), 177–182. https://doi.org/10.1177/0963721412445309

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Gangestad, S. W., & Snyder, M. (2000). Self-monitoring: Appraisal and reappraisal. *Psychological Bulletin*, *126*(4), 530–555. https://doi.org/10.1037/0033-2909.126.4.530

Heller, D., Komar, J., & Lee, W. B. (2007). The dynamics of personality states, goals, and well-being. *Personality and Social Psychology Bulletin*, *33*(6), 898–910. https://doi.org/10.1177/0146167207301010

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Holman, D. J., & Hughes, D. J. (2021). Transactions between Big-5 personality traits and job characteristics across 20 years. *Journal of Occupational and Organizational Psychology*, *94*(3), 762–788. https://doi.org/10.1111/joop.12332

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparametirized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/1082-989X/98/$3.00

Hu, L. T., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hughes, D. J. (2018). Psychometric validity: Establishing the accuracy and appropriateness of psychometric measures. In P. Irwing, T. W. Booth, & D. J. Hughes (Eds.), *Wiley handbook of psychometric testing: A multidisciplinary approach to survey, scale, and test development* (pp. 751–779). John Wiley & Sons. https://doi.org/10.1002/9781118489772.ch24

Hughes, D. J., & Batey, M. (2017). Using personality questionnaires for selection. In H. W. Goldstein, E. D. Pulakos, C. Semedo, & J. Passmore (Eds.). *The Wiley Blackwell handbook of the psychology of recruitment, selection and employee retention (pp. 151–181)*. John Wiley & Sons.

Hughes, D. J., & Evans, T. R. (2018). Putting "emotional intelligences" in their place: Introducing the integrated model of affect-related individual differences. *Frontiers in Psychology*, *9*, Article 2155. https://doi.org/10.3389/fpsyg.2018.02155

Irwing, P., Cook, C., Pollet, T. V., & Hughes, D. J. (2020). Comedian's trait level and stage personalities: Evidence

for goal directed personality adaptation. *Personality and Social Psychology Bulletin*, *46*(4), 590–602. https://doi.org/10.1177/0146167219867963

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *Big Five Inventory (BFI)*. https://doi.org/10.1037/t07550-000

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). Guilford Press.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: theory and research* (2nd ed., pp. 102–138). Guilford Press.

Judge, T. A., Simon, L. S., Hurst, C., & Kelley, K. (2014). What I experienced yesterday is who I am today: Relationship of work motivations and behaviors to within-individual variation in the five-factor model of personality. *Journal of Applied Psychology*, *99*(2), Article 199. https://doi.org/10.1037/a0034485

Judge, T. A., & Zapata, C. P. (2015). The person-situation debate revisited: Effect of situation strength and trait activation on the validity of the Big Five personality traits in predicting job performance. *Academy of Management Journal*, *58*(4), 1149–1179. https://doi.org/10.5465/amj.2010.0837

Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review*, *30*(7), 865–878. https://doi.org/10.1016/j.cpr.2010.03.001

Kudret, S., Erdogan, B., & Bauer, T. N. (2019). Self-monitoring personality trait at work: An integrative narrative review and future research directions. *Journal of Organizational Behavior*, *40*(2), 193–208. https://doi.org/10.1002/job.2346

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, Article 863. https://doi.org/10.3389/fpsyg.2013.00863

Le, H., Schmidt, F. L., Harter, J. K., & Lauver, K. J. (2010). The problem of empirical redundancy of constructs in organizational research: An empirical investigation. *Organizational Behavior and Human Decision Processes*, *112*, 112–125. https://doi.org/10.1016/j.obhdp.2010.02.003

Lievens, F., Lang, W. B., De Fruyt, F., Corstjens, J., de Vijver, M., & Bledlow, R. (2018). The predictive power of people intra-individual variability across situations: Implementing whole trait theory in assessment. *Journal of Applied Psychology*, *103*, 753–771. https://doi.org/10.1037/apl0000280

MacCallum, R. C., Browne, M. W., & Cai, L. (2006). Testing differences between nested covariance structure models: Power analysis and null hypotheses. *Psychological Methods*, *11*(1), Article 19. https://doi.org/10.1037/1082-989X.11.1.19

McCabe, K. O., & Fleeson, W. (2016). Are traits useful? Explaining trait manifestations as tools in the pursuit of goals. *Journal of Personality and Social Psychology*, *110*(2), 287–301. https://doi.org/10.1037/a0039490

McNeish, D., An, J., & Hancock, G. R. (2017). The thorny relation between measurement quality and fit index cutoffs in latent variable models. *Journal of Personality Assessment*, *1*, 1–10. https://doi.org/10.1080/00223891.2017.1281286

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, *39*(3), 479–515. https://doi.org/10.1207/s15327906mbr3903_4

Minbashian, A., Wood, R. E., & Beckmann, N. (2010). Task-contingent conscientiousness as a unit of personality at work. *Journal of Applied Psychology*, *95*(5), 793–806. https://doi.org/10.1037/a0020016

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*(2), 246–268. https://doi.org/10.1037/0033-295x.102.2.246

Moskowitz, D. S., & Zuroff, D. C. (2004). Flux, pulse, and spin: Dynamic additions to the personality lexicon. *Journal of Personality and Social Psychology*, *86*(6), 880–893. https://doi.org/10.1037/0022-3514.86.6.880

Moskowitz, D. S., & Zuroff, D. C. (2005). Robust predictors of flux, pulse, and spin. *Journal of Research in Personality*, *39*, 130–147. https://doi.org/10.1016/j.jrp.2004.09.004

Mõttus, R., Wood, D., Condon, D. M., Back, M. D., Baumert, A., Constatini, G., Epskamp, S., Greiff, S., Johnson, W., Lukazewski, A., Murray, A., Revelle, W., Wright, A., Yarkoni, T., Zeigler, M., & Zimmerman, J. (2020). Descriptive, predictive and explanatory personality research: Different goals, different approaches, but a shared need to move beyond the Big Few traits. *European Journal of Personality*, *33*(6), 1175–1201. https://doi.org/10.1002/per.2311

Nestler, S., & Back, M. D. (2017). Using cross-classified structural equation models to examine the accuracy of personality judgments. *Psychometrika*, *82*(2), 475–497. https://doi.org/10.1007/s11336-015-9485-6

Nettle, D. (2006). Psychological profiles of professional actors. *Personality and Individual Differences*, *40*(2), 375–383. https://doi.org/10.1016/j.paid.2005.07.008

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776–783. https://doi.org/10.1037/h0043424

Parrigon, S., Woo, S. E., Tay, L., & Wang, T. (2017). CAPTION-ing the situation: A lexically derived taxonomy of pcyhological situation characteristics. *Journal of Personality and Social Psychology*, *112*(4), 642–681. http://dx.doi.org/10.1037/pspp0000111

Paulhus, D. L., & Martin, C. L. (1988). Functional flexibility: A new conception of interpersonal flexibility. *Journal of Personality and Social Psychology*, *55*(1), 88–101. https://doi.org/10.1037/0022-3514.55.1.88

Perunovic, W. Q. E., Heller, D., Ross, M., & Komar, S. (2011). The within-person dynamics of intrinsic and extrinsic motivation, affective states, and cultural identification: A diary Study of bicultural individuals. *Social Psychological and Personality Science*, *2*, 635–641. https://doi.org/10.1177/1948550611405071

Preacher, K. J., & Coffman, D. L. (2006). *Computing power and minimum sample size for RMSEA* [Computer software]. http://quantpsy.org/

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., & Funder, D. C. (2014). The situational eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology*, *107*(4), 677–718. https://doi.org/10.1037/a0037250

Rojon, C., McDowall, A., & Saunders, M. N. (2015). The relationships between traditional selection assessments and workplace performance criteria specificity: A comparative meta-analysis. *Human Performance*, *28*(1), 1–25. https://doi.org/10.1080/08959285.2014.974757

Sauerberger, K. S., & Funder, D. C. (2017). Behavioral change and consistency across contexts. *Journal of Research in Personality*, *69*, 264–272. https://doi.org/10.1016/j.jrp.2016.04.007

Sauerberger, K. S., & Funder, D. C. (2020). The Riverside Situational Q-sort. In J. F. Rauthmann, R. A. Sherman, & D. C. Funder (Eds.), *The Oxford handbook of psychological situations* (pp. 287–298). Oxford University Press

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Shaffer, J. A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation: A guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods*, *19*(1), 80–110. https://doi.org/10.1177/1094428115598239

Sherman, R. A., Rauthmann, J. F., Brown, N. A., Serfass, D. G., & Jones, A. B. (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of Personality and Social Psychology*, *109*, 872–888. https://doi.org/10.1037/pspp0000036

Snyder, M. (1974). Self–monitoring of expressive behavior. *Journal of Personality and Social Psychology*, *30*, 526–537. https://doi.org/10.1037/h0037039

Snyder, M., & Gangestad, S. (1986). On the nature of self-monitoring: Matters of assessment, matters of validity. *Journal of Personality and Social Psychology*, *51*(1), 125–139. https://doi.org/0022-3514/86/$00.75

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, *88*(3), Article 500. https://doi.org/10.1037/0021-9010.88.3.500

Widaman, K. F., Little, T. D., Preacher, K. J., & Sawalani, G. M. (2011). On creating and using short forms of scales in secondary research. In K. H. Trzesniewski, M. B. Donnellan, & R. E. Lucas (Eds.), *Secondary data analysis: An introduction for psychologists* (pp. 39–61). https://doi.org/10.1037/12350-003.

Wirth, R. J. & Edwards, M. C. (2007). Item factor analysis: Current approaches and future directions. *Psychological Methods*, *12*(1), 58–79. https://doi.org/10.1037/1082-989x.12.1.58

Wood, R. E., Beckmann, N., Birney, D. P., Beckmann, J. F., Minbashian, A., & Chau, R. (2019). Situation contingent units of personality at work. *Personality and Individual Differences*, *136*, 113–121. https://doi.org/10.1016/j.paid.2018.01.026

Wright, J. C., & Mischel, W. (1987). A conditional approach to dispositional constructs: The local predictability of social behavior. *Journal of Personality and Social Psychology*, *53*(6), 1159–1177. https://doi.org/10.1037/0022-3514.53.6.1159

Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. In M. E. Sobel & M. P. Becker (Eds.), Sociological methodology 2000 (pp. 165–200). ASA.

Zeigler-Hill, V., Li, H., Masri, J., Smith, A., Vonk, J., Madson, M. B., & Zhang, Q. (2013). Self-esteem instability and academic outcomes in American and Chinese college students. *Journal of Research in Personality*, *47*, 455–463. https://doi.org/10.1016/j.jrp.2013.03.010