Spring 2010

The hierarchical structure of emotional expressivity: scale development and nomological implications

John Jeffrey Humrichouse
University of Iowa

Recommended Citation
Humrichouse, John Jeffrey. "The hierarchical structure of emotional expressivity: scale development and nomological implications." PhD (Doctor of Philosophy) thesis, University of Iowa, 2010.
https://doi.org/10.17077/etd.1m2x7c8p

THE HIERARCHICAL STRUCTURE OF EMOTIONAL EXPRESSIVITY:

SCALE DEVELOPMENT AND NOMOLOGICAL IMPLICATIONS

by

John Jeffrey Humrichouse

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

May 2010

Thesis Supervisor: Professor David Watson

ABSTRACT

Integrating existing models of emotional expressivity, the 3-level hierarchical model contains a general factor of emotional expressivity vs. inexpressivity at the highest level; relatively independent factors of positive and negative expressivity at the second-order level; and discrete expressivity factors of sadness, hostility, guilt/shame, fear, joviality, confidence and amusement at the lowest level. The bottom-up analytic strategy consisted of identifying first the structure of the discrete affects; subsequent second-order factor analyses supported the existence of the higher order factors. The Iowa Scales of Emotional Expressivity (ISEE)—a hierarchical set of scales—systematically incorporate the level of abstraction of the items to assess each level of the hierarchy. Structural analyses replicated across college student ($N = 387$) and young adult ($N = 344$) samples with strong comparability coefficients. Striking differences existed in comparisons of the nomological relations of the general factor level vs. second-order level—Positive and Negative Expressivity demonstrated differential relations with Extraversion and Neuroticism and incremental predictive validity beyond Positive and Negative Affect, respectively. The ISEE demonstrated convergent and discriminant validity with existing scales and through multi-trait multi-method analyses of self-other agreement and test-retest data. Although test-retest correlations were less than optimal, the ISEE improve upon existing measures of emotional expressivity by extending the assessment to the discrete affect level and by creating Positive and Negative Expressivity scales with improved discriminant validity and clearer differential relations.

Abstract Approved: ____
Thesis Supervisor

____
Title and Department

____
Date

THE HIERARCHICAL STRUCTURE OF EMOTIONAL EXPRESSIVITY:
SCALE DEVELOPMENT AND NOMOLOGICAL IMPLICATIONS

by

John Jeffrey Humrichouse

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

May 2010

Thesis Supervisor: Professor David Watson

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

John Jeffrey Humrichouse

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Psychology at the May 2010 graduation.

Thesis Committee: ___________________________
David Watson, Thesis Supervisor

____
Lee Anna Clark

____
Kristian Markon

____
Jerry Suls

____
Walter Vispoel

ABSTRACT

Integrating existing models of emotional expressivity, the 3-level hierarchical model contains a general factor of emotional expressivity vs. inexpressivity at the highest level; relatively independent factors of positive and negative expressivity at the second-order level; and discrete expressivity factors of sadness, hostility, guilt/shame, fear, joviality, confidence and amusement at the lowest level. The bottom-up analytic strategy consisted of identifying first the structure of the discrete affects; subsequent second-order factor analyses supported the existence of the higher order factors. The Iowa Scales of Emotional Expressivity (ISEE)—a hierarchical set of scales—systematically incorporate the level of abstraction of the items to assess each level of the hierarchy. Structural analyses replicated across college student ($N = 387$) and young adult ($N = 344$) samples with strong comparability coefficients. Striking differences existed in comparisons of the nomological relations of the general factor level vs. second-order level—Positive and Negative Expressivity demonstrated differential relations with Extraversion and Neuroticism and incremental predictive validity beyond Positive and Negative Affect, respectively. The ISEE demonstrated convergent and discriminant validity with existing scales and through multi-trait multi-method analyses of self-other agreement and test-retest data. Although test-retest correlations were less than optimal, the ISEE improve upon existing measures of emotional expressivity by extending the assessment to the discrete affect level and by creating Positive and Negative Expressivity scales with improved discriminant validity and clearer differential relations.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

There are clear differences in the extent to which people express their emotions. These differences in emotional expressions have long interested researchers and are relevant to several areas of psychology such as personality (e.g., emotional expressivity as a stable trait; Gross & John, 1995; Kring, Smith, & Neale, 1994), social psychology (e.g., nonverbal communication; Riggio & Riggio, 2002), psychopathology (e.g., flat affect; Kring, Kerr, Smith, & Neale, 1993), and health psychology (e.g., immunological functioning; Berry & Pennebaker, 1998). Emotionally expressive individuals report higher levels of self-esteem, well-being, life satisfaction, social closeness and lower levels of social anhedonia (Gross & John, 1997; King & Emmons, 1990; Kring et al., 1994). Lack of emotional expression is implicated in several psychological disorders such as schizophrenia (Lee et al., 2008), depression (Sloan et al., 1997), social anxiety (Kashdan & Breen, 2008; Kashdan et al., 2007), and post-traumatic stress disorder (Tull et al., 2007). With such clear individual differences in emotional expressivity—the tendency to express affective states through nonverbal means—and the broad relevance of emotional expressivity to the field of psychology, the importance of reliably and validly assessing this domain should be obvious.

Emotional expressivity has been modeled at different levels of abstraction including at a general factor level, a dimensional level, and a discrete emotions level; however, currently there is only limited support for a hierarchical model that integrates these different levels of abstraction. The importance of recognizing the existence of different levels of abstraction within a hierarchical structure and theoretical and practical issues related to developing hierarchical assessments are addressed in greater detail in Chapter 1. Certainly, a clear hierarchical model and a hierarchical assessment that allows researchers to assess emotional expressivity at different levels of abstraction would be beneficial to assessing the domain in a valid manner.

A review of existing measures of emotional expressivity in Chapter 2 provides evidence of construct validity and, more specifically, provides evidence of internal consistency, test-retest reliability, and convergent and discriminant validity of existing measures. Two of the existing measures modeled emotional expressivity at different levels of abstraction—one measure assesses a general factor and the other assesses a hierarchical model—and these disparate conceptualizations have yet to be integrated into a clear, comprehensive hierarchical model. By building from these existing scales and models, an even more reliable and valid assessment of emotional expressivity might be achieved through this scale development project.

In Chapter 3, results from a preliminary study of a large sample of newlywed couples replicated previous structural analyses of emotional expressivity. Several problematic items in one widely used scale of emotional expressivity brought into question the structural validity of the scale and subsequently resulted in issues related to discriminant validity. In this newlywed sample, Multi-Trait Multi-Method (MTMM; Campbell & Fiske, 1959) analyses of self- and spouse-ratings simultaneously tested convergent validity and discriminant validity. Removing the problematic items and recognizing the differences in the level of abstraction of the item content led to exploratory analyses on a reduced item pool and resulted in very brief revised scales with improved discriminant validity and clearer differential relations in the nomological network. Overall, this preliminary study provided initial support for an integrated hierarchical model of emotional expressivity.

Extending this model to the discrete affect level resulted in a three-level hierarchical model of emotional expressivity including: a general factor at the highest level, positive and negative expressivity at the second-order level, and discrete affects at the lowest level. To begin to test the three-level hierarchical model, the development of the Iowa Scales of Emotional Expressivity (ISEE) followed the substantive, structural and external phases of scale development (Clark & Watson, 1995; Loevinger, 1957). The substantive phase, covered in Chapter 4, focused on conceptualizing the constructs and multiple rounds of item writing and revisions.

The structural phase, covered in Chapter 5, focused on creating scales aligned with the theoretical structural model. The general analytic strategy consisted of a series of factor analyses in a bottom-up approach. Factor analyses identified the structure of the discrete affect level and then second-order factor analyses supported the hierarchical model. The structural phase focused on creating a set of hierarchical scales aligned with each level of the hierarchical model. Selecting the strongest markers of the higher order factors maximized the amount of common variance related to the higher order factors and minimized the specific variance of the lower order factors.

An examination of the nomological network, covered in Chapter 6, supported the conceptual connections of emotional expressivity with personality and trait affect. Although emotional expressivity can be conceptualized at various levels of abstraction, the nomological network explicates the validity of these constructs at different levels of abstraction. Overall, the interrelations between Positive Expressivity (PE), Positive Affect (PA) and Extraversion (E)—as well as the interrelations between Negative Expressivity (NE), Negative Affect (NA) and Neuroticism (N)—strongly supported the validity of the second-order level of emotional expressivity.

The external phase of scale development, covered in Chapter 7, addressed convergent and discriminant validity. Strong correlations between the ISEE and existing measures of emotional expressivity supported convergent validity. MTMM analyses of self- and other-ratings simultaneously tested convergent and discriminant validity. Importantly, the lower levels of the proposed hierarchical model demonstrated both convergent and discriminant validity, that is, convergent correlations exceeded discriminant correlations even though discriminant correlations within a hierarchical structure are expected to be significant.

Analyses of short-term test-retest reliability, covered in Chapter 8, revealed that the ISEE have less than optimal dependability. Short-term test-retest reliability, termed dependability, minimizes the potential for true change and focuses on the estimation of measurement error (Cattell, Eber, & Tatsuoka, 1970; Chmielewski & Watson, 2009; Watson, 2004). To facilitate the interpretation of dependability, Watson (2004) recommended including other measures to serve as benchmarks or points of comparison. Accordingly, this dependability study also included existing scales of emotional expressivity, and conceptually related measures of trait affect and personality.

Finally, Chapter 9 covers a general discussion of structural validity, the nomological network, convergent and discriminant validity, and test-retest reliability. The strengths and limitations of the current research are discussed and future directions are proposed. The three-level hierarchical model of emotional expressivity links the general factor level to the discrete affect level and was conceptualized within a well-articulated nomological network. The significance of systematically accounting for the level of abstraction of the items to target specific constructs at different levels of a hierarchy has implications for extending the lower order structure of personality. The ISEE allow researchers to select a brief scale of general emotional expressivity, two brief scales assessing positive and negative expressivity or multiple discrete affect scales to meet their specific research needs; alternatively, including all of the scales allows for direct comparisons across the hierarchical levels of emotional expressivity. Creating scales and demonstrating the reliability and validity of scales is an iterative process, such that this research is just the beginning of the process of validating the ISEE.

CHAPTER 1. HIERARCHICAL MODELS AND ASSESSMENTS

In general, hierarchical models allow for a domain to be modeled at various levels of breadth, while integrating these various levels of abstraction and providing structure to a domain. Currently, there is mixed support for hierarchical models of emotional expressivity that integrate these different levels of abstraction (Dobbs et al., 2007; Gross & John, 1998; Trierweiler et al., 2002). An integrated hierarchical structure would facilitate establishing interconnections between other hierarchical models. To provide a framework for understanding the subsequent reviews of existing scales and to facilitate the integration of prior research, this introduction discusses existing hierarchical structures.

Hierarchical Models

In the field of personality, the universal hierarchical model of personality links the basic factors of personality at different levels of abstraction and integrates previous models of normal and abnormal personality into a single model (Markon, Krueger, & Watson, 2005). Although descriptions of the basic factors at each level of abstraction are beyond the scope of this introduction, a general description of the model includes: (a) two factors similar to α and β (Digman, 1997) at the highest level of abstraction; (b) at the next level, α splits to create the Big Three model of Negative Emotionality, Disinhibition, and Positive Emotionality; (c) then Disinhibition splits into Unconscientious Disinhibition and Disagreeable Disinhibition, but Negative and Positive Emotionality remain very similar at the four-factor level; (d) at the lowest level, Positive Emotionality splits into Extraversion (E) and Openness (O) and the other factors of the Big Five model, Neuroticism (N), Conscientious (C) and Agreeableness (A) generally remain the same.

Beyond identifying the basic factors at different levels of the hierarchy (i.e., Big Two, Big Three, Big Five) this hierarchical model identifies replicable relations between the factors and emphasizes that factors exist at different levels of abstraction. More specifically, Negative Emotionality remains very similar across the three-, four- and five-factor levels of abstraction and Positive Emotionality remains very similar across the three- and four-factor levels of abstraction. This important contribution allows for the integration of research based on these differing levels of the hierarchical model. Ultimately, the universal model of personality established a framework among the broadest factors of personality to explicate the nomological network among them.

Similarly, a three-level hierarchical model of affect links the basic factors of affect at different levels of abstraction and integrates previous models into a single scheme (Tellegen, Watson, & Clark, 1999a, 1999b; Watson & Tellegen, 1999). At the highest level is a general bipolar factor of happiness; the next level consists of two relatively independent dimensions of Positive Affect (PA) and Negative Affect (NA), and the lowest level consists of multiple specific affects (joy, fear, sadness, anger, etc.). The key to this hierarchical structure is the recognition that the general bipolar factor of happiness and independent dimensions of PA and NA are better viewed as different levels of abstraction within a hierarchical model, rather than as competing models at the same level of abstraction. At the highest level of this model the general bipolar factor of happiness accounts for the tendency for PA and NA to be moderately negatively correlated. Therefore, the hierarchical model of affect accounted for both the bipolarity of pleasantness-unpleasantness and the independence of PA and NA, effectively resolving a debate that occupied the literature for decades (Watson & Tellegen, 1999).

Both the universal model of personality and the three-level hierarchical model of affect are conceptually relevant to the development of a hierarchical model of emotional expressivity. Although the breadth of emotional expressivity is narrower than the factors of the various levels of the universal model of personality, emotional expressivity has conceptual connections to the broad factors of E (or Positive Emotionality) and N (or Negative Emotionality). On the other hand, emotional expressivity could exist at similar levels of abstraction as the three-level hierarchical model of affect. In this framework both emotional expressivity and trait affect could be viewed as related components encompassed by the broader factors of personality.

Building from existing research and modeling emotional expressivity at similar levels of abstraction as the three-level hierarchical model of affect can facilitate efforts to model emotional expressivity. In particular, at the discrete affect level there are likely to be strong structural similarities between the models of emotional expressivity and trait affect. For example, the existence of a construct covering feelings of sadness corresponds with the existence of a construct covering expressions of sadness. At higher levels of abstraction the structural similarities between affect and emotional expressivity are less certain. Whereas in the three-level hierarchical model of affect, PA and NA are moderately negatively correlated (Tellegen et al., 1999a, 1999b), in the existing models of emotional expressivity—to be discussed in the next chapter—positive and negative expressivity are moderately to strongly positively correlated. The structural relations within these models and the connections between these models are important to understanding these domains and clarifying the nomological network.

Although constructs can exist at various levels of abstraction within hierarchical models, there is a trade-off between bandwidth (i.e., increased breadth or range of coverage) versus fidelity (i.e., increased accuracy of coverage). “Greater fidelity is achieved at the loss of bandwidth and increased bandwidth comes at the price of fidelity” (Hogan & Roberts, 1996, p. 627). Narrower traits, specifically matched to a criterion, increase prediction, but extremely narrow traits have restricted range of applicability and risk becoming circular explanations of behavior. Although narrower constructs are less complex and are better for predictive purpose, Funder argued, “[g]lobal traits…have real explanatory power. The recognition of a pattern of behavior is a bona fide explanation of each of the behaviors that comprise it. Indeed, the more global a trait is, the more explanatory power it has” (1991, pp. 35-36).

Within a hierarchy, people generally prefer a basic level that is relatively broad and informative compared to either a superordinate level that is broader and less informative or a subordinate level that is narrower and more informative (John, Hampson, & Goldberg, 1991). Identifying the basic level within a hierarchy requires finding the optimal balance such that a construct is as broad and as informative as possible. A direct way to compare the trade-off between bandwidth versus fidelity is to assess multiple levels of varying bandwidths within a hierarchical structure. "In general, then, research on individual differences is most fruitfully conducted simultaneously at different levels of construct breadth" (Hampson, John, & Goldberg, 1986, p. 53).

Hierarchical Assessments

Working from a clear hierarchical model, scales could be created to assess the various levels of abstraction, that is, a hierarchical assessment would assess each level of the hierarchical model. To support the hierarchical model, the scales of the lower levels must demonstrate both the common variance and specific variance at each level of abstraction. Within a hierarchical assessment the common variance among the lower level factors must be accounted for by a factor at the next highest level; however, to necessitate the existence of the lower levels of a hierarchy, these factors must also demonstrate specific variance (Costa & McCrae, 1995). In other words, without specific variance of the factors at one level of a hierarchy then only the next higher level of the hierarchy is needed to assess the common variance.

The recognition of the importance of demonstrating the common and specific variance at each level of the hierarchy has practical implications for how to create a hierarchical assessment. The NEO-PI-R (Costa & McCrae, 1992) contributed to the literature on the Big Five Model of personality by hierarchically assessing both five broad, global domains and 30 more narrowly focused facets, that is, two levels of assessment are possible by administering this instrument. More specifically, in the NEO-PI-R the six facet scales within each of five broad domains can be added together to calculate the Big Five domain scores. This method of adding the lower level scales into higher level scores combined both the common variance and the specific variances of the lower factors into the higher level scores. In this case, the specific variance of the lower scales simply added noise to the assessment of the common variance of the higher order factors. Alternatively, to minimize this concern, factor scoring weights are applied or an abbreviated assessment with the purest markers of the factors was created (i.e., the NEO-FFI; Costa & McCrae, 1992).

Alternatively, within the literature on the Big Three model of personality, the SNAP (Clark, 1993) utilized separate scales in a hierarchical assessment of three broad, global factors—Positive Temperament, Negative Temperament, and Disinhibition—and 12 primary traits. The scales assessing the lower level primary traits contain specific variance, but the separate assessments of the broader higher order factors do not include the specific variance of the primary traits. Additionally, for practical reasons, the creation of abbreviated scales assessing higher order factors minimizes the length of assessment when only the higher order factors are of interest for the specific research purpose.

Created prior to research supporting the three-level hierarchical model of affect, the Positive and Negative Affect Schedule – Expanded Form (PANAS-X) assessed two levels of affect—the higher order factors of PA and NA and 11 lower order discrete affects (Watson & Clark, 1999). Within this hierarchical assessment, several items are scored as indicators of both the lower order specific affects and the higher order factors. For example, “afraid” is scored on the specific Fear scale and on the higher order NA scale. Whereas the fear scale contains several other items related to fear (e.g., scared and nervous), the NA scale also contains a broader range of affective terms (e.g., irritable and guilty). PA and NA contain a representative mixture of the affective terms that are the purest markers of the higher order factors. This method maximized the common variance among the items while potentially minimizing the contaminating effects of the specific variance of the items. In other words, validity of the discrete affect scales depends on demonstrating specific variance unaccounted for by the higher order factors and thus the purest markers of the discrete affect scales may contain relatively greater amounts of variance unaccounted for by the higher order factors.

Creating items that specifically target different levels of the hierarchy would be a systematic way to create a hierarchical assessment. For instance, a superordinate level of emotional expressivity would have wide bandwidth and lower fidelity by including more inclusive and less descriptive items (e.g., “I am emotionally expressive”) and a subordinate level of emotional expressivity would have narrower bandwidth and higher fidelity by including more specific and descriptive items (e.g., “My anger comes out in the tone of my voice”). In relation to a broad general factor, the items of a more general nature would be expected to contain more common variance, whereas items with more specific item content would likely contain relatively less common variance and relatively more specific variance. In a hierarchical assessment, the more general items would be strong markers of the general factor and the specific items would be strong markers of the lower level factors. Therefore, scales could be created at each level of the hierarchy.

In summary, establishing the hierarchical structure of emotional expressivity within the framework of existing hierarchical models of personality and affectivity will serve to strengthen all models. Additionally, each level of the hierarchical model of emotional expressivity could be assessed with separate scales for that level (i.e., a hierarchical assessment). Next, existing measures of emotional expressivity will be reviewed, specifically examining the level of abstraction of the item content to integrate these existing models into a hierarchical structure.

CHAPTER 2. EXISTING MEASURES OF EMOTIONAL EXPRESSIVITY AND STRUCTURAL ANALYSES

Several factor analytically derived measures of emotional expressivity already exist in the literature (Gross & John, 1995; King & Emmons, 1990; Kring et al., 1994). Factor analysis is a statistical procedure that can be used, among other things, to reduce the data from multiple items to a fewer number of meaningful dimensions or factors. Factors are latent variables that are extracted based on the patterns of intercorrelations among a set of items. This section covers two types of factor analyses: exploratory factor analysis and confirmatory factor analysis. Exploratory factor analysis seeks to identify factors and provides item factor loadings, that is, a quantitative estimate of the proportion of the items' variance that is attributable to the factor. Alternatively, rather than seeking to identify factors, Confirmatory Factor Analysis (CFA) tests a hypothesized factor structure and provides quantitative fit indices based on how well the proposed structure accounts for the data. As will be reviewed towards the end of this chapter, recent factor analyses (Dobbs, Sloan, & Karpinski, 2007) have not found strong support for the internal structure of some of the existing measures.

Emotional Expressivity Scale (EES)

The EES (Kring et al., 1994) is based on a general disposition to display emotions outwardly. The EES is the only measure to be reviewed that is based solely on a general factor model of emotional expressivity without items explicitly tapping valence content or nonverbal channel specifications (e.g., facial, vocal). The EES contains 17 items (e.g., "I think of myself as emotionally expressive") with respondents indicating the extent to which the item applies to them (1 = never true and 6 = always true). Interestingly, 11 out of the 17 items of the scale are negatively keyed (e.g., "Even if I am experiencing strong feelings, I don't express them outwardly" and "I hold my feelings in"). A principal component analysis (PCA) of the EES, based on responses from 373 undergraduate students, revealed the first factor accounted for 23% of the total variance with all 17 items loading greater than |.30| on the general factor; however, the researchers did not report individual item factor loadings. A coefficient alpha of .90 supported internal consistency and 102 undergraduate students retested after a 4-week interval yielded a test-retest reliability of .90. The EES converged with other measures of emotional expressivity, correlated positively with E, N and life satisfaction, and negatively with social anhedonia. The EES failed to predict behavioral assessments of positive and negative expressivity (assessed while viewing film clips) but related to an overall composite based on behavioral assessments. A cited reference search of the Web of Science database for the original EES scale development article (Kring et al., 1994) revealed 108 citations (as of 3/22/10).

Berkeley Expressivity Questionnaire (BEQ)

The BEQ (Gross & John, 1995) is based on a two-stage model of emotions consisting of an emotional impulse in the first stage and then emotionally expressive behaviors in the second stage. The BEQ assesses emotional expressivity as a hierarchical model including facets of Impulse Strength (e.g., “I have strong emotions”), Positive Expressivity (e.g., “When I’m happy, my feelings show”) and Negative Expressivity (e.g., “It is difficult for me to hide my fear”). This self-report measure contains 16 items, with an agreement response format (1= strongly disagree to 7 = strongly agree), scored as a total scale or separately for the individual facets. Development of the BEQ utilized three large undergraduate samples, a derivation sample ($N = 470$) and two replication samples (Ns = 394 and 528). Based on the derivation sample, PCA revealed the first factor accounted for 33% of the total variance and all 16 items loaded |.30| or greater on the general factor. For the BEQ Total scale, coefficient alpha greater than .80 and an average interitem correlation (AIC) of approximately .25 supported internal consistency and 68 undergraduate students retested after a 2- to 3-month interval yielded a test-retest reliability of .86.

A Varimax-rotated PCA three-factor solution accounted for 51% of the total variance—Impulse Strength, Negative Expressivity and Positive Expressivity factors accounted for 33%, 10%, and 9% of the variance, respectively. Most items loaded greater than |.40| on the designated factor; however, the researchers failed to report cross-loadings. In three samples, the facets of the BEQ demonstrated only moderate internal consistency: (a) Impulse Strength (6-items, αs = .73 - .78), (b) Negative Expressivity (6-items, αs = .68 - .72), and Positive Expressivity (4-items, αs = .65 - .71). All three facets intercorrelated approximately .50; these associations become substantially higher if disattenuated to correct for unreliability. The BEQ Total scale converged with other measures of emotional expressivity, correlated positively with E and N, and negatively with emotional control. Emotional expressivity corresponded with affective experience, that is, Positive and Negative Expressivity scales differentially related to PA and NA; however, NA was more strongly related to Impulse Strength than to Negative Expressivity. The Positive and Negative Expressivity scales differentially predicted behavioral assessments of positive and negative expressions in a film-viewing paradigm. A cited reference search of the Web of Science database for the original BEQ scale development article (Gross & John, 1995) revealed 61 citations (as of 3/22/10).

Emotional Expressiveness Questionnaire (EEQ)

The EEQ (King & Emmons, 1990) assesses expressiveness more generally rather than limiting the construct to nonverbal expressiveness. Based on a hierarchical model, the EEQ includes facets of expression of positive emotion (e.g., “I laugh a lot”), expression of negative emotion (e.g., “When I am angry people around me usually know”), and expression of intimacy (e.g., “I often tell people that I love them”). This measure contains 16 items, with an agreement response format (1 = do not agree to 7 = agree), scored as a total scale or separately for the individual facets. Based on a sample of 299 undergraduate students, a Varimax-rotated PCA three-factor solution—Expression of positive emotion (7-items), Expression of intimacy (4-items), and Expression of negative emotion (5-items)—accounted for 14%, 12%, and 10% of the total variance, respectively. Factor loadings on the designated factors ranged from .74 - .24 and four items cross-loaded greater than .30. The EEQ total scale demonstrated moderate internal consistency with a coefficient alpha of .78; however, in another study the EEQ Total scale and facets demonstrated lower internal consistency with coefficient alphas ranging from .53 to .69 (Gross & John, 1998). The EEQ converged with other measures of emotional expressivity and correlated negatively (r = -.24) with the Ambivalence Over Emotional Expressiveness Questionnaire (AEQ). A cited reference search of the Web of Science database for the original EEQ scale development article (King & Emmons, 1990) revealed 182 citations (as of 3/22/10); however, the article primarily focused on the AEQ development.

Summary of Existing Scales

The creation of the EES, BEQ and EEQ helped to establish the construct validity of emotional expressivity and provided support for the validity of self-report assessments of this construct. The EES and BEQ are the most extensively used measures and have the most validating support. The validity of the EES is supported by evidence of internal consistency, strong test-retest reliability, and convergence with other measures; however, the EES only assesses emotional expressivity based on a general factor model and does not assess separate dimensions of positive and negative expressivity. Accordingly, the EES item content is more general and abstract and only targets the assessment of a general factor of emotional expressivity. Alternatively, the BEQ item content is a mixture of general, intermediate and specific abstractions and all items are indicators of both individual facets and the general factor. Although strong test-retest reliability and convergence with other measures of the construct supported the validity of the BEQ, the facets of Impulse Strength and Positive and Negative Expressivity demonstrated only moderate internal consistency and the researchers failed to report the cross-loadings. The EEQ and its facets also demonstrated lower internal consistency and the researchers only reported limited evidence of validation. The next section will further examine the structure of emotional expressivity, primarily focusing on the EES and BEQ.

Structure of Emotional Expressivity

On the basis of a broad conceptualization of expressivity, Gross and John (1998) conducted a structural analysis of six expressivity-related questionnaires including: the EES, the BEQ, the EEQ, the Self-Monitoring Scale (SMS; Snyder, 1974), the Affect Communication Test (ACT; Friedman et al., 1980), and the Affect Intensity Measure (AIM; Larsen & Diener, 1987). Out of six expressivity measures, the EES and BEQ correlated the strongest ($r = .78$; Gross & John, 1998). This strong correlation between a general factor scale (i.e., EES) and the total score of a multifaceted scale (i.e., BEQ) provided general support for a hierarchical model. Analyses designed to identify facets of expressivity omitted the EES because it only assesses a general factor of emotional expressivity.

From the other five faceted scales (83 items) a Varimax-rotated PCA revealed five factors: Expressive Confidence, Positive Expressivity, Negative Expressivity, Impulse Intensity, and Masking. A three-level hierarchical model fit the data the best (CFI = .90, $\chi^2/df = 3.6$). The highest level contained a general factor of expressivity; the second-order contained Expressive Confidence, Masking, and Core Emotional Expressivity and then, at the lowest level, Core Emotional Expressivity further broke down into Positive Expressivity, Negative Expressivity and Impulse Intensity. Similar to the three-facet structure of the BEQ, the three facets of Core Emotional Expressivity correlated with a mean of .41.

The three facets of Core Emotional Expressivity only correlated weakly with Expressive Confidence and Masking. The Expressive Confidence "factor describes individuals who have great confidence in their expressive skills, feel comfortable performing in public and at social gatherings, and enjoy being in the limelight" (Gross & John, 1998, p. 175). The Masking factor "involve[s] perceived discrepancies between the inner experience and outer expression of emotion or attempts at masking the expression of one's inner feelings for self-presentational purposes" (Gross & John, 1998, p. 175). These weak correlations with Expressive Confidence and Masking identified the outer limits of the construct of *emotional* expressivity. In fact, very few items from Expressive Confidence and Masking explicitly contained content tapping specific emotions or emotions and feelings more generally. Furthermore, the items of the Self-Monitoring Scale (Snyder, 1974)—which is conceptually the least similar to emotional expressivity as outlined for this proposal—primarily defined the Expressive Confidence and Masking facets in this model.

The structural analyses failed to include a test of the reduced model of Core Emotional Expressivity—comprised of Positive Expressivity, Negative Expressivity and Impulse Intensity—as a separate model in these data (Gross & John, 1998). Additionally, Gross and John (1997) in a previous study that tested the BEQ hierarchical model, with three facets of Positive Expressivity, Negative Expressivity and Impulse Strength, omitted the relevant fit indices. Ultimately, these structural analyses of existing expressivity questionnaires resulted in the creation of scales by selecting items from the five different scales; however, these new scales demonstrated only moderate internal consistency. The Negative Expressivity scale (11 items, $\alpha = .72$, AIC .19) contained items tapping negative terms such as: angry, disappointment, fear, upset, pity, and disgusted. The Impulse Intensity scale (11 items, $\alpha = .76$, AIC .22) contained items relating to experiencing strong emotions that are difficult to suppress. The Positive Expressivity scale (13 items, $\alpha = .85$, AIC .30) contained items tapping positive terms such as: joy, energetic, enthusiastic, happy, and laugh.

Dobbs et al. (2007), in a more recent structural analysis, tested the three-factor hierarchical model of the BEQ with both PCA and CFA; relevant fit indices were provided for the latter. First, a Varimax-rotated PCA three-factor solution accounted for 49% of the total variance. An examination of the factor loadings revealed that 5 out the

16 items had cross-loadings greater than .30 (see Table 2, Dobbs et al., 2007). Two items from the BEQ Negative Expressivity scale (i.e., "People often do not know what I am feeling" and "I have learned it is better to suppress my anger") cross-loaded on the positive expressivity factor; in fact, the second item loaded more strongly on positive expressivity (loading .38) than on the designated factor of negative expressivity (loading only .18). Two items from the BEQ Impulse Strength scale cross-loaded on other factors; one item (i.e., "I am sometimes unable to hide my feelings, even though I would like to") loaded approximately equally on the negative expressivity factor and the other item (i.e., "I have strong emotions") cross-loaded on the positive expressivity factor. Finally, one item from the BEQ Positive Expressivity scale (i.e., "I am an emotionally expressive person") cross-loaded more strongly on both the negative expressivity factor and the impulse strength factor (loading .53 and .39, respectively) than on the designated factor (loading .38).

More generally, cross-loadings are expected within a hierarchical structure, that is, items containing common variance of the higher order factor result in cross-loadings when examining the lower order factors. Furthermore, an oblique rotation allows the lower order factors to be correlated, thus minimizing the magnitude of the cross-loadings. However, in the current structural analyses the cross-loadings are greater than the expected primary loadings based on the scoring of the facets. Moreover, these cross-loadings may partially account for the moderately strong intercorrelations among the BEQ facets: positive and negative expressivity, r = .47; positive expressivity and impulse strength, r = .47; negative expressivity and impulse strength, r = .41 (Dobbs et al., 2007).

CFA of the BEQ items failed to fit the data (Dobbs et al., 2007). For these analyses, the model included three correlated error variances to account for the three reverse-keyed items. CFA of the BEQ failed fit a three-factor hierarchical model (χ^2 (98) = 435.7; CFI = .78, RMSEA = .10), a three-factor orthogonal model (χ^2 (101) = 620.9; CFI = .66, RMSEA = .12), or a one-factor model (χ^2 (101) = 564.1; CFI = .70, RMSEA =

.11). Modifying the model (i.e., removal of items, allowing cross-loadings, adding additional correlated error variances) also failed to achieve adequate fit or could not be theoretically supported. Although it is difficult to achieve adequate fit indices with item-level CFA, these analyses revealed that achieving adequate fit for the BEQ required more than simply removing some problematic items.

Dobbs et al. (2007) also examined the fit of the EES (Kring et al., 1994) based on a general factor model of emotional expressivity. A PCA one-factor solution accounted for 47% of the total variance with the 17 items loading from .81-.42. Initial CFA of the EES revealed a poor fit for a one-factor model (χ^2 (119) = 639; CFI = .84, RMSEA = .11). After modifying the model to include 21 correlated error terms, a one-factor model adequately fit the data (χ^2 (98) = 230.5; CFI = .96, RMSEA = .06). Adding correlated error terms to achieve adequate fit indicated there are covariances among the items unaccounted for by the general factor. However, the items of EES do not contain content related to specific emotions, valence, or nonverbal channel and therefore a clear second factor that could account for these covariances is unlikely. Overall, these results supported a general factor of emotional expressivity, but again it is difficult to achieve an acceptable fit based on item-level CFA.

Another recent structural analysis of emotional expressivity utilized multi-method structural equation modeling based on ratings of four emotion terms for each of the following: love, joy, fear, anger, shame, and sadness (Trierweiler et al., 2002). Participants and two peers rated “how often they showed their emotions to other people when they experienced them on 4-point frequency scales (responses ranged from never or almost never to almost always or always)” (Trierweiler et al., 2002, p. 1027). Within the four emotion terms, two-item parcels created two observed measures of each of the six specific emotions. Structural equation modeling tested three nested models: a general expressivity model, a two-factor valence model and a discrete emotions model. The general expressivity model fit the worst of the three models, (χ^2 (54) = 1332.8; RMSEA

= .22; BIC = 9905.01) and the two-factor valence model also failed to fit the data adequately, (χ^2 (53) = 718.9; RMSEA = .16; BIC = 9297.21). The multidimensional discrete emotions model fit the data the best, (χ^2 (39) = 59.14; RMSEA = .02; BIC = 8723.82). Importantly, these CFAs compared a discrete emotions model to a general factor model and a two-factor valence model rather than examining these three levels within a hierarchical structure.

To examine the potential of a hierarchical structure, I conducted second-order factor analyses of the correlation matrix of the discrete emotions (see Table 5 for self-ratings, Trierweiler et al., 2002). Second-order factor analyses tested for higher order factors above the discrete emotions model, that is, a one-factor second-order factor analysis tested for a general factor and a two-factor second-order factor analysis tested for positive and negative expressivity. A second-order one-factor PCA solution accounted for 50% of the total variance with discrete emotions factors loading from .84 to .55 on the general factor. A second-order Promax-rotated two-factor PCA solution accounted for 70% of the total variance and the two factors correlated .39. The negative discrete emotions loaded on the first factor from .86 to .67 and the positive discrete emotions loaded on the second factor from .96 to .91, with all cross-loadings less than |.10|.

In addition Trierweiler et al., (2002) supported the multidimensionality of emotional expressivity through the following: (a) mean level differences in expression of discrete emotions, (b) differing levels of self-peer convergence for the discrete emotions, and (c) differential relations between discrete emotions and personality. Participants reported expressing positive emotions (i.e., joy and love) more frequently than negative emotions (i.e., sadness, fear, anger, and shame). More specifically, participants reported the highest and lowest mean levels of expressing joy and shame, respectively. In line with the trait visibility effect, results confirmed the highest and lowest levels of self-peer convergence for joy and shame, respectively. The relations between emotional expressivity and personality will be discussed in more detail in the next section.

Emotional Expressivity and the Big Five

Determining the relations among emotional expressivity, positive and negative expressivity and existing constructs—or, in other words, situating these constructs within a nomological network—provides a greater understanding of the domain (Cronbach & Meehl, 1955). Broadly speaking, E is the Big Five trait most strongly related to emotional expressivity; that is, people who tend to be social and gregarious are also expressive in general (Kring et al., 1994). Riggio and Riggio (2002), in a recent meta-analysis, reported an effect size of $r = .39$, $p < .01$ for the relation between E and emotional expressiveness—two-thirds of the studies reported a significant correlation. The meta-analysis also reported an effect size of $r = -.01$ for the relation between N and emotional expressivity. A broader conceptualization of expressiveness, rather than *emotional* expressivity or specifying separate dimensions of positive and negative expressivity, potentially resulted in the lack of a significant relation with N. In contrast, the EES general factor of emotional expressivity correlated with N ($r = .21$; Kring et al., 1994) and the BEQ Total scale also correlated with N ($r = .29$; Gross & John, 1995).

In the structural analysis of six expressivity questionnaires discussed earlier, Gross and John (1998) selected items from these questionnaires to form the three facets of Core Emotional Expressivity: Positive Expressivity, Negative Expressivity and Impulse Intensity. These facets represented the same basic structure as the original BEQ but the scales contained more items and, consequently, have slightly higher coefficient alphas. When these revised facets were correlated with the Big Five, Positive Expressivity correlated more strongly with E ($r = .58$) than with N ($r = .04$) and Negative Expressivity correlated more strongly with N ($r = .47$) than E ($r = .27$). Furthermore, N correlated more strongly with Negative Expressivity ($r = .47$) than with Impulse Intensity ($r = .34$).

Reliably and validly assessing positive and negative expressivity could explicate important issues of discriminant validity masked by focusing solely on a general expressivity factor. Positive and negative expressivity based on the compilation of items from the various measures of expressivity (Gross & John, 1998) demonstrated relatively clearer differential relations with E and N, and BEQ Positive Expressivity correlated more strongly with E (r = .41) than N (r = .04); however, BEQ Negative Expressivity correlated comparably with N (r = .26) and E (r = .21). Furthermore, compared to BEQ Negative Expressivity, Impulse Strength correlated more strongly with N (r = .38 vs. .26; Gross & John, 1995). The lack of a clear differential pattern in the relations of Negative Expressivity with E and N, as well as the stronger convergence between Impulse Strength and N than between Negative Expressivity and N, suggests that the BEQ Negative Expressivity scale may have poor convergent and discriminant validity.

Additionally, the discrete emotions level demonstrated differential relations with the Big Five (Trierweiler et al., 2002). Across both self- and peer-ratings, the discrete emotions of joy and love correlated with E ($r \approx .50$), A ($r \approx .35$) and Intellect ($r \approx .20$), whereas the discrete emotions of sadness, fear, anger, shame correlated with N ($r \approx .25$). The peer-ratings generally corroborated the self-ratings but, in addition, the peer-ratings indicated correlations of anger and shame with A ($r \approx$ -.35 and $\approx$ .20, respectively).

In summary, two recent structural analyses (Dobbs et al., 2007; Trierweiler et al., 2002) supported either a general factor model or a discrete emotions model. The current state of the literature leaves the domain of emotional expressivity without a clear structure and without an understanding of how a general model and a discrete emotions model are related. However, second-order factor analyses of the correlations among the discrete emotions suggested that higher order factors of positive and negative expressivity existed above the discrete emotions level. Moreover, reliably and validly assessing positive and negative expressivity could explicate important issues of discriminant validity and demonstrate clearer differential relations with E and N.

CHAPTER 3. A PRELIMINARY STUDY: THE COUPLES ASSESSMENT PROJECT

Emotional expressivity is an important aspect of interpersonal relationships, and expressing one's emotions appropriately may be of great importance within a marriage. As part of a larger research study, the Couples Assessment Project (CAP), newlywed couples completed a questionnaire packet and a videotaped interaction task. For the purposes of this scale development project, the results focused on the BEQ (Gross & John, 1995), personality and trait affect. This study consisted of a large sample (i.e., 394 participants; 197 couples) and allowed for a strong test of the psychometric properties of the BEQ, such as internal consistency and structural validity. CAP began before Dobbs et al. (2007) published a structural analysis of the BEQ; however, the replication of structural problems of the BEQ in a large community sample would further support the need for a new measure of emotional expressivity.

Previous reviews of existing measures and structural analyses supported a general factor model of emotional expressivity and a discrete emotions model; however, these reviews and structural analyses provided very limited support for a hierarchical model including positive and negative expressivity. However, the differential relations of positive and negative expressivity with E and N indicated the importance of this level of abstraction for understanding the nomological network of emotional expressivity. Therefore, CAP included the BEQ (Gross & John, 1995) to clarify this level of abstraction and to examine the relations of positive and negative expressivity with personality and trait affect.

In addition to testing the original structure of the BEQ, exploratory analyses examined alternative factor solutions. A one-factor principal factor analysis (PFA) and a Promax-rotated three-factor PFA solution tested the general factor of the BEQ and the three correlated facets of Impulse Strength, Positive Expressivity, and Negative Expressivity, respectively. To support the general factor, all items need to load at least

.30 in the one-factor solution. To support the lower order structure, all items need to load at least .30 on the designated factor, with cross-loadings less than .30, in the three-factor solution. Additionally, a two-factor solution would reveal how the second factor splits from the general factor and what remains part of the larger general factor. Finally, the removal of poor items (i.e., items that fail to load > .30 on the designated factor or items that cross-load > .30 on the other factors) and taking into account the level of abstraction of the item content could yield support for an alternative factor solutions and ultimately a simplified hierarchical model.

In CAP, self- and spouse-ratings of the BEQ comprised two methods within MTMM analyses (Campbell & Fiske, 1959), and allowed for simultaneous tests of convergent and discriminant validity. Agreement between the self- and spouse-ratings (i.e., a variable-centered approach) addressed convergent validity; comparing the level of the self-spouse agreement correlations to other correlations within the MTMM addressed discriminant validity. The correlation between the same variable assessed by two different methods (e.g., self-spouse agreement of positive expressivity) should be stronger than the correlation between two different variables assessed by the same method (e.g., self-rated positive expressivity and self-rated negative expressivity) or the correlation between two different variables assessed by two different methods (e.g., self-rated positive expressivity and spouse-rated negative expressivity).

CAP also included measures of personality and trait affect to examine the nomological network of emotional expressivity. To replicate previous studies, emotional expressivity needed to correlate significantly with E. Beyond a general factor of emotional expressivity, the inclusion of the BEQ and, more specifically, the facets of positive and negative expressivity allowed for an examination of differential relations with personality and trait affect. Based on previous studies (Gross & John, 1997, 1998), I hypothesized that Positive Expressivity would converge with E and PA, whereas Negative Expressivity would converge with N and NA. Finally, the discriminant validity of Negative Expressivity and Impulse Strength in relation to N and NA needed to be addressed.

Prior research (Kring et al., 1998) documented substantial sex differences for emotional expressivity, with females reporting greater levels of emotionally expressivity. The newlywed sample in CAP ensured equivalent numbers of males and females in the sample to facilitate examining sex differences. Based on previous studies, I predicted females would report significantly higher levels of emotional expressivity; however, even though significant sex differences exist on the BEQ, these mean level differences were not expected to influence its factor structure (Gross & John, 1995).

Method

Participants

CAP consisted of a sample of 202 newlywed couples. Based on addresses listed on marriage licenses, we mailed invitation letters to recently married couples from Johnson and Linn County in eastern Iowa. Each couple received $90 in exchange for their participation. The mean age of the sample (N = 372) was 32.4 (SD = 10.8) years (husbands, M = 33.5, SD = 11.6; wives, M = 31.3, SD = 10.0). On average the newlyweds had known their spouses for 6.1 years and dated for 4.0 years prior to marriage. For the participants reporting ethnicity, the sample consisted of 87% Caucasian, 2% Asian American, 3% African-American, 1% Native American, 4% Foreign born, 3% Other, and 2% multi-ethnicities (9% of the sample did not report ethnicity).

Procedure and Measures

One to four couples participated in small group sessions that lasted approximately 2 hours. Couples completed a large questionnaire packet that contained the BEQ (Gross & John, 1995). The response format was changed from a 1 to 7 *disagree* to *agree* format to a 1 to 5 *not like me* to *like me* format because the BEQ items were intermixed with another scale. For the spouse-ratings, the BEQ also was adapted to create separate wife and husband versions (e.g., “I often express…” was changed to “S/he often expresses…”)

and the response format was also adapted (e.g., 1 to 5 *not like me* to *like me* was changed to 1 to 5 *not like her/him* to *like her/him*).

Participants also rated themselves on the Big Five Inventory (BFI; Benet-Martinez & John, 1998; John & Srivastava, 1999). The BFI is a 44-item measure of E (8-items), N (8-items), C (9-items), A (9-items) and O (10-items). The scale consists of a 5-point response scale from *strongly disagree* to *strongly agree*.

Participants completed the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). The PANAS consists of two 10-item higher order scales assessing PA (e.g., active, alert, excited) and NA (e.g., nervous, sad, afraid). Participants rated different mood terms on how they generally feel with a 5-point response scale from *very slightly or not at all* to *extremely*.

Results

Initial Analyses

The initial sample consisted of 202 couples, but after removing five couples due to missing ratings from one or both members the resulting sample consisted of 197 couples, 394 participants. On the BEQ items, one participant was missing two items that were imputed based on the average of the five items that correlated most strongly with the missing item in the rest of the sample. Descriptive statistics and internal consistency reliabilities for the self- and spouse-ratings on the BEQ are reported in Table 1. Coefficient alphas for the positive and negative expressivity facets in this sample were below .80. The reported means for the BEQ are based on the average of the items with a 5-point response scale. Cohen's *d*, reported in the column on the right, indicated significant sex differences on all scales with females being more expressive in both the self- and spouse-ratings.

Structural Analyses of the BEQ

The next set of analyses examined the internal structure of the BEQ self-ratings. Coefficient alpha was .89 and the AIC was .34 for the 16 items of the BEQ. A one-factor

PFA solution accounted for 82% of the common variance and supported the existence of a general factor of emotional expressivity (see Table 2). All but one item loaded above .30 on this general factor. The items relating to being expressive and having strong emotions defined this factor. The top-loading item tapped general emotional expressivity without explicitly mentioning valence content ("I am an emotionally expressive person"). Items referencing specific emotions (i.e., fear nervous, and anger) tended to have lower loadings than the more general items, or in other words, items including specific emotions contained variance that is unaccounted for by this general factor. Five out the top seven items defining the factor are from the Impulse Strength scale of the BEQ. The reverse-keyed items from the Negative Expressivity scale loaded the weakest on this general factor and the item explicitly referring to suppressing emotions loaded the weakest of all items. Overall, this one-factor solution supported the general factor of emotional expressivity.

In Table 2, one-factor PFA solutions, reported separately for females and males, supported the general factor; however, a few items loaded less than .30. Comparability coefficients supported the replicability of the general factor across the sexes (i.e., correlations were .996 and .994 in the female and male data, respectively). More specifically, regression-based factor scoring weights, calculated separately for the females and males, were cross applied to each sample yielding factor scores that were strongly correlated.

Before testing the structure of the BEQ based on its three-factor model, both Promax- and Varimax-rotated two-factor PFAs examined how a second factor split from the general factor. Although the Promax-rotation revealed a cleaner solution, a Varimax-rotation revealed only one additional splitter item (i.e., items with loadings > .30 on more than one factor). Overall the Promax- and Varimax-rotation yielded similar solutions; therefore, Table 3 only reports the former. A Promax-rotated two-factor PFA solution accounted for 97% of the common variance with rotated eigenvalues of 4.4 and 2.5.

Experiencing strong emotions and expressing negative emotions defined the first factor. The top two loading items are “I experience my emotions very strongly” and “I have strong emotions.” The items from the Positive Expressivity scale defined the second factor and all four positive expressivity items loaded greater than .30 on this factor. The top two loading items on the second factor were “Whenever I feel positive emotions, people can easily see exactly what I am feeling” and “When I’m happy, my feelings show.” The second factor captured more than just positive expressivity and included two items that tapped expressivity more generally.

This two-factor solution failed to reveal simple structure. Compared to the one-factor solution, the first factor in the two-factor solution captured more variance related to intensity or strength of emotions. The two splitter items contained more abstract item content without referencing specific emotions (i.e., “I am an emotionally expressive person” and “What I’m feeling is written all over my face”). Alternatively, the items with more specific content tapping valence loaded more cleanly on one of the two factors. Consistent with existence of a hierarchical general factor, the first factor correlated .58 with the second factor.

Organized within the existing facet scales of the BEQ, a Promax-rotated three-factor PFA solution failed to reveal clean simple structure (see Table 4). Clear marker items generally recreated the facets of the BEQ—Impulse Strength, Positive Expressivity and Negative Expressivity; however, three items cross-loaded greater than .30, two items failed to load strongest on the designated factor, and one item failed to load greater than .30 on any factor. In particular three of the six items of the Negative Expressivity scale either cross-loaded or failed to load on the factor. Again, the more abstract items tended to split across the factors.

A Promax-rotated three-factor PCA solution revealed a similar factor structure to the PFA solution. The first factor contained five of the same top loading items and the second factor contained three of the same items from the Positive Expressivity scale.

However, the three reverse-keyed items now defined the third factor, which likely reflected method variance related to reverse-keying. Ultimately the Promax-rotated three-factor PFA solution revealed fewer splitter items compared to the Promax-rotated three-factor PCA solution (three and five items, respectively).

To summarize the results of these factor analyses: (a) one-factor PFAs supported the general factor of emotional expressivity, (b) comparability coefficients established a replicable general factor structure in both females and males, (c) in the two-factor solution positive expressivity items and two general emotional expressivity items split from the general factor, (d) a Promax-rotated three-factor PFA failed to reveal a clean simple structure, but loosely recreated the BEQs three facets of Impulse Strength, Positive Expressivity, and Negative Expressivity, (e) in the three-factor solution several items split across factors or failed to load greater than .30 on the designated factor, and (f) the more abstract items loaded strongly in the one-factor solution but tended to split across factors in the two- and three-factor solutions.

Convergent and Discriminant Validity of the BEQ

At the BEQ facet level, the lack of clean simple structure and the existence of several cross-loading items necessitated an examination of discriminant validity. In Table 5 two MTMM matrices simultaneously addressed the convergent and discriminant validity of the BEQ. Self- and spouse-ratings of the BEQ formed the two methods of the MTMM. Partialing sex from these correlations controlled for inflated correlations due to significant sex differences. The upper MTMM contained the original facets of the BEQ based on 16 items; whereas, in the lower MTMM the removal of five problematic items left a 5-item Impulse Strength scale, a 3-item Positive Expressivity scale, and a 3-item Negative Expressivity scale. The removal of problematic items tested the influence of these items on convergent and discriminant correlations at the level of the facets. As previously noted, the items with more abstract item content tended to split across factors when examining the structure of the facets. Although these more abstract items loaded strongly on the general factor, including these items in analyses focused at the level of the facets reduced the clarity of the results.

First, along the reliability diagonals, parentheses bracket the coefficient alphas of these scales. In the lower MTMM matrix, the removal of the problematic items revealed only a slight drop in coefficient alphas. Similarly, along the validity diagonal (in bold) the self-spouse agreement correlations revealed only a slight drop from the upper to the lower MTMM. These significant self-spouse agreement correlations supported convergent validity. Next, the correlations in the validity diagonal exceeded any other correlation in its row or column of heterotrait-heteromethod block. This provided support for the discriminant validity of these scales. However, in the top MTMM the correlations on the validity diagonal failed the more stringent test of discriminant validity; that is, that the correlations of the validity diagonal needed to exceed the other correlations in their row or column of the monomethod triangles. Alternatively, in the lower MTMM, with the problematic items removed, some of the correlations of the validity diagonal exceeded the other correlations in their row or column of the monomethod triangles. Lastly, both MTMMs met the criteria of displaying a similar pattern of correlations in the heterotrait triangles for both the heteromethod and monomethod blocks.

Importantly, comparisons between the upper and lower MTMM matrices revealed that the correlation between Positive and Negative Expressivity dropped from .50 to .25 within the self-ratings monomethod triangle and from .46 to .28 within the spouse-ratings monomethod triangle. By removing the cross-loading items, these 3-item Positive and Negative Expressivity scales now passed the more stringent test of discriminant validity. Only the correlations with Impulse Strength still failed the more stringent test of discriminant validity.

Exploratory Scale Development

The BEQ facets suffered from the inclusion of more abstract items that created problems with the internal structure and discriminant validity. For instance, the items "I am an emotionally expressive person" and "What I'm feeling is written all over my face" tapped the general factor; however, both items split across factors in the two- and three-factor solutions. The next set of exploratory analyses took into account the level of abstraction of the BEQ items.

A simplified hierarchical model conceptualized emotional expressivity as a general factor with two subdimensions of positive and negative expressivity. Moreover, taking into account the level of abstraction of the BEQ items, three 3-item scales assessed: general emotional expressivity, positive expressivity and negative expressivity. Three more abstract items, without valence content, comprised a general emotional expressivity scale (i.e., "What I'm feeling is written all over my face," "I am an emotionally expressive person," and "People often do not know what I am feeling"). Although problematic in analyses at the facet level, these abstract items tended to be strong markers of the general factor and therefore created an observed measure of the general factor of emotional expressivity. The 3-item Positive and Negative Expressivity scales corresponded to the revised scales from the lower MTMM in Table 5 and contained the top three loading items from the three-factor solution. These scales clearly lacked comprehensiveness but approximated equivalent representations of positive (i.e., positive, laugh, and happy) and negative (i.e., fear, negative, and nervous or upset") expressivity.

To test that all nine items (3 general, 3 positive and 3 negative) supported a general factor, a PFA one-factor solution accounted for 93% of the common variance and the factor loadings ranged from |.76| - |.37| (see Table 6). A coefficient alpha of .82 also supported the internal consistency of the nine items. Next, in lower half of Table 6 a Promax-rotated two-factor PFA of just the six positive and negative expressivity items accounted for 97% of the common variance and supported two clear factors. The three positive expressivity items defined the first factor and the three negative expressivity items defined the second factor. All of the positive and negative items loaded greater than

|.45| on their designated factor with cross-loadings less than |.15|. With sex partialed out, the 3-item positive and negative expressivity scales correlated significantly with each other (r = .25) and with the 3-item general expressivity scale (rs = .57 and .47, respectively). Overall, these scales contained too few items to assess these constructs reliably; however, these exploratory analyses supported a simplified hierarchical model and further clarified issues related to the level of abstraction of the BEQ items.

The results in Table 7 continued to compare the original BEQ with the revised 3-item scales in relation to the BFI and PANAS scales. In the upper half of Table 7, the BEQ Total scale correlated about equally with E and N; however, Positive Expressivity correlated more strongly with E, whereas Negative Expressivity and Impulse Strength correlated more strongly with N. It is noteworthy that NA correlated more strongly with Impulse Strength than with Negative Expressivity (.32 vs. .18).

In the lower half of Table 7, the 3-item general expressivity scale demonstrated a slightly different pattern of relations than the BEQ Total scale and the 3-item positive and negative expressivity scales revealed clearer differential relations with the BFI and PANAS. Compared to the BEQ Total scale that correlated more strongly with N (r = .42) than E (r = .30), the 3-item general expressivity scale correlated more strongly with E (r = .42) than N (r = .20). After removing the three problematic items from the original BEQ Negative Expressivity scale, the correlation between the reduced 3-item negative expressivity scale and N increased substantially (.39 to .53), whereas the correlation with E decreased substantially (.22 to .02). The differential pattern of positive and negative expressivity replicated in the relations with the PANAS. Removing three items from the original BEQ Negative Expressivity scale doubled the correlation between negative expressivity and NA (.18 to .36).

In addition, exploratory regression analyses with these 3-item scales demonstrated the incremental predictive validity of positive and negative expressivity in predicting E and N. In both the upper and lower halves of Table 8, in the first step of these regression analyses positive and negative expressivity predicted E and N, respectively. In the next step, PA accounted for incremental variance in E and NA accounted for incremental variance in N. Next, reversal of the order of entry for the predictors demonstrated the differential predictive validity of emotional expressivity and trait affect. Comparing the R^2 change from Step 2 for the prediction of E revealed PA contributed more unique variance than positive expressivity (.15 and .09, respectively). Similarly, comparing the R^2 change from Step 2 for the prediction of N revealed NA contributed more unique variance than negative expressivity (.25 and .11, respectively). At the same time, however, these analyses also demonstrated the non-redundancy of emotional expressivity and trait affect, that is, positive and negative expressivity explained meaningful variance beyond PA and NA. Finally, in Step 3, negative expressivity is entered to predict E and positive expressivity is entered to predict N; neither contributed any additional unique variance. The third step of these regression analyses clearly demonstrated the differential predictive validity of positive and negative expressivity.

Discussion

The three factor hierarchical structure of the BEQ received only limited support in these analyses. Coefficient alphas below .80 indicated the BEQ facet scales are unreliable. Previous research (Gross & John, 1995, 1997) also reported coefficient alphas below .80 and another study reported alphas of .63 for both positive and negative expressivity (Dobbs et al., 2007). The low coefficient alphas likely resulted from too few items assessing these constructs; however, several problematic items also reduced internal consistency. In particular, three of the six items of the Negative Expressivity scale either cross-loaded or failed to load on the factor. Unfortunately, the two articles related to the BEQ scale development omitted the cross-loadings (Gross & John, 1995, 1997). However, importantly, four of the five problematic items in the CAP data replicated problematic items from another recent structural analysis of the BEQ (Dobbs et al., 2007). Additionally, a few of the problematic items lacked content validity, such as "I am an emotionally expressive person" as a positive expressivity item and "People often don't know what I am feeling" as a negative expressivity item.

This lack of clear internal structure raised issues of discriminant validity among the BEQ facets. The MTMM matrix for the original BEQ revealed that correlations between two different traits assessed by the same method exceeded the correlations between the same trait assessed by two different methods. This lack of discriminant validity is not surprising considering the results of the factor analyses. With several cross-loading items one would expect significant correlations between these factors. The second MTMM matrix of the BEQ scales confirmed that after the removal of problematic items, the correlation between positive and negative expressivity dropped from .50 to .25. Clearly, these cross-loading items created problems related to discriminant validity, and further brought into question the internal structure of the BEQ.

More importantly, the results from this preliminary study pointed out the importance of considering the level of abstraction and content of the items. In particular, some of the BEQ items clearly demonstrated strong loadings on the general factor but these same items split across factors at the facet level. These analyses identified items functioning at different levels of abstraction.

Then, exploratory analyses systematically took into account the level of abstraction of the items and led to the creation of three 3-item scales: a general emotional expressivity scale, a positive expressivity scale and a negative expressivity scale. The constellation of correlations among the 3-item scales supported a simplified hierarchical structure. If considered factor loadings on a latent general emotional expressivity factor, multiplying the correlations between the 3-item general expressivity scale and the 3-item positive and negative expressivity scales (*r*s = .57 and .47, respectively) approximated the correlation between positive and negative expressivity ($r = .25$).

Clearer differential relations with the BFI and PANAS also supported these exploratory 3-item scales. First, both the original BEQ Total scale and the revised 3-item general emotional expressivity scale correlated significantly with E and N. However, a recent meta-analysis reported a significant relation between emotional expressivity and E but also reported essentially no relation between emotional expressivity and N (Riggio & Riggio, 2002). This meta-analysis primarily contained measures assessing expressivity as a general factor rather than specifying separate dimensions of positive and negative expressivity. Second, although previous research (Gross & John, 1997) already reported differential relations—with Positive Expressivity related more strongly to E and Negative Expressivity related more strongly to N—in the current study, the removal of problematic items revealed even more striking differential relations of positive and negative expressivity. These clearer differential relations highlight the importance of reliably and validly assessing positive and negative expressivity.

These exploratory 3-item positive and negative expressivity scales also demonstrated clearer differential relations with PA and NA. Importantly, however, the correlations between the corresponding expressivity scales and E and N exceeded the correlations with the corresponding trait affect scales. Within the nomological network, both expressivity and affectivity aligned more closely with E and N than with each other. Although emotional expressivity and trait affect moderately correlated, the magnitude of the correlations suggested quite a bit of differentiation between emotional expressivity (i.e., the tendency to express affective states through nonverbal means) and trait affect (i.e., what one generally experiences or feels).

Hierarchical regressions further demonstrated the differential relations of positive and negative expressivity with E and N. Indeed, negative expressivity accounted for a greater amount of the variance in N than positive expressivity accounted for in E, which is noteworthy, considering a meta-analysis provided little support for a relation between emotional expressivity and N (Riggio & Riggio, 2002). Furthermore, these regression analyses demonstrated the non-redundancy of emotional expressivity and trait affect.

Whereas E and N both contained an experiential component—such that E related strongly to PA and N related strongly to NA—E and N also contained an expressive component.

In summary, these exploratory analyses and the creation of the revised 3-item scales demonstrated the potential to create a hierarchical assessment, that is, separate scales of general emotional expressivity, positive expressivity and negative expressivity. In particular these exploratory analyses highlighted the importance of taking into account the level of abstraction of the item content. By removing the more abstract items, the revised 3-item scales of positive and negative expressivity more clearly aligned with E/PA and N/NA, respectively. The next series of studies built from these exploratory analyses and simplified hierarchical model to create a hierarchical assessment of emotional expressivity.

Table 1. Descriptive Statistics for Self- and Spouse-Ratings of the BEQ in the CAP Sample

Scale (# of items)	α	AIC	*M* (*SD*)	Female *M* (*SD*)	Male *M* (*SD*)	*d*
Self-ratings						
BEQ Total Scale (16)	.89	.34	3.14 (.73)	3.50 (.66)	2.78 (.61)	**1.13**
Positive Expressivity (4)	.76	.44	3.71 (.84)	4.01 (.78)	3.42 (.80)	**0.75**
Negative Expressivity (6)	.68	.26	2.85 (.72)	3.13 (.71)	2.58 (.62)	**0.83**
Impulse Strength (6)	.86	.51	3.04 (1.0)	3.52 (.86)	2.55 (.88)	**1.11**
Spouse-ratings						
BEQ Total Scale (16)	.89	.34	3.23 (.72)	3.61 (.55)	2.85 (.66)	**1.25**
Positive Expressivity (4)	.76	.44	3.83 (.77)	4.08 (.60)	3.57 (.83)	**0.70**
Negative Expressivity (6)	.71	.29	2.99 (.75)	3.32 (.65)	2.66 (.69)	**0.98**
Impulse Strength (6)	.86	.51	3.07 (.96)	3.58 (.78)	2.55 (.85)	**1.26**

N = 394.

BEQ = Berkeley Expressivity Questionnaire.

Table 2. 1-Factor PFAs of the BEQ in the CAP Sample

BEQ Items	Factor 1	Female Factor 1	Male Factor 1
I am an emotionally expressive person.	.79	.77	.72
I experience my emotions very strongly.	.76	.75	.70
I have strong emotions.	.74	.74	.69
What I'm feeling is written all over my face.	.73	.74	.64
I am sometimes unable to hide my feelings, even though I would like to.	.69	.66	.64
I sometimes cry at sad movies.	.67	.59	.56
My body reacts very strongly to emotional situations.	.66	.60	.65
When I'm happy, my feelings show.	.63	.65	.54
Whenever I feel positive emotions, people can easily see exactly what I am feeling.	.60	.58	.52
Whenever I feel negative emotions, people can easily see exactly what I am feeling.	.60	.63	.55
There have been times when I have not been able to stop crying even though I tried to stop.	.55	.44	.46
It is difficult for me to hide my fear.	.55	.50	.50
No matter how nervous or upset I am, I tend to keep a calm exterior. **(R)**	.41	.35	.33
People often do not know what I am feeling. **(R)**	.38	.33	.24
I laugh out loud when someone tells me a joke that I think is funny.	.32	.35	.23
I've learned it is better to suppress my anger than to show it. **(R)**	.15	.15	.02

N = 394 (Female N = 197; Male N = 194).

BEQ = Berkeley Expressivity Questionnaire.

(R) = Reverse-keyed items.

Note: BEQ items reprinted from *Personality and Individual Differences*, v. 19, J.J. Gross and O.P. John, “Facets of emotional expressivity: Three self-report factors and their correlates,” pp. 555-568, Copyright 1995, with permission from Elsevier.

Table 3. Promax-Rotated 2-Factor PFA of the BEQ in the CAP Sample

BEQ Items	Factor 1	Factor 2
I experience my emotions very strongly.	**.83**	-.05
I have strong emotions.	**.78**	-.02
My body reacts very strongly to emotional situations.	**.75**	-.07
There have been times when I have not been able to stop crying even though I tried to stop.	**.72**	-.18
I sometimes cry at sad movies.	**.67**	.04
I am sometimes unable to hide my feelings, even though I would like to.	**.61**	.14
I am an emotionally expressive person.	**.54**	**.34**
Whenever I feel negative emotions, people can easily see exactly what I am feeling.	**.53**	.12
It is difficult for me to hide my fear.	**.50**	.09
No matter how nervous or upset I am, I tend to keep a calm exterior. **(R)**	**.42**	.01
Whenever I feel positive emotions, people can easily see exactly what I am feeling.	-.02	**.79**
When I'm happy, my feelings show.	.03	**.76**
People often do not know what I am feeling. **(R)**	-.07	**.57**
What I'm feeling is written all over my face.	**.41**	**.43**
I laugh out loud when someone tells me a joke that I think is funny.	-.01	**.43**
I've learned it is better to suppress my anger than to show it. **(R)**	-.03	.22

N = 394.

BEQ = Berkeley Expressivity Questionnaire.

(R) = Items were reverse-keyed prior to factor analysis.

Note: Factor loadings ≥ .30 are in bold.

Note: BEQ items reprinted from *Personality and Individual Differences*, v. 19, J.J. Gross and O.P. John, "Facets of emotional expressivity: Three self-report factors and their correlates," pp. 555-568, Copyright 1995, with permission from Elsevier.

Table 4. Promax-Rotated 3-Factor PFA of the BEQ in the CAP Sample

BEQ Items	Impulse Strength	Positive Expressivity	Negative Expressivity
I have strong emotions.	**.88**	.02	-.12
I experience my emotions very strongly.	**.88**	-.02	-.05
I sometimes cry at sad movies.	**.63**	.04	.07
My body reacts very strongly to emotional situations.	**.62**	-.08	.20
There have been times when I have not been able to stop crying even though I tried to stop.	**.61**	-.19	.18
I am sometimes unable to hide my feelings, even though I would like to.	**.37**	.10	**.39**
Whenever I feel positive emotions, people can easily see exactly what I am feeling.	.02	**.78**	-.01
When I'm happy, my feelings show.	.11	**.77**	-.08
I laugh out loud when someone tells me a joke that I think is funny.	.06	**.44**	-.09
I am an emotionally expressive person.	**.53**	**.35**	.04
No matter how nervous or upset I am, I tend to keep a calm exterior. **(R)**	.03	-.07	**.58**
Whenever I feel negative emotions, people can easily see exactly what I am feeling.	.18	.05	**.53**
It is difficult for me to hide my fear.	.22	.04	**.44**
What I'm feeling is written all over my face.	.16	**.38**	**.41**
People often do not know what I am feeling. **(R)**	-.21	**.53**	.23
I've learned it is better to suppress my anger than to show it. **(R)**	-.12	.20	.15

N = 394.

BEQ = Berkeley Expressivity Questionnaire.

(R) = Items were reverse-keyed prior to factor analysis.

Note: Factor loadings ≥ .30 are in bold.

Note: BEQ items reprinted from *Personality and Individual Differences*, v. 19, J.J. Gross and O.P. John, “Facets of emotional expressivity: Three self-report factors and their correlates,” pp. 555-568, Copyright 1995, with permission from Elsevier.

Table 5. MTMM for Self- and Spouse-Ratings of the BEQ and Revised BEQ in the CAP Sample

BEQ Scales (# items)	Self-Ratings 1	2	3	Spouse-Ratings 4	5	6
Self-ratings						
1. Positive Expressivity (4)	(.76)					
2. Negative Expressivity (6)	.50	(.68)				
3. Impulse Strength (6)	.50	.49	(.86)			
Spouse-ratings						
4. Positive Expressivity (4)	**.42**	.25	.25	(.76)		
5. Negative Expressivity (6)	.29	**.39**	.30	.46	(.71)	
6. Impulse Strength (6)	.31	.28	**.42**	.53	.51	(.86)

Revised BEQ Scales (# items)	Self-Ratings 1	2	3	Spouse-Ratings 4	5	6
Self-ratings						
1. Positive Expressivity (3)	(.72)					
2. Negative Expressivity (3)	.25	(.65)				
3. Impulse Strength (5)	.32	.47	(.85)			
Spouse-ratings						
4. Positive Expressivity (3)	**.35**	.17	.16	(.70)		
5. Negative Expressivity (3)	.19	**.38**	.32	.28	(.64)	
6. Impulse Strength (5)	.23	.30	**.40**	.41	.48	(.84)

N = 394.

BEQ = Berkeley Expressivity Questionnaire.

Note: All correlations are partialed by sex. All correlations significant at $p < .001$. Convergent correlations are in bold.

Table 6. 1-Factor and Promax-Rotated 2-Factor PFAs of the 3-Item BEQ in the CAP Sample

3-Item BEQ	Factor 1
What I'm feeling is written all over my face.	**.76**
When I'm happy, my feelings show.	**.73**
Whenever I feel positive emotions, people can easily see exactly what I am feeling.	**.72**
I am an emotionally expressive person.	**.70**
Whenever I feel negative emotions, people can easily see exactly what I am feeling.	**.59**
It is difficult for me to hide my fear.	**.53**
I laugh out loud when someone tells me a joke that I think is funny.	**.38**
No matter how nervous or upset I am, I tend to keep a calm exterior. **(R)**	**-.37**
People often do not know what I am feeling. **(R)**	**-.47**

3-Item BEQ	Factor 1	Factor 2
When I'm happy, my feelings show.	**.80**	.01
Whenever I feel positive emotions, people can easily see exactly what I am feeling.	**.75**	.03
I laugh out loud when someone tells me a joke that I think is funny.	**.47**	-.06
Whenever I feel negative emotions, people can easily see exactly what I am feeling.	.06	**.63**
It is difficult for me to hide my fear.	.08	**.58**
No matter how nervous or upset I am, I tend to keep a calm exterior. **(R)**	-.12	**-.54**

N = 394.

BEQ = Berkeley Expressivity Questionnaire.

(R) = Reverse-keyed items.

Note: Factor loadings ≥ .30 are in bold.

Note: BEQ items reprinted from *Personality and Individual Differences,* v. 19, J.J. Gross and O.P. John, "Facets of emotional expressivity: Three self-report factors and their correlates," pp. 555-568, Copyright 1995, with permission from Elsevier.

Table 7. Correlations Between BEQ and 3-Item BEQ with PANAS and BFI in the CAP Sample

BEQ Self-Ratings	PA	NA	E	N	O	A	C
BEQ Total Scale	.11	**.25***	**.30***	**.42***	.09	.07	.08
Positive Expressivity	**.29***	.06	**.48***	.12	.17*	.18 *	.13
Negative Expressivity	-.03	.18*	**.22***	**.39***	-.06	.04	.03
Impulse Strength	.07	**.32***	.15*	**.47***	.12	.01	.06

3-Item BEQ Self-Ratings	PA	NA	E	N	O	A	C
3-Item General Expressivity	.15*	.06	**.42***	**.20***	.04	.18*	.14*
3-Item Positive Expressivity	**.30***	.02	**.44***	.06	.16*	.16 *	.13*
3-Item Negative Expressivity	-.14*	**.36***	.02	**.53***	-.07	-.07	-.10

N = 388-394.

*$p < .01$.

Note: Correlations ≥ .20 in bold.

BEQ = Berkeley Expressivity Questionnaire. PA = Positive Affect. NA = Negative Affect. E = Extraversion. N = Neuroticism. O = Openness. A = Agreeableness. C = Conscientiousness.

Table 8. Hierarchical Regressions for BFI, PANAS, and 3-Item BEQ in the CAP Sample

Criterion Extraversion	Predictor	R^2	R^2 Change	β	Beta
Step 1	Pos Exp	.19		1.13*	.44*
Step 2	Pos Exp	.34	.15	.82*	.32*
	PA			.48*	.40*
Step 1	PA	.25		.59*	.50*
Step 2	PA	.34	.09	.48*	.40*
	Pos Exp			.82*	.32*
Step 3	PA	.34	.00	.47*	.39*
	Pos Exp			.85*	.33*
	Neg Exp			-.07	-.03

Criterion Neuroticism	Predictor	R^2	R^2 Change	β	Beta
Step 1	Neg Exp	.28		1.34*	.53*
Step 2	Neg Exp	.54	.25	.88*	.35*
	NA			.60*	.54*
Step 1	NA	.43		.74*	.66*
Step 2	NA	.54	.11	.60*	.54*
	Neg Exp			.88*	.35*
Step 3	NA	.54	.00	.60*	.53*
	Neg Exp			.94*	.37*
	Pos Exp			-.17	-.07

N = 388.

* $p < .001$

Pos Exp = 3-Item Positive Expressivity. Neg Exp = 3-Item Negative Expressivity. PA = Positive Affect. NA = Negative Affect.

Note: For Step 1, *df* Model = 1; *df* Error = 386; *df* Corrected Total = 387. For Step 2, *df* Model = 2; *df* Error = 385; *df* Corrected Total = 387. For Step 3, *df* Model = 3; *df* Error = 384; *df* Corrected Total = 387.

CHAPTER 4. ISEE DEVELOPMENT

The review of the existing measures of emotional expressivity and the results of the preliminary study indicated the need for a new measure of emotional expressivity. Justification for creating a new measure needed to demonstrate improvements over existing measures while accounting for existing evidence of construct validity. Construct validity cannot be established by a single study but rather is an ongoing process based on the accumulation of all evidence supporting the existence of the underlying construct. First, Loevinger (1957) and later Clark and Watson (1995) emphasized scale development and validation around the concept of construct validity. Clark and Watson (1995) specifically outlined scale development guidelines for substantive, structural, and external phases of scale development. This chapter focuses on the substantive phase.

In conjunction with a literature review (see Chapter 2), the substantive phase entailed a clear conceptualization and written out descriptions of the target constructs. Next, the substantive phase involved creating an initial item pool and evaluating the relevance, comprehensiveness (i.e., making sure not to leave out anything important) and representativeness (i.e., maintaining the right balance or the proportionality of the number of items). Furthermore, a panel of experts (i.e., members of the Watson and Clark labs) reviewed the items to identify problematic items and to strengthen the initial item pool. Then pilot-testing examined item-response distributions, interitem correlations and the potential of the items to form distinct factors.

Affect-Focused Perspective

Affect—the hedonic tone of one's waking experience that varies in intensity, duration and activating triggers—is a broad construct that incorporates both emotions and moods (Watson & Vaidya, 2003). Emotions are intense, but brief, responses that are triggered by an activating event; whereas, moods are less intense, can last for hours, and can occur without being triggered by an activating event (Watson & Vaidya, 2003). Emotions are a coordinated response incorporating one's subjective experience (e.g., the feeling of fear), an expressive component (e.g., a facial expression of fear), a physiological reaction (e.g., increased heart rate), and a behavioral response (e.g., fleeing). Moods are primarily the subjective experience of emotions that may or may not be accompanied by nonverbal behaviors. However, emotions and moods are not mutually exclusive, and "moods and emotions dynamically interact in important ways. Emotions can lead to particular moods and moods can alter the probability that particular emotions will be triggered" (Davidson, 1994, p. 53).

Previous conceptualizations of emotional expressivity emphasized an emotion-focused perspective consisting of strong emotional reactions and "manifest[ing] emotional impulses behaviorally" (Gross & John, 1995, p. 555). This conceptualization of emotional expressivity developed from a two-stage model of emotions with an emotional impulse in the first stage and then emotionally expressive behaviors in the second stage. Working from this two-stage model Gross and John "aimed to develop a measure that would assess both the strength of emotional response tendencies and the degree to which these behavioral impulses are expressed" (1995, p. 556).

The current conceptualization of emotional expressivity emphasizes an affect-focused perspective. Broadening the conceptualization of emotional expressivity with an affect-focused perspective shifted the emphasis more towards expressive behaviors. Although an affect-focused perspective still incorporates expressions resulting from the general emotion generation model (i.e., brief intense reactions to activating events; Gross & John, 1995; Levenson, 1994; Plutchik, 1990), an affect-focused perspective also incorporates expressive behaviors indicative of affective states that are less dependent on an initial generative emotional impulse. An affect-focused versus emotion-focused perspective parallels the distinction between affect and emotions, that is, affect includes both emotions and less intense states that can exist for longer durations without a clear activating trigger.

Previous researchers have suggested thresholds exist for experiencing and expressing affect (Ekman, 1992; Kring & Gordon, 1998). For example, two people could have similar internal affective experiences but have different thresholds for expressing these affective experiences. In this example, the person with a lower threshold for expressivity would generally express lower intensity affective states. Therefore, high levels of the continuous dimension of emotional expressivity are less dependent on intense emotional impulses. In the affect-focused perspective, the variability from high to low expressivity is more dependent on the threshold for expressing lower intensity affective states.

This broader affect-focused perspective of emotional expressivity accounts for subtler, less prototypical expressions that are indicative of lower intensity affective states. For example, a person may express an irritable mood through his or her tone of voice without meeting all of the criteria of an emotion of anger (i.e., subjective experience, an expressive component, a physiological reaction, and a behavioral response). A slight smile with raised eyebrows may reveal a person's good mood; however, this expression is less likely to result from an emotional impulse per se. Subtle expressions of sadness, such as sad eyes, may be another example of incorporating less intense expressions as a result of an affect-focused perspective. The explicit requirement for an activating event or trigger followed by a brief, intense expression certainly would seem to miss a substantial portion of everyday expressions of sadness. In essence, this would be an issue of restricted range in emotional expressivity, considering that as little as 17% of one's waking day is spent experiencing an emotion (Watson, 2000).

Communicative Function

Functional accounts of emotions recognized that emotions evolved as a complex system to provide an adaptive response while emphasizing that their subsystems—the subjective experience, an expressive component, a physiological reaction, and a behavioral response—may serve different functions (Keltner & Gross, 1999). Whereas the functions of some of these subsystems may be more directed internally towards the adaptive emotional response of the individual, the function of an expressive component of an emotional response may be directed towards the interpersonal level and thus primarily serve a communicative function.

The current conceptualization emphasizes genuine expressions of internal affective states rather than inhibited or feigned emotional expressions, that is, emotional expressions are accompanied by the corresponding affective state and are communicating that state. As an illustrative exercise, Kring et al. (1994) briefly addressed a conceptual framework that crossed high and low expressivity with high and low experience that resulted in four cells: (a) genuine expression, (b) not genuine expression, (c) genuine lack of expression, and (d) not genuine lack of expression. Not genuine expression and not genuine lack of expression are conceptually similar to the constructs of expressive confidence and masking described by Gross and John (1998). Although expressive confidence and masking were within the hierarchical model of expressivity both correlated only weakly with core emotional expressivity (*r*s = .15).

Three-Level Hierarchical Model of Emotional Expressivity

This section describes the three levels of the hierarchical model and the scales targeted to assess the constructs of each level. The three-level hierarchical model of emotional expressivity contains a general bipolar factor of emotional expressivity at the highest level, the two relatively independent dimensions of positive and negative expressivity at the next level, and multiple discrete affect-expressivity at the lowest level.

The general trait of emotional expressivity accounts for the tendency to express affective states through nonverbal behaviors, such as facial expressions, vocal characteristics (e.g., volume, tone, and rate of speech), eye contact, and body language (e.g., posture, and gestures) (Gross & John, 1995, 1997; Kring et al., 1994). Emotional expressivity refers to naturalistic expressions typically occurring without much deliberative thought, conscious attention or intention. Emotional expressivity is a continuous dimension with high levels indicated by people generally revealing their internal affective state through nonverbal behavior; that is, even in the absence of strong emotions, it is possible to tell how a person with a high level of emotional expressivity feels. A person with a low level of emotional expressivity may experience varied affective states, but only rarely do his or her facial expressions, body language and vocal characteristics convey these experiences outwardly.

In this hierarchical assessment, item content specifically phrased to target the general factor level form an observed measure of the general factor. Items intended to tap the broad general factor of emotional expressivity are more abstract without valence content or reference to specific nonverbal channels (e.g., “I am emotionally expressive”). Modeled after the EES, the general scale is designed to assess a bipolar factor of expressiveness versus inexpressiveness with a relative balance of reverse-keyed items. The general factor scale was expected to be approximately 5 to 8 items long (with an AIC of .45, 5 items yield > .80 alpha; with an AIC of .40, 8 items produce > .80 alpha).

The construct of general emotional expressivity encompasses two narrower second-order constructs—positive and negative expressivity. Positive expressivity encompasses expressions of PA including such affective states as excitement, confidence, and amusement. Negative expressivity encompasses expressions of NA including such affective states as sadness, fear, irritability, and guilt. The constructs of positive and negative expressivity are modeled after the experiential constructs of PA and NA (Watson & Tellegen, 1985; Watson et al., 1988; Watson & Clark, 1999) and are loosely modeled after Positive and Negative Expressivity of the BEQ (Gross & John, 1995). Positive and negative expressivity link the general factor to the discrete affects level and integrate the domain of emotional expressivity.

The relation between positive and negative expressivity needs to be clarified as a result of the poor discriminant validity exhibited in the preliminary study. The lack of internal consistency and multiple cross-loading items of the BEQ obscured the true relation between positive and negative expressivity. Although positive and negative emotional expressivity were moderately correlated, the differential relations with independent constructs of E and N and PA and NA supported of the relative independence of positive and negative expressivity. However, the relative independence of positive and negative expressivity questioned the existence of a meaningful general factor of emotional expressivity. Overall, the results from the preliminary study and more specifically the exploratory analyses suggested the expected range for the relation between positive and negative expressivity to be in the .30 - .50 range.

Modeled after the literature on the hierarchical assessment of affective experience (Watson et al., 1988; Watson & Clark, 1999) the positive and negative expressivity scales contain a representative subset of the discrete affect items. These items tap the discrete affects with more descriptive item content (e.g., valence and specific nonverbal behaviors). As unipolar scales, positive and negative expressivity only tap high levels of emotional expressivity. Although the general emotional expressivity scale incorporates reverse-keyed items, incorporating reverse-keyed positive and negative expressivity items—that contain discrete affective states—could have greater potential to be interpreted as masking, which was found to be only weakly negatively related to emotional expressivity (Gross & John, 1998). Positive and negative expressivity scales are expected to be approximately 10 items long (with an AIC of .30, 10 items yield > .80 alpha).

Published measures of emotional expressivity lack separate scales at the discrete affect level. Therefore, the affective terms included in the seven Basic Positive and Negative Emotion scales of the PANAS-X — joviality, self-assurance, attentiveness, fear, hostility, guilt, sadness—served as a starting point for the discrete affect scales in the current model (Watson & Clark, 1999). Starting from the empirically supported structure of the affective terms comprising the PANAS-X greatly facilitated the process of creating a pool of affective terms that are related without being redundant.

Additionally, consideration of other sets of affective terms (Diener et al., 1995; Izard, 1991; Tellegen et al., 1999b; Trierweiler et al., 2002) and item content from published measures of emotional expressivity were also reviewed (Friedman et al., 1980; Gross & John, 1995, 1998; King & Emmons, 1990; Kring et al., 1994) and yielded additional item content (i.e., affective terms and expressive nonverbal behaviors).

The discrete affect level of three-level hierarchical model initially targeted eight discrete affects. The four positive expressivity discrete affects included joviality, self-assurance and attentiveness—the three positive discrete affects of the PANAS-X—and additionally included an amusement scale. Previous scales of emotional expressivity included items relating to laughing (Gross & John, 1995; King & Emmons, 1990). For instance, five out of the seven items defining the first factor of the EEQ explicitly tapped laughing (King & Emmons, 1990). Moreover, in the scales created from the structural analysis of the five measures of expressivity (Gross & John, 1998), 6 out of the 13 items assessing the positive expressivity facet contained the word "laugh." The four negative expressivity discrete affects included sadness, fear, hostility, and guilt—the four negative discrete affects of the PANAS-X. Other models also contain very similar versions of these four basic negative affects (Diener et al., 1995; Trierweiler et al., 2002). The eight potential discrete affect scales—fear, hostility, guilt, sadness, joviality, self-assurance, attentiveness, and amusement—will be relatively narrow and possibly contain about six items each (with an AIC of .40, 6 items produce .80 alpha).

The proposed model focuses on these eight potential discrete affects rather than attempting to model every possible discrete affect exhaustively. First, the current literature lacks general consensus about the lower order of affect (Watson & Vaidya, 2003). Second, modeling the discrete affect level required multiple markers of each category and pragmatic reasons limited the scope of this scale development project. Third, this circumscribed number of lower order affects provided an initial test for the three-level hierarchical model of emotional expressivity; if it is successful, later attempts to extend and to determine the boundaries could incorporate additional discrete affects.

As mentioned in Chapter 1, the three-level hierarchical model of affectivity, containing a general bipolar factor of happiness, two relatively independent dimensions of PA and NA, and multiple specific affects (Tellegen et al., 1999a, 1999b; Watson & Tellegen, 1999), served as a template for the three-level hierarchical model of emotional expressivity. Both models posit a general bipolar factor that accounts for the relation between the second-order factors. In the experiential model, the general happiness versus unhappiness factor accounts for the weak to moderate negative correlation (approximately r = -.30) between the relatively independent constructs of PA and NA. The general expressivity versus inexpressivity factor accounts for the moderate positive correlation (approximately r = .30 - .50) between the relatively independent constructs of positive and negative expressivity. At the second-order level both PA and NA and positive and negative expressivity, account for the common variance between the similarly valenced discrete affects. At the lowest order level the discrete affects contain specific variance. Obvious connections exist between trait affect and emotional expressivity, that is, several of the discrete affects that a person feels are also the same discrete affects that a person expresses.

There are several reasons for aligning the three-level hierarchical models of affect and emotional expressivity. First, parsimony supported similar hierarchical models of affect and emotional expressivity. Second, models based on relatively similar levels of abstraction facilitate establishing conceptual connections between them. Third, matching affective and expressive subsystems fit with the communicative function of expressivity (i.e., linking internal affective experiences with expressions communicated to others).

Creation of Initial Item Pool

The initial item pool consisted of items at the general emotional expressivity and discrete affect levels. Nine homogeneous item composites (HIC; Hogan, 1983 as cited in

Watson et al., 2007) were created to assess: general emotional expressivity, hostility, fear, sadness, guilt, joviality, self-assurance, attentiveness and amusement. The initial item pool incorporated the concept of thresholds for expressing affective states, that is, item content incorporated various levels of difficulty to tap both higher and lower intensity affective states. Many of items explicitly contained conditional language (e.g., "If I'm feeling cheerful…" or "When I am nervous…"). Additionally, some items emphasized the communicative function of emotional expressivity (e.g., "People can see when I'm feeling…" or "People can tell when I'm feeling…").

The members of the Watson and Clark labs reviewed the items by selecting the three best and the three worst items from each HIC. Then group discussion addressed problematic items and potential ways to improve the initial item pool. This review process and further refinement resulted in a 73-item initial item pool (10 items for the general emotional expressivity HIC and 8 items for each of the discrete affect HICs, except for only 7 items for the attentiveness HIC). The initial 73-item pool incorporated an agreement response format (1= *disagree* to 5 = *agree*).

Pilot-Testing

The pilot-testing aimed to evaluate the initial item pool for problematic items (i.e., skewed response distributions), to examine the interitem correlations within the HICs to identify redundant items, to potentially identify HICs that have too few items and to examine the potential of the HICs to form distinct factors.

Participants and Procedure

The sample consisted of 213 participants recruited through two methods: (a) an email recruitment and online assessment, and (b) an extra credit opportunity for undergraduate students enrolled in a summer class. Participants recruited through email voluntarily completed a 73-item online survey at the time and place of their choosing. Out of 345 email invitations, 188 people (i.e., friends, family members, family members' friends, former students, professors, and research assistants) completed the online survey for a 54.5% response rate. The online participants consisted of 127 women and 61 men with their ages distributed as follows: 18-24 (9%); 25-34 (30%); 35-44 (16%); 45-54 (16%); 55-64 (17%); 65-older (13%). Additionally, 25 undergraduate students completed a questionnaire packet in a group testing session. The questionnaire packet started with the same 73 items as the online survey but also contained two additional scales of emotional expressivity and emotional contagion, and measures of personality, affectivity and social desirability—only results related to the 73 items are reported here. The student participants consisted of 16 women and 8 men with their ages distributed as follows: 18-24 (84%); 25-34 (12%). One participant omitted sex and age information.

Preliminary Analyses

Three dropped cases resulted from missing more than 10% of the data. In the rest of the sample, 168 participants completed all items and 42 participants omitted items—26 participants omitted only one item and 16 participants omitted between two to seven items. A total of 86 missing items—imputed from the average of the five most strongly correlated items in the rest of the sample—equaled 0.6% of the 15,330 total item responses.

Next, preliminary analyses examined the item response distributions, interitem correlations, and coefficient alphas of the HICs. Most items demonstrated means slightly above 3 on the 5-point scale, standard deviations of about one and negatively skewed distributions; however, the reverse-keyed items of the general emotional expressivity HIC demonstrated means closer to 2.5 and positively skewed distributions.

A preliminary test of the three-level hierarchical model of emotional expressivity consisted of a series of PFAs to examine support for the general factor, the two factors of positive and negative expressivity, and the discrete affects. A one-factor PFA of the initial item pool accounted for 40% of the common variance with item factor loadings ranging from |.73| - |.27|. The items from the general emotional expressivity HIC ("People can usually tell what kind of mood I'm in" and "My behavior usually shows how I'm feeling") loaded the strongest on this factor. All but seven items loaded greater than |.35| on this general factor. Compared to the discrete affect items, the general items contained more common variance related to the general factor as indicated by being the top loading items in the one-factor solution. Thus, the general concept of writing items at two levels of abstraction yielded promising results.

Based only on the discrete affect items (i.e., general expressivity items omitted), both Varimax- and Promax-rotated two-factor PFAs supported the hypothesized two-factor solutions. Although the Promax-rotation revealed a clearer pattern and allowed the two factors to correlate as expected, the Varimax-rotation allowed for a more stringent test to examine splitter items or cross-loadings. The Varimax-rotated two-factor PFA accounted for 53% of the common variance; 29 of the 31 items from the positive expressivity HICs loaded greater than .35 on the first factor and 29 of the 32 items from the negative expressivity HICs loaded greater than .35 on the second factor. Only two of the items from the positive expressivity HICs cross-loaded greater than .35 and none of the items from the negative expressivity HICs cross-loaded greater than .35. Overall, these one-factor and two-factor PFA solutions provided initial support for a general factor of emotional expressivity and two factors of positive and negative expressivity.

Separate PFAs of the items from the positive and negative expressivity HICs demonstrated the potential of the discrete affects to define distinct factors. Both Varimax- and Promax-rotated four-factor PFAs of the negative expressivity items revealed four interpretable factors. In Table 9, the Varimax-rotated four-factor PFA accounted for 84% of the common variance and the items from the hostility, sadness, guilt, and fear HICs generally defined the four factors, respectively. Both Varimax- and Promax-rotated four-factor PFAs of the positive expressivity items revealed three interpretable factors with an uninterpretable fourth factor. In Table 10, a Varimax-rotated three-factor PFA accounted for 85% of the common variance and the items from the self-assurance, joviality and amusement HICs generally defined the three factors, respectively.

The attentiveness HIC failed to emerge as a distinct factor and demonstrated the lowest AIC and coefficient alpha. Moreover, attentiveness may be more meaningful as a short-term construct than as a long-term trait and attentiveness may be expressed more subtly compared to some of the other discrete affects. Both of these considerations likely made it more difficult for these items to define a clear factor. Within the attentiveness HIC the items containing the word *interest* hung together better (AIC = .36) than the other items (AIC = .25). After shifting the emphasis more towards interest, this potential affect may be re-examined in future attempts to expand this measure to incorporate more discrete affects; however, at this point, the preliminary analyses and further consideration of the item content resulted in dropping the attentiveness HIC from the revised item pool to focus on creating a stronger assessment of the remaining seven discrete affects.

After dropping the items from the attentiveness HIC, both Varimax- and Promax-rotated three-factor PFAs of the remaining positive expressivity items revealed three interpretable factors. In Table 11, a Varimax-rotated three-factor PFA accounted for 91% of the common variance and the items from the self-assurance, joviality and amusement HICs generally defined the three factors, respectively.

Stopping short of creating preliminary scales, pilot-testing focused on how items functioned and resulted in several items being removed or revised. Overall, pilot-testing the items in the remaining HICs evaluated the item response distributions, interitem correlations, AICs, coefficient alphas, factor loadings from the one-factor and two-factor PFAs, and the four-factor and three-factor PFAs of the negative and positive expressivity HICs, respectively. Although the fear and sadness HICs emerged as factors, analyses indicated lower than optimal interitem correlations, AICs and coefficient alphas and therefore the revised item pool contained several new items for these two HICs. In summary, pilot-testing the initial 73-item pool resulted in 22 dropped items, 5 revised items and 17 new items to improve the item pool for the next two studies.

Table 9. Varimax-Rotated 4-Factor PFA of the Negative Expressivity HICs in the Pilot-Testing Sample

HIC	Item	I	II	III	IV
Hostil	My anger comes out in my tone of voice.	**.78**	.09	.11	.07
Hostil	If I'm feeling grouchy, it shows.	**.72**	.21	.13	.08
Hostil	Even if I'm only a little bit irritated, it still comes out in my voice.	**.66**	.20	.19	.15
Hostil	When I'm even a little bit irritated, everyone seems to know.	**.60**	.15	.31	.04
Hostil	People can tell when I'm feeling cranky.	**.55**	.33	.13	.07
Guilt	If I'm angry with myself, I show it.	**.53**	.25	.07	.27
Hostil	When I'm irritable, other people know to stay away.	**.53**	.04	.16	.15
Hostil	I raise my voice when I'm angry.	**.48**	.19	.15	.06
Sad	The slightest disappointment will show on my face.	**.48**	.32	.34	.19
Guilt	If I'm angry with myself, others can see it in my face.	**.46**	.28	.19	.32
Hostil	I will glare at people when I'm angry.	**.46**	.02	.01	.20
Sad	If I'm feeling discouraged, my shoulders will slump a little.	**.40**	.11	.32	.26
Sad	Even if I'm only a little disappointed, others can hear it in my voice.	**.37**	.32	.29	.24
Sad	People know when I'm sad.	.41	**.58**	.09	.02
Sad	I cry when I'm feeling sad.	.13	**.58**	.14	.11
Sad	People can tell if I'm feeling a little blue.	.42	**.54**	.07	-.05
Guilt	If I'm dissatisfied with myself, it will show in my face.	.30	**.54**	.16	.21
Sad	Sometimes my eyes tear up at sad movies.	.08	**.51**	.14	.09
Fear	I scream when I'm really frightened.	.02	**.45**	.02	.25
Sad	I tend to speak more quietly when I'm sad.	.25	**.44**	.16	.14
Guilt	I find it hard to look at someone if I'm feeling guilty.	.19	.13	**.75**	.18
Guilt	If I'm feeling ashamed I won't make eye contact.	.21	.06	**.68**	.16
Guilt	If I'm feeling guilty, it shows.	.21	.34	**.67**	.24
Guilt	People can tell if I'm feeling guilty.	.15	.45	**.53**	.18
Guilt	I hang my head when I'm ashamed.	.25	.22	**.42**	.21
Fear	When I am nervous, it shows.	.19	.19	.17	**.72**
Fear	My voice will shake when I'm nervous.	.16	-.02	.10	**.65**
Fear	I think my nervousness shows through.	.12	.11	.27	**.58**
Fear	I get a little shaky when I'm scared.	.11	.37	.01	**.53**
Fear	The slightest bit of fear seems to come through in my voice.	.11	.30	.19	**.41**
Fear	I speak more quickly when I'm scared.	.09	.30	.15	**.39**
Fear	I fidget when I'm anxious.	.14	.00	.25	**.32**

N = 210.

Hostil = Hostility HIC. Sad = Sadness HIC. Guilt = Guilt HIC. Fear = Fear HIC.

Note: The highest factor loading for each item is in bold.

Table 10. Varimax-Rotated 3-Factor PFA of the Positive Expressivity HICs in the Pilot-Testing Sample

HIC	Item	I	II	III
Self sure	If I'm feeling confident, it shows.	**.79**	.27	.08
Self sure	People can see when I'm feeling confident by the way I handle myself.	**.78**	.15	.22
Self sure	If I'm feeling strong, people can hear it in my voice.	**.78**	.18	.10
Self sure	I speak a little more loudly when I'm feeling confident.	**.68**	.11	.15
Self sure	If I'm feeling sure of myself then others can tell.	**.66**	.32	.16
Self sure	My voice reveals my confidence.	**.64**	.26	.18
Self sure	You can see my confidence in my face.	**.60**	.21	.10
Jovial	If I'm feeling cheerful, I sound quite chipper.	**.59**	.43	.16
Attent	My face generally shows if I'm interested in something.	**.53**	.42	.08
Attent	When I'm feeling attentive I sit up straight.	**.53**	.31	.08
Attent	People can tell if I'm feeling alert.	**.49**	.32	.09
Self sure	When I'm feeling proud, I walk a little taller.	**.44**	.06	.18
Jovial	I am quite animated when I'm excited.	**.42**	.39	.36
Attent	People can tell if I'm interested in the topic of conversation.	**.37**	.22	.13
Jovial	If I'm happy then my face will show it.	.37	**.62**	.21
Jovial	I beam with joy when I'm really happy.	.33	**.61**	.29
Jovial	My enthusiasm shows.	.44	**.59**	.18
Jovial	People can tell when I'm feeling good.	.50	**.52**	.16
Attent	If something is interesting to me, then it shows.	.43	**.45**	-.01
Attent	I will nod my head to show that I'm attentive.	.18	**.45**	.02
Attent	I make eye contact when I'm feeling attentive.	.27	**.43**	.16
Jovial	I can hardly sit still if I'm excited.	.20	**.37**	.33
Jovial	When I'm excited I start talking really fast.	.21	**.33**	.21
Laugh	I will laugh at the silliest little things.	.19	-.09	**.76**
Laugh	The slightest joke can make me laugh.	.17	.05	**.76**
Laugh	I'll laugh out loud when I'm watching TV.	.05	.34	**.63**
Laugh	Once I start laughing then I'll laugh at almost anything.	.24	-.17	**.61**
Laugh	If I find something funny, I laugh out loud.	.09	.44	**.56**
Laugh	I laugh at my own stories.	.17	.32	**.51**
Laugh	When I'm laughing really hard, my whole body shakes.	.05	.24	**.50**
Laugh	Sometimes I laugh so hard that my eyes get watery.	.04	.26	**.48**

N = 210.

Self sure = Self-Assurance HIC. Jovial = Joviality HIC. Attent = Attentiveness HIC. Laugh = Amusement HIC.

Note: The highest factor loading for each item is in bold.

Table 11. Varimax-Rotated 3-Factor PFA of the Positive Expressivity HICs After Dropping the Attentiveness HIC in the Pilot-Testing Sample

HIC	Item	I	II	III
Self sure	If I'm feeling confident, it shows.	**.80**	.25	.06
Self sure	People can see when I'm feeling confident by the way I handle myself.	**.80**	.16	.20
Self sure	If I'm feeling strong, people can hear it in my voice.	**.78**	.14	.11
Self sure	If I'm feeling sure of myself then others can tell.	**.70**	.32	.11
Self sure	My voice reveals my confidence.	**.66**	.22	.17
Self sure	I speak a little more loudly when I'm feeling confident.	**.65**	.11	.16
Self sure	You can see my confidence in my face.	**.64**	.18	.08
Jovial	If I'm feeling cheerful, I sound quite chipper.	**.57**	.45	.10
Jovial	People can tell when I'm feeling good.	**.51**	.50	.09
Self sure	When I'm feeling proud, I walk a little taller.	**.42**	.16	.14
Jovial	I beam with joy when I'm really happy.	.36	**.66**	.17
Jovial	If I'm happy then my face will show it.	.42	**.59**	.12
Jovial	My enthusiasm shows.	.48	**.58**	.09
Laugh	If I find something funny, I laugh out loud.	.10	**.51**	.48
Jovial	I can hardly sit still if I'm excited.	.17	**.50**	.23
Jovial	I am quite animated when I'm excited.	.40	**.49**	.28
Jovial	When I'm excited I start talking really fast.	.19	**.41**	.12
Laugh	I will laugh at the silliest little things.	.15	.07	**.77**
Laugh	The slightest joke can make me laugh.	.14	.16	**.75**
Laugh	Once I start laughing then I'll laugh at almost anything.	.21	-.06	**.64**
Laugh	I'll laugh out loud when I'm watching TV.	.06	.42	**.57**
Laugh	When I'm laughing really hard, my whole body shakes.	.05	.32	**.45**
Laugh	I laugh at my own stories.	.18	.41	**.44**
Laugh	Sometimes I laugh so hard that my eyes get watery.	.08	.31	**.43**

N = 210.

Self sure = Self-Assurance. Jovial = Joviality HIC. Laugh = Amusement HIC.

Note: The highest factor loading for each item is in bold.

CHAPTER 5. STRUCTURAL VALIDITY

In the structural phase, Loevinger emphasized "the distinction between the discovery of structure in a set of items and the imposition of structure upon the items" (1957, pp. 663-664). The structural phase aimed to create unidimensional, homogeneous scales (i.e., scales that measure one thing and only one thing) aligned with the theoretical structural model discussed earlier. In other words, the theoretical structural model informed the structural analyses but I also attempted to "listen to the data" to maximize structural fidelity. A bottom-up approach delineated the discrete affect structure first and then built from the first-order level to align the ISEE with the proposed three-level hierarchical model of emotional expressivity.

Structural Analyses

To support the unidimensionality of a scale, in addition to a strong coefficient alpha (i.e., > .80), the average interitem correlation (AIC) and all individual interitem correlations should fall in the range of .15 - .50 with the individual interitem correlations clustered near the AIC (Clark & Watson, 1995). Within this range, the target AIC of the scale should match the theoretical conceptualization of the construct—closer to .15 for broader constructs and closer to .50 for narrower constructs. Interitem correlations less than .15 indicate a lack of homogeneous item content, whereas, interitem correlations greater than .50 indicate redundancy between items.

In addition to examining estimates of internal consistency, factor analysis informed item selection. Factor analysis extracts the common variance from the interitem correlation matrix and then provides factor loadings, that is, quantitative estimates of the proportion of the items' variance attributable to the underlying latent factor (Floyd & Widaman, 1995). Items with factor loadings greater than .30 indicated potential candidate items for the scales. Clear marker items loaded strongly on one factor (i.e., greater than .30) with generally weak loadings on the other factors. Then aggregating several marker items cancelled out the specific variances of the items (i.e., variance not attributable to the factor) and left a scale score as a more reliable and valid indicator of the underlying dimension.

The theoretical three-level hierarchical model guided considerations related to the appropriate factor analyses to conduct. More specifically, to utilize factor analysis in hierarchical scale construction, the level of abstraction of the items needed to be considered in relation to the level of analysis. The inclusion of more abstract items (e.g., "I am emotionally expressive") obscured previous factor analyses designed to identify the lower order structure. Additionally, the proposed three-level hierarchical model of emotional expressivity necessitated oblique rotations (i.e., Promax) to allow correlated lower order factors.

To examine the proposed hierarchical structure, a second-order factor analysis of the correlation matrix of the discrete affect factors tested for the existence of higher order factors (Floyd & Widaman, 1995). Similar to extracting the common variance from the interitem correlation matrix, second-order factor analysis extracted the common variance from the correlation matrix of the first-order factors. Hierarchical factor solutions provided factor loadings of the first-order factors on higher order factors. More specifically, a second-order one-factor PFA tested for a higher order general factor of emotional expressivity and a second-order Promax-rotated two-factor PFA tested for the higher order factors of positive and negative expressivity.

Item Selection

After the bottom-up approach delineated the hierarchical structure, the creation of hierarchical scales depended on selecting items containing the greatest amount of variance related to each level of abstraction. Selecting the strongest markers of the higher order factors maximized the amount of common variance related to the higher order factors and minimized the specific variance related to lower order factors. For instance, the more abstract general emotional expressivity items needed to contain more common variance related to the general factor than the discrete affect items. Alternatively, a summed total of all items as an indicator of a general factor combined items containing specific variance related to the discrete affect factors.

Item selection involved balancing psychometric adjustments (i.e., raising coefficient alpha or AIC, removing high or low individual interitem correlations) against substantive concerns (i.e., matching the items to the theoretical construct, maintaining breadth and representiveness of item content, and avoiding redundancy). The primary item selection criteria included the following: (a) the item clearly marked the factor in both samples (i.e., load > .30 on the expected factor with low cross-loadings < .30), (b) the interitem correlations remained under .60, (c) item removal maintained coefficient alpha above .80, (d) item content was evaluated in relation to other marker items to maintain breadth and avoid redundancy.

Structural Validity of the EES and BEQ

Additionally, Clark and Watson (1995) recommended the inclusion of comparison scales in the structural phase. Based on the literature review in Chapter 2, the EES (Kring et al., 1994) and the BEQ (Gross & John, 1995) were included to examine the structural validity, psychometric properties and descriptive statistics of these measures. Having the same sample of participants complete existing measures of emotional expressivity and the revised item pool for the ISEE facilitated comparisons across measures. Then in the external phase of scale development, covered in a later chapter, convergent relations among these measures are addressed.

Method

The remaining chapters cover analyses of data from two studies: Study 1 is based on a Student Sample and Study 2 reports data from the Iowa Longitudinal Personality Project (ILPP) Sample. Accordingly, this section outlines the methods of these studies including the primary samples and procedures. Chapter 6 provides additional relevant information about the secondary samples and methods (i.e., the Retest Sample of Study 1 and the Informant Sample of Study 2).

Participants and Procedure

Student Sample

The Student Sample consisted of 387 undergraduate students from the University of Iowa. Participant recruitment occurred through two methods: (a) 334 students from Elementary Psychology participated as a partial fulfillment of a course research exposure requirement, and (b) 53 students from other psychology classes participated in exchange for extra credit. In small group testing sessions, participants completed questionnaire packets that took about 30-45 minutes to complete. The sample consisted of 57% females and 43% males with a mean age of 19.6 (range 18-34; one student omitted age). A subset of the Student Sample completed a 2-week online retest assessment and Chapter 6 provides more details about the Retest Sample.

ILPP Sample

On four prior occasions (1996, 1999, 2002, 2005) the Iowa Longitudinal Personality Project assessed a sample of young adults. Previous articles reported on the first three phases of this longitudinal project (Vaidya, Gray, Haig, Mroczek, & Watson, 2008; Vaidya, Gray, Haig, & Watson, 2002). The original sample of 759 students participated in 1996 as part of group testing sessions for an introductory psychology class. The assessments in 1999, 2002 and 2005 mailed a battery of questionnaires with responses returned from 392, 299, and 271 participants, respectively. In 2005, the 4th Phase of ILPP augmented the sample of 271 original ILPP participants with 127 new participants (e.g., spouses, friends, co-workers of the original ILPP participants) who provided other-ratings on the original participants and joined the ongoing longitudinal study by also completing self-ratings.

The 5th Phase of ILPP began in the summer of 2009. The University of Iowa Alumni Association provided up-to-date mail and email addresses for the 392 ILPP participants from the 2nd Phase—the first longitudinal assessment. Additionally, contact information from the 3rd and 4th Phases augmented the list of addresses. An email and a mailed postcard invited participants to complete an online assessment or to wait for a mailed packet of questionnaires. Internet searches (i.e., www.phonenumber.com, www.dexonline.com, www.people.yahoo.com, www.zabasearch.com, www.whitepages.com, www.intelius.com, and www.ussearch.com) attempted to locate mailing addresses and phone numbers for cases with incorrect mailing addresses. Additionally, I called participants to confirm addresses and inform participants of the two methods of participation. Finally, mailed reminder postcards and emailed reminder invitations served as a final step in the recruitment process.

In the 5th Phase, 344 participants completed either an online version ($N = 142$) or a mailed packet of questionnaires ($N = 202$) relating to personality, affectivity, life events, and psychopathology. The sample consisted of 253 (74%) females and 91 (26%) males. Based on data collected from a previous assessment, ages were available for 261 participants with a mean age of 31.1 (range 30-39) for the 5th Phase assessment. Ages were unavailable for 83 participants—originally 4th Phase Informants—who were incorporated to augment the 4th and 5th Phase self-ratings sample. Participants received $30 for filling out the questionnaires, which took approximately 1 hour to complete.

Measures

<u>Measures in both the Student and ILPP Samples</u>

Both the Student and ILPP Samples completed the 68-item revised item pool intended to assess general emotional expressivity (8 items) and the seven discrete affects—fear (9 items), hostility (8 items), guilt (9 items), sadness (11 items), joviality (7 items), self-assurance (8 items), amusement (8 items). The items were administered with a 1 to 5 response format (1= *disagree* and 5 = *agree*).

<u>Additional measures only in the Student Sample</u>

The EES (Kring et al., 1994) assesses emotional expressivity based on a general disposition to display emotions outwardly. The EES contains 17 items of a general nature without valence content or nonverbal channel specifications (e.g., “I think of myself as emotionally expressive"). Reverse-keyed items (e.g., "I hold my feelings in") comprise the majority of the scale. Participants indicated the extent to which the item applies to them on a 1 to 5 response scale (1 = *never true* and 5 = *always true*). See Chapter 2 for a more extensive review of the EES.

The BEQ (Gross & John, 1995) assesses emotional expressivity based on a hierarchical model. Scored either as a total scale or separately for the individual facets, the BEQ yields scores for a Total Scale (16 items), Positive Expressivity (4 items), Negative Expressivity (6 items), and Impulse Strength (6 items). Participants indicated agreement on 1 to 5 response format (1= *strongly disagree* to 5 = *strongly agree*). See Chapter 2 for a more extensive review of the BEQ.

Results

Initial Analyses

The Student Sample consisted of 387 participants with complete data. The ILLP Sample consisted of 347 participants; however, missing data resulted in the removal of three participants' data, for a final *N* of 344. In the rest of the ILPP sample, 319 participants completed all items and 25 participants omitted some items; 20 participants omitted only one item and five participants omitted between two to five items. A total of 36 missing items—imputed from the average of the five most strongly correlated items in the rest of the sample—equaled 0.2% of the 23,392 total item responses.

Structural Analyses

An initial top-down overview of the complete 68-item revised item pool provided tentative support for a broad general factor and two factors of positive and negative expressivity. In the Student Sample, a PFA one-factor solution of the revised 68-item pool accounted for 38% of the common variance (21% of the total variance in PCA) with all but seven items loading greater than |.30|. In the ILPP Sample, a PFA one-factor solution of the revised 68-item pool accounted for 44% of the common variance (27% of the total variance in PCA) with all but four items loading greater than |.30|.

In the Student Sample, the Promax-rotated PFA two-factor solution of the 60 discrete affect items accounted for 55% of the common variance (30% of the total variance in PCA); two clear factors, which correlated .26 with one another, emerged. In the ILPP Sample, the Promax-rotated PFA two-factor solution of the 60 discrete affect items accounted for 63% of the common variance (38% of the total variance in PCA); once again, two clear factors, which correlated .34 with one another, emerged. Although a few items failed to load on the general factor or on positive or negative expressivity factors, the next set of factor analyses—aimed at identifying the maximum number of discrete affect factors—retained all items.

Velicer's minimum average partial (MAP) test (Velicer, 1976 as cited in O'Connor, 2000a; Velicer, Eaton, & Fava, 2000) is a statistically based method of determining the number of factors to retain. The MAP test partials the first principal component from the correlations among a set of items and then calculates the average squared coefficient of the off-diagonal partial correlations. This process is repeated on the partialed correlation matrix—the second principal component is partialed and the average squared coefficient of the off-diagonal partial correlations is calculated—and continues iteratively until the number of components extracted and partialed is one less than the number of items. The number of factors to retain is indicated by the *minimum* average partial from this series, that is, the point where partialing the component is still accounting for systematic variance. The MAP test program (i.e., map.sas), retrieved online (O'Connor, 2000b), provided results for both the original MAP test (Velicer, 1976 as cited in O'Connor, 2000a) and the revised MAP test (Velicer, Eaton, & Fava, 2000).

Based on data for the 60 discrete affect items in the Student sample, the original MAP Test indicated nine factors could be extracted, whereas the revised MAP Test indicated seven factors could be extracted. Based on data for the 60 discrete affect items in the ILPP sample, both the original MAP Test and the revised MAP Test indicated eight factors could be extracted. See Table 12 for the average squared partial correlations and the average 4th power partial correlations. Table 12 demonstrates that although the program outputs results relevant to the revised MAP Test, the output of the 4th power partial correlations clearly needs to be extended to another digit or two to the right of the decimal point to make definitive comparisons.

Parallel analysis is another quantitative technique that provides information regarding the number of factors that should be retained (O'Connor, 2000a). It involves computing eigenvalues based on random data (representing the same numbers of variables and observations) and then comparing them to actually observed eigenvalues. As long as the actual eigenvalues exceed the simulated eigenvalues, it is reasonable to argue that meaningful variance is being modeled. However, when the random eigenvalues begin to exceed the actual values, the most likely explanation is that one is fitting error variance. The parallel analysis program (i.e., parallel.sas) was retrieved online (O'Connor, 2000c). Table 13 presents the actual and simulated eigenvalues for both the Student and ILPP Samples. These results supported the extraction of seven factors in the Student Sample and five factors in the ILPP Sample. Compared to parallel analyses on random data sets, parallel analyses based on permutations of actual raw data sets (i.e., column-wise random shuffling; rawpar.sas; O'Connor, 2000d) yielded nearly identical simulated eigenvalues and indicated the same numbers of factors to extract.

Based on the MAP test and parallel analyses, a series of factor analyses extracted three to eight factors in the discrete affect items; however, the factors ultimately needed to be well defined (i.e., a minimum of three clear markers) for the solution to be retained. In the Student Sample, seven well-defined factors emerged that were clearly interpretable as the seven discrete affect HICs. The eighth factor failed to have three clear marker items. In the ILPP Sample, five factors emerged, with three factors interpretable as representing the positive expressivity HICs (i.e., joviality, self-assurance and amusement); however, the negative expressivity HICs remained combined in unexpected ways—the fear and guilt HICs defined one factor and hostility and sadness HICs defined a second factor. In the ILPP Sample the sixth and subsequent factors remained uninterpretable.

To reconcile the factor solutions from the Student and ILPP Samples, the 37 negative expressivity items were examined separately in both samples. Four clear factors representing the four discrete affect HICs emerged in Student Sample, but two items from each of the fear, sadness, and guilt HICs failed to have their strongest loading on the expected factor. The remaining 31 markers items were then reexamined in the ILPP sample and a clear four-factor solution emerged.

After identifying a cleaner structure for the negative discrete affects, the positive and negative discrete affect items were recombined and clear seven-factor solutions emerged in both samples. Based on iterative analyses testing various combinations and the removal of specific items, the discrete affect scales finally included 44 items (6 items apiece for Sadness, Hostility, Guilt/Shame, Fear, and Amusement; and 7 items for Joviality and Confidence). In general, the discrete affects emerged more clearly in the Student Sample and therefore item removal more likely resulted from analyses in the ILPP Sample. Some items failed to meet all of the previously outlined item selection criteria, such as a few interitem correlations exceeded .60; moreover, in both samples the coefficient alpha for amusement remained below .80.

The Sadness scale contains items with affective terms clearly related to the core mood state, such as feeling down, depressed, or blue. The item about one's eyes revealing sadness (i.e., "If I'm feeling sad, other people can see it in my eyes") may capture subtle expressions of sadness whereas crying when sad probably has a higher threshold for the outward expression of sadness. The Hostility scale contains items about expressing anger and irritability through one's voice; other items incorporated expressing irritability in one's facial expressions and glaring. Expressing anger by glaring at another person probably has a higher threshold for outward expression of anger. The Guilt/Shame scale contains items about expressing guilt and shame through such nonverbal behaviors as avoiding eye contact and hanging or lowering one's head. The Fear scale contains items with affective terms such as fear, scared, afraid, and nervous; the items tapped such specific content as being shaky and jittery and having one's voice reveal fear and nervousness. The item, "I speak more quickly when I'm nervous," was envisioned in relation to public speaking, which is one of the most commonly reported fears (Seim & Spates, 2010). The Joviality scale, named after the analogous scale in the PANAS-X (Watson & Clark, 1999), contains items with positive affective terms such as cheerful, happy, enthusiasm, excited, and joy. The Confidence scale contains items about feeling strong and sure of oneself, and revealing this through one's voice and the way one handles oneself. The Amusement scale largely taps content related to laughing, with items covering a range of the spectrum from finding something amusing, to laughing out loud, to laughing so hard your eyes get watery. The item, "I laugh at my own stories," gets at the social nature of laughter and story telling.

In the Student Sample, the Promax-rotated PFA seven-factor solution accounted for 97% of the common variance (53% of the total variance in a seven-factor PCA); all 44 items loaded greater than .30 on the expected factor (see Table 14). In the ILPP Sample, the Promax-rotated PFA seven-factor solution accounted for 96% of the common variance (59% of the total variance in a seven-factor PCA); all 44 items loaded greater than .30 on the expected factor (see Table 15).

Based on a visual inspection of the factor loadings (i.e., the so-called "eyeball test") the structure of the discrete affects replicated well across both samples. All items loaded greater than .30 on the expected factor and very few items cross-loaded greater than .30 on secondary factors. Only one item replicated a cross-loading greater than .30 across both samples (i.e., "When I'm even a little bit irritated, everyone seems to know," which loaded primarily on hostility with a cross-loading on sadness). More generally, items containing "irritated" or "irritable"—which were intended to tap hostility—tended to cross-load on the sadness factor. Additionally, Varimax-rotated seven-factor PFAs provided a more stringent test of the cross-loadings. All items loaded greater than .30 on the expected factor across both samples; however, as expected, several items cross-loaded on other factors—typically another factor of the same valence. Sadness and hostility items tended to hang together and the guilt/shame and fear items tended to hang together.

Comparability coefficients provide a more formal, quantitative index of how well the factor structure replicates across the samples. This method involved calculating regression based factor scoring weights within each sample, then applying both sets of scoring weights to each sample and correlating the two sets of factor scores based on the Student and ILPP sample solutions. In Table 16 the comparability coefficients—ordered based on the strength of match—ranged from .98 to .94. The comparability coefficients for discrete affects demonstrated remarkable similarity across the samples, that is, .01 was the largest difference between the comparability coefficients of the same factor across the two samples (e.g., Sadness = .97 vs. .98 in the Student and ILPP samples, respectively). In general, all comparability coefficients exceeded the conventional .90 benchmark for factor similarity and supported strong structural replication (Everett, 1983).

In Table 17 and Table 18, the progression of correlation matrices demonstrated how the discrete affects emerged across successive factor analyses. A strong correlation between a factor score and discrete affect scale indicated the scale contained variance that was substantially related to the factor. All discrete affect scales correlated strongly with the general factor—all correlations exceeded .50. The inclusion of 24 negative expressivity items compared to only 20 positive expressivity items tipped the general factor slightly towards negative expressivity. In both samples, the extraction of a second factor revealed a clear split between the negative and positive expressivity discrete affect scales—the correlations exceeded .60 between scales and the expected factor. For the two- through seven-factor solutions, the strongest correlation for the discrete affect scales are in bold and underlined. Confidence emerged in the three-factor matrices, then Sadness and Hostility split from Guilt/Shame and Fear in the four-factor matrices, and then Sadness and Hostility emerged separately in the five-factor matrices. Continued in Table 18, Guilt/Shame split from Fear in the six-factor matrices; however, in the ILPP Sample, the Joviality scale failed to correlate clearly with one factor and instead correlated .57, .59, .59 across the first, third, and fifth factors. In the seven-factor matrices, each of the discrete affect scales clearly correlated with one of the factors. Across both samples, the successive factor analyses extracted the factors in similar order.

Overall, the correlations between the discrete affect scales and the one-, two- and seven-factor scores strongly supported the hierarchical structure of emotional expressivity. All discrete affect scales correlated strongly with the factor score based on the general factor (across both samples the mean r = .66; all mean correlations throughout the rest of this document are r-to-z transformed, averaged and then transformed back to rs). The two-factor matrix indicated a clear split between the positive and negative expressivity scales (across both samples the mean r = .78 with the expected factor score whereas the mean r = .17 with other factor score). Moreover, in the seven-factor matrix, each discrete affect scale clearly aligned with its corresponding factor (across both samples the mean r = .91 with its corresponding factor).

Building from the bottom-up

In the top half of Table 19, the seven discrete affect scales correlated moderately to strongly in both the Student Sample (below the diagonal) and the ILPP Sample (above the diagonal). As expected, the within-valence discrete affect scales tended to correlate more strongly with each other than with the cross-valence discrete affect scales. The within-valence discrete affect scales correlated with a mean of .41 and .53 in the Student and ILPP Samples, respectively; whereas, the cross-valence discrete affect scales correlated with a mean of .21 and .26 in the Student and ILPP Samples, respectively. In the ILPP Sample, the Sadness scale exhibited higher than expected cross-valence correlations with Joviality and Confidence (*r*s = .47 and .39, respectively). Overall, the discrete affect scales correlated slightly higher in the ILPP Sample than in the Student Sample.

In the lower half of Table 19, the seven discrete affect factors revealed a similar pattern of correlations as the discrete affect scales (factor correlations were derived from the Promax-rotated seven-factor PFAs in Tables 13 and 14). The within-valence discrete affect factors tended to correlate more strongly than the cross-valence discrete affect factors. The within-valence discrete affect factors correlated with a mean of .38 and .42 in the Student and ILPP Samples, respectively; whereas, the cross-valence discrete affect factors correlated with a mean of .17 and .18 in the Student and ILPP Samples, respectively.

Next, a second-order factor analysis extracted the common variance from the correlation matrix of the discrete affect factors. A second-order one-factor PFA tested for a general factor of emotional expressivity and a second-order Promax-rotated two-factor PFA tested for positive and negative expressivity factors. In the upper half of Table 20, results from second-order one-factor PFAs supported a general factor—all discrete affect factors loaded approximately .40 to .60 on this higher order dimension in both samples. In the lower half of Table 20, results from second-order Promax-rotated two-factor PFAs supported two clear higher order factors of negative and positive expressivity in both samples. All four negative discrete affect factors clearly loaded of the first factor and all three positive discrete affect factors clearly loaded on the second factor. The higher order factors of negative and positive expressivity correlated .43 and .36 in the Student and ILPP Samples, respectively. In the ILPP Sample, the sadness factor loaded the strongest on the general factor (i.e., loading .72) and also demonstrated a moderate cross-loading in the two-factor solution. Overall, these second-order factor analyses supported the existence of both a broad higher order factor of emotional expressivity and intermediate level factors of negative and positive expressivity.

Creating the Positive and Negative Expressivity Scales

The method for creating the Positive Expressivity (PE) and Negative Expressivity (NE) scales aimed to select a representative subset of the discrete affect items that contained the most common variance related to these higher order factors. In both samples, a Promax-rotated two-factor PFA of the complete set of 44 items from the seven discrete affect scales revealed two clear factors. The negative expressivity items defined the first factor and the positive expressivity items defined the second factor. However, the wide range of factor loadings clearly revealed that some of the discrete affect items contained more common variance related to the higher order factors. In addition to the previously outlined item selection criteria, the method for selecting items for the PE and NE scales emphasized selecting the strongest markers of the higher order factors. An iterative process examined several combinations of items with the interitem correlations as a key criterion. Ideal items exhibited relatively similar levels of interitem correlations within discrete affects as across similarly valenced discrete affects (e.g., sadness items correlated together approximately the same as sadness and fear items; this minimizes the covariances among items that would be unaccounted for by broad factors of positive and negative expressivity). Additionally, to maintain the representativeness of the discrete affects in relation to the positive and negative expressivity factors, PE and NE contained equal numbers of items from each discrete affect. Thus, NE contained 12 items (3 items from Sadness, Hostility, Guilt/Shame, and Fear), and PE contained 12 items (4 items from Joviality, Confidence, and Amusement).

In Tables 21 and 22, the column farthest on the right reports PFA one-factor solutions for the 24 items of PE and NE. In the Student Sample, a PFA one-factor solution of the 24 items accounted for 52% of the common variance (23% of the total variance in a PCA)—loadings ranged from .59 - .30. In the ILPP Sample, a PFA one-factor solution of the 24 items accounted for 57% of the common variance (30% of the total variance in a PCA)—loadings ranged from .67 - .29. Importantly, these one-factor solutions revealed intermixed factor loadings of the PE and NE items and similar average loadings for the PE and NE items (i.e., .46 vs. .44 and .55 vs. .48 in the Student and ILPP Samples, respectively).

In Tables 21 and 22, Promax-rotated PFA two-factor solutions of the 24 items indicated two clear factors with the first factor defined by all of the PE items and the second factor defined by all of the NE items. In the Student Sample, a two-factor solution of the 24 items accounted for 79% of the common variance (37% of the total variance in a PCA) with loadings ranging from .71 - .37 on the expected factors and all cross-loadings less than .20. In the ILPP sample, a two-factor solution of the 24 items accounted for 85% of the common variance (45% of the total variance in a PCA) with loadings ranging from .80 - .31 on the expected factors and all cross-loadings less than .30. The joviality items tended to be the strongest markers of the positive expressivity factor and the amusement items tended to be somewhat weaker markers of this factor. In both samples, the item "If I'm feeling down, it shows" loaded the strongest on the negative expressivity factor. The two factors correlated .30 and .32 and the PE and NE scales correlated .29 and .32 in the Student and ILPP Samples, respectively. In Table 23 the comparability coefficients exceeded .99.

Creating the ISEE General Scale

Similar to creating PE and NE, the method for creating the ISEE General scale aimed to select the items that contained the most common variance related to the higher order general factor. Moreover, the bottom-up approach of creating a hierarchical set of measures relied on selecting items aligned with the higher order factors built from the lower order factors. More specifically, the three stages included: (a) creating the discrete affect scales, (b) then using these scales in the process of creating PE and NE, and (c) using PE and NE in the process of creating the ISEE General scale. However, in addition to the common variance of the general factor, the discrete affect items clearly contained variance related to positive and negative expressivity, as well as specific variance related to the discrete affect factors. Therefore the inclusion of the eight general emotional expressivity items specifically targeted the common variance of the general factor.

In the Student Sample, a PFA one-factor solution of the 8 general expressivity items with the 24 PE and NE items accounted for 52% of the common variance (24% of the total variance in a PCA) with loadings ranging from |.71| - |.27|. All but one item loaded greater than |.30| on this general factor. Moreover, two general expressivity items loaded the strongest on this general factor (i.e., "I am emotionally expressive" and "If I'm experiencing a strong emotion, it shows"). All eight general expressivity items were among the top 11 loading items and all general expressivity items loaded greater than |.45| in the one-factor solution.

In the ILPP Sample, a PFA one-factor solution of the 8 general expressivity items and the 24 PE and NE items accounted for 57% of the common variance (31% of the total variance in a PCA) with loadings ranging from |.79| - |.26|. All but three items loaded greater than |.30| on this general factor. Moreover, the same two general expressivity items loaded the strongest on this general factor (i.e., "I am emotionally expressive" and "If I'm experiencing a strong emotion, it shows"). Seven of the general expressivity items were among the top 13 loading items and all general expressivity items loaded greater than |.45| on the one-factor solution. As expected, compared to the more specific discrete affect items, the general expressivity items contained more common variance related to the general factor.

The method for selecting items for the ISEE General scale included the previously outlined item selection criteria. An iterative process examined several combinations of items with the interitem correlations as a key criterion. In Table 24, PFA one-factor solutions of the six items selected for the ISEE General scale supported the general factor with loadings ranging from |.73| - |.50| and from |.82| - |.55| in the Student and ILPP Samples, respectively. It is noteworthy that the PE and NE scales correlated approximately equivalently with the ISEE General scale (*r*s = .50 and .49, respectively in the Student data and *r*s = .59 and .53, respectively in the ILPP data).

Descriptive Statistics

Table 25 reports the descriptive statistics for all the ISEE in the Student and ILPP Samples. The coefficient alphas all exceed .80, except for the Hostility (α = .77) and Fear (α = .78) scales in the Student Sample and the Amusement scale in both samples (αs = .74 and .75 in Student and ILPP Samples, respectively). AICs ranged from .51 to .29. The AICs of discrete affect scales exceeded the AICs of PE and NE—which is as expected, given that the latter are broader scales. However, the AIC of the ISEE General scale exceeded those for PE and NE (.40 vs. .29 and .29 in the Student Sample and .47 vs. .36 and .36 in the ILPP Sample). Clearly, the different levels of abstraction of the discrete affect items and the general emotional expressivity items influenced the AICs that were obtained.

Reporting the mean levels of the scales—based on the average of the items of each scale on a 1 to 5 response scale—allowed for analyses of sex differences across all scales and for comparisons across scales with differing numbers of items. Cohen's *d* quantified the effect size of the mean level sex differences. Females reported significantly higher levels of emotional expressivity on all scales except for Guilt/Shame and Confidence in the Student Sample. The sex difference for Confidence in the Student Sample indicated that males reported higher levels of Confidence compared to females. These moderate to strong sex differences tended to be slightly larger in the ILPP Sample compared to the Student Sample.

Comparisons across samples, but within sex, also revealed significant mean level differences. Compared to males in the ILPP Sample, males in the Student Sample reported significantly higher mean levels of ISEE General (d = 0.28), NE (d = 0.68), Sadness (d = 0.43), Hostility (d = 0.41), Guilt/Shame (d = 0.67), Fear (d = 0.70), Joviality (d = 0.26) and Confidence (d = 0.36). Compared to females in the ILPP Sample, females in the Student Sample reported significantly higher mean levels of NE (*d* = 0.23), Guilt/Shame (*d* = 0.24), and Fear (*d* = 0.62). All other mean level comparisons across samples, but within sex, yielded nonsignificant differences ($p > .05$).

In general, females reported higher levels of emotional expressivity than males and males in the Student Sample reported higher levels of expressivity than males in the ILPP Sample. Additionally, across sexes and samples, all mean levels for PE and the positive expressivity discrete affect scales (range 3.61 – 4.44) exceeded all mean levels for NE and the negative expressivity discrete affect scales (2.37 – 3.53). Within sexes and samples, the means of the ISEE General scale always fell between the means of PE and NE.

Table 26 reports the descriptive statistics for the EES and BEQ in the Student Sample. The EES and BEQ Total scale demonstrated strong coefficient alphas. Coefficient alphas for the BEQ facets ranged from .81 to .71. AICs ranged from .44 to .30. The EES and the ISEE General scale demonstrated comparable effect sizes for sex differences (*d* = 0.60 vs. *d* = 0.52, respectively); however, the sex differences on the BEQ tended to be larger compared to the EES or the ISEE scales. More specifically, the effect sizes for the BEQ Total scale and BEQ Positive Expressivity more than doubled the effect sizes for the ISEE General scale and PE.

Additionally, the descriptive statistics for the revised BEQ 3-item scales of general expressivity, positive expressivity and negative expressivity are reported as a point of comparison. It is interesting to note that the sex differences on the revised BEQ 3-item general expressivity scale were comparable to the EES and ISEE General scale.

Structural Validity of the EES and BEQ

In the Student Sample, a PFA one-factor solution of the17 EES items accounted for 92% of the common variance (49% of the total variance in a PCA) with loadings ranging from |.81| - |.43| (see Table 27). All 11 reverse-keyed items loaded positively on the factor and all six positively keyed items loaded negatively on the factor. The two items, “I am not very emotionally expressive” and “I think of myself as emotionally expressive” defined the ends of this bipolar factor (i.e., lack of emotional expressivity versus emotional expressivity). The lowest loading item, “I am often considered indifferent by others,” less clearly tapped emotional expressivity.

In the Student Sample a PFA one-factor solution of the 16 BEQ items accounted for 77% of the common variance (37% of the total variance in a PCA) with item factor loadings ranging from |.77| - |.29| (see Table 28). The positively keyed items loaded positively on the factor and the reverse-keyed items loaded negatively on the factor. The two items, “I am an emotionally expressive person” and “People often do not know what I am feeling” defined the ends of this bipolar factor (i.e., emotional expressivity versus lack of emotional expressivity). The item “I’ve learned it is better to suppress my anger than show it” had a low loading.

In the Student Sample (see Table 29), a Promax-rotated PFA three-factor solution of the 16 BEQ items accounted for all of the common variance (59% of the total variance in a PCA). The facets of the BEQ, Impulse Strength, Negative Expressivity and Positive Expressivity emerged and all items loaded strongest on the expected factor and greater than |.30|. Four items cross-loaded greater than .30 and two of these four items (“I am an emotionally expressive person” and “What I’m feeling is written all over my face”) also cross-loaded in both the preliminary study (i.e., CAP) and in another structural analysis (Dobbs et al., 2007). This Promax-rotation revealed correlated factors; Positive and Negative Expressivity correlated .32 and Positive and Negative Expressivity correlated .45 and .49 with Impulse Strength, respectively.

As a comparison to the BEQ results from the preliminary study, the revised 3-item scales of general expressivity, positive expressivity and negative expressivity were examined in the Student data. A PFA one-factor solution of all nine items supported a general factor with factor loadings ranged from |.75| - |.28| (see Table 30). Next, in the lower half of Table 30 a Promax-rotated two-factor PFA of just the six positive and negative expressivity items supported two clear factors. The three positive expressivity items defined the first factor and the three negative expressivity items defined the second factor with all primary loadings greater than |.45| and all cross-loadings less than |.15|.

Additionally, a Promax-rotated two-factor PFA of the 3 general items with the 6 positive and negative expressivity items tested the potential of the general expressivity items to increase the correlation between positive and negative expressivity. All nine items loaded greater than .30 on the first or second factor and the three general expressivity items loaded on the first factor with one general item splitting across both factors. As expected, the correlation between the factors increased to .47 compared to .31 without the general items.

Calculating comparability coefficients tested the replication of the BEQ factor structure across the Student Sample and CAP. In Table 31 comparability coefficients ranged from .99 to .94. In general the structure of the BEQ is stable across both samples; however, the comparability coefficients for the BEQ more closely resembled those for the ISEE discrete affect scales than for PE and NE. Moreover, the comparability coefficients for the revised BEQ 3-item positive and negative expressivity factors exceeded the comparability coefficients of the full BEQ (see Table 32).

Table 12. MAP Tests in the Student and ILPP Samples

Student			ILPP		
Factor	Squared	4^{th} Power	Factor	Squared	4^{th} Power
0	.0493	.0069	0	.0798	.0139
1	.0211	.0020	1	.0278	.0029
2	.0135	.0010	2	.0128	.0010
3	.0111	.0005	3	.0117	.0006
4	.0089	.0003	4	.0098	.0004
5	.0083	.0003	5	.0086	.0003
6	.0079	.0002	6	.0084	.0003
7	.0075	**.0002**	7	.0084	.0002
8	.0075	.0002	8	**.0080**	**.0002**
9	**.0074**	.0002	9	.0082	.0002
10	.0077	.0002	10	.0085	.0002
11	.0081	.0002	11	.0089	.0003
12	.0084	.0002	12	.0091	.0003
13	.0088	.0003	13	.0094	.0003

N = 387 (Student Sample). N = 344 (ILPP Sample).

Note: Bold average squared partial correlations and bold average 4^{th} power partial correlations indicate the number of factors to retain according to the output of the MAP program.

Table 13. Parallel Analyses in the Student and ILPP Samples

Student				ILPP			
Factor	Actual EV	Simulated EV	Permutation EV	Factor	Actual EV	Simulated EV	Permutation EV
1	12.15	1.86	1.86	1	15.95	1.94	1.93
2	5.84	1.79	1.79	2	6.80	1.84	1.84
3	3.47	1.73	1.72	3	2.91	1.78	1.78
4	2.91	1.68	1.68	4	2.57	1.72	1.73
5	2.05	1.63	1.63	**5**	**2.05**	**1.68**	**1.67**
6	1.76	1.60	1.59	6	1.61	1.63	1.63
7	**1.64**	**1.56**	**1.55**	7	1.43	1.59	1.60
8	1.45	1.52	1.52	8	1.42	1.55	1.55
9	1.32	1.49	1.48	9	1.13	1.51	1.52
10	1.21	1.45	1.45	10	1.06	1.48	1.48

N = 387 (Student Sample). N = 344 (ILPP Sample).

Note: The point before simulated and permutation eigenvalues (EV) exceed actual eigenvalues (in bold) suggests the number of meaningful factors to retain.

Table 14. Promax-Rotated 7-Factor PFA of the ISEE Discrete Affects in the Student Sample

ISEE Discrete Affect Items	1	2	3	4	5	6	7
Confidence							
If I'm feeling confident, it shows.	**.79**	.07	.02	.01	-.02	.02	-.08
People can see when I'm feeling confident by the way I handle myself.	**.70**	.05	.00	.05	.09	.00	-.06
If I'm feeling strong, people can hear it in my voice.	**.66**	.00	.01	.10	-.07	-.01	.08
If I'm feeling sure of myself then others can tell.	**.59**	.10	-.04	.07	.04	-.04	.01
I speak a little more loudly when I'm feeling confident.	**.58**	-.12	.04	-.08	.05	.06	.11
When I'm feeling proud, I walk a little taller.	**.57**	.00	.03	-.12	.01	.12	.02
My voice reveals my confidence.	**.47**	.17	.06	-.09	.18	-.09	-.10
Joviality							
If I'm feeling cheerful, people can see it in my smile.	-.04	**.76**	.03	.02	-.15	-.01	-.03
If I'm happy then my face will show it.	.10	**.71**	-.02	-.03	-.02	.05	.01
My enthusiasm shows.	.12	**.65**	.02	.02	.07	-.12	-.03
People can tell when I'm feeling happy.	.01	**.60**	.10	.03	.02	-.07	.11
If I'm feeling excited, it shows.	.00	**.56**	-.01	-.04	.06	.10	.05
I beam with joy when I'm really happy.	.08	**.53**	-.01	.11	-.11	.08	.13
I am quite animated when I'm excited.	-.02	**.43**	-.04	.02	.08	.06	.14
Guilt/Shame							
If I'm feeling guilty, it shows.	.04	-.05	**.84**	.09	-.10	-.01	.04
I find it hard to look at someone if I'm feeling guilty.	-.05	.07	**.75**	-.20	.04	.06	.03
If I'm feeling ashamed I won't make eye contact.	-.07	.08	**.71**	-.12	.07	.07	.00
People can tell if I'm feeling guilty.	.14	.00	**.62**	.12	-.10	-.04	-.12
If I'm ashamed of myself, it will show in my face.	.15	-.04	**.43**	.16	-.06	.08	.05
I hang my head when I'm ashamed.	.02	.03	**.39**	.07	-.05	.15	-.06
Sadness							
People know when I'm sad.	-.03	.09	-.09	**.73**	.01	-.04	-.03
People can tell if I'm feeling a little blue.	.10	-.14	.04	**.71**	-.10	-.08	.11
People can tell if I'm feeling a little depressed.	.02	-.10	-.01	**.65**	.16	.03	.06

Table 14. Continued

ISEE Discrete Affect Items	1	2	3	4	5	6	7
If I'm feeling sad, other people can see it in my eyes.	-.04	.17	.01	**.60**	.07	.08	-.06
If I'm feeling down, it shows.	-.03	.07	.01	**.54**	.20	.19	-.12
I cry when I'm feeling sad.	-.18	.19	-.07	**.42**	-.05	.29	.01
Hostility							
My anger comes out in my tone of voice.	-.02	.05	-.06	-.01	**.73**	.01	.01
I will glare at people when I'm angry.	.07	-.05	-.10	.01	**.61**	.03	.11
Even if I'm only a little bit irritated, it still comes out in my voice.	.03	-.06	.06	.22	**.54**	-.02	.00
When I'm irritable, people can see it in my face.	-.01	-.01	.10	.28	**.52**	-.08	-.01
When I'm even a little bit irritated, everyone seems to know.	-.01	-.07	.15	.32	**.48**	-.10	.00
I raise my voice when I'm angry.	.15	.02	-.13	-.16	**.46**	.16	-.03
Fear							
I get a little shaky when I'm scared.	.10	.09	-.01	.07	-.08	**.64**	.06
I get jittery when I'm afraid.	.05	.02	.04	.03	.02	**.60**	.09
If I'm feeling a little nervous, you can hear it in my voice.	-.04	-.08	.21	.02	.07	**.59**	-.08
I think my nervousness shows through.	-.10	-.08	.23	-.04	.11	**.55**	.01
I speak more quickly when I'm nervous.	.12	.01	-.03	-.05	.00	**.48**	-.02
The slightest bit of fear seems to come through in my voice.	-.11	.00	.23	.19	.06	**.34**	.00
Amusement							
If I find something funny, I laugh out loud.	-.05	.10	.00	-.06	.11	-.07	**.67**
I'll laugh out loud when I'm watching TV.	.03	.06	.01	-.13	.13	.03	**.62**
The slightest joke can make me laugh.	-.08	.07	.06	.20	-.15	-.04	**.57**
Sometimes I laugh so hard that my eyes get watery.	-.04	.08	-.11	.00	.01	.09	**.46**
If I find something amusing, it shows.	.04	.24	.10	-.03	.09	-.10	**.45**
I laugh at my own stories.	.22	-.11	-.09	.12	-.11	.17	**.43**

N = 387.

Note: The highest factor loading for each item is in bold.

Table 15. Promax-Rotated 7-Factor PFA of the ISEE Discrete Affects in the ILPP Sample

ISEE Discrete Affect Items	1	2	3	4	5	6	7
Confidence							
If I'm feeling strong, people can hear it in my voice.	**.80**	.06	.07	-.05	.01	-.03	.03
If I'm feeling confident, it shows.	**.79**	.10	-.10	.07	-.06	-.01	-.01
People can see when I'm feeling confident by the way I handle myself.	**.74**	.07	-.01	.05	-.11	-.04	.01
If I'm feeling sure of myself then others can tell.	**.72**	.14	-.06	.06	.00	.03	-.09
My voice reveals my confidence.	**.62**	.17	-.01	-.06	.07	.07	-.10
I speak a little more loudly when I'm feeling confident.	**.52**	.08	.13	-.23	.21	-.11	.13
When I'm feeling proud, I walk a little taller.	**.37**	.36	.04	-.14	.12	-.02	.13
Joviality							
If I'm feeling cheerful, people can see it in my smile.	.06	**.72**	.07	-.03	-.03	.05	.02
My enthusiasm shows.	.11	**.64**	-.03	.02	.03	.05	.03
If I'm feeling excited, it shows.	.08	**.60**	.05	.12	-.04	-.02	-.03
I beam with joy when I'm really happy.	.23	**.59**	.02	.07	.02	-.01	.04
People can tell when I'm feeling happy.	.16	**.57**	-.03	.18	-.09	.17	-.07
If I'm happy then my face will show it.	.15	**.49**	-.08	.12	-.07	.33	.01
I am quite animated when I'm excited.	.14	**.37**	.13	.08	.01	.10	-.02
Guilt/Shame							
If I'm feeling guilty, it shows.	.13	-.04	**.75**	.15	-.05	.03	-.10
I find it hard to look at someone if I'm feeling guilty.	-.07	.06	**.74**	-.08	.03	-.01	.08
If I'm feeling ashamed I won't make eye contact.	-.08	.13	**.73**	-.15	.04	.04	.10
People can tell if I'm feeling guilty.	.09	-.03	**.62**	.14	-.02	.03	-.14
If I'm ashamed of myself, it will show in my face.	-.12	.11	**.58**	.35	-.10	-.04	-.08
I hang my head when I'm ashamed.	.07	.03	**.54**	-.04	.06	-.14	.18
Sadness							
If I'm feeling sad, other people can see it in my eyes.	.03	.11	-.03	**.70**	.01	-.05	.14
People know when I'm sad.	-.09	.21	.06	**.68**	.06	-.13	-.07
People can tell if I'm feeling a little depressed.	.02	.01	.02	**.66**	.16	-.07	.06

Table 15. Continued

ISEE Discrete Affect Items	1	2	3	4	5	6	7
If I'm feeling down, it shows.	-.02	.11	.14	**.60**	.08	-.04	.12
People can tell if I'm feeling a little blue.	-.06	.39	-.02	**.51**	.20	-.18	.03
I cry when I'm feeling sad.	-.08	.13	.07	**.38**	-.10	.20	.12
Hostility							
My anger comes out in my tone of voice.	-.03	.09	.01	.10	**.66**	.17	-.02
I will glare at people when I'm angry.	.13	-.04	-.05	-.05	**.65**	-.04	.09
I raise my voice when I'm angry.	-.10	.11	-.04	.07	**.59**	.02	-.01
When I'm even a little bit irritated, everyone seems to know.	.08	-.20	.08	.41	**.46**	.03	-.14
Even if I'm only a little bit irritated, it still comes out in my voice.	-.06	-.09	.13	.37	**.46**	.07	-.03
When I'm irritable, people can see it in my face.	.12	-.14	.01	**.44**	.40	.01	.07
Amusement							
I'll laugh out loud when I'm watching TV.	-.03	-.01	-.02	-.17	.07	**.75**	.00
If I find something funny, I laugh out loud.	-.10	.15	.03	-.09	.10	**.67**	-.03
The slightest joke can make me laugh.	.02	-.02	.08	-.04	-.02	**.50**	-.01
Sometimes I laugh so hard that my eyes get watery.	-.01	.15	-.04	-.03	-.05	**.49**	.17
If I find something amusing, it shows.	.12	.25	-.06	.13	.04	**.45**	-.09
I laugh at my own stories.	-.04	.30	.01	-.01	.09	**.34**	.01
Fear							
I get a little shaky when I'm scared.	-.05	.13	-.04	.08	-.04	-.04	**.74**
I get jittery when I'm afraid.	-.08	.09	.03	.02	.08	-.02	**.69**
If I'm feeling a little nervous, you can hear it in my voice.	.20	-.23	.24	.21	-.08	.15	**.41**
I speak more quickly when I'm nervous.	.27	-.15	.10	-.01	.05	.00	**.39**
I think my nervousness shows through.	.00	-.13	.25	.30	-.09	.10	**.38**
The slightest bit of fear seems to come through in my voice.	-.12	-.05	.24	.18	.02	.09	**.35**

N = 344.

Note: The highest factor loading for each item is in bold.

Table 16. Comparability Coefficients from Promax-Rotated 7-Factor PFAs of the ISEE Discrete Affects in the Student and ILPP Samples

	Scores based on ILPP Solution						
Student Data	1	2	3	4	5	6	7
Scores based on Student Solution							
1 Guilt/Shame	**.98**	.32	.49	.54	.16	.27	.27
2 Confidence	.18	**.97**	.02	.15	.51	.23	.30
3 Fear	.61	.07	**.97**	.53	.07	.19	.26
4 Sadness	.50	.25	.43	**.97**	.28	.38	.26
5 Joviality	.18	.51	.12	.33	**.97**	.08	.64
6 Hostility	.33	.30	.26	.56	.04	**.95**	.30
7 Amusement	.18	.28	.20	.20	.57	.17	**.94**

	Scores based on Student Solution						
ILPP Data	1	2	3	4	5	6	7
Scores based on ILPP Solution							
1 Guilt/Shame	**.98**	.60	.17	.13	.71	.51	.13
2 Sadness	.65	**.98**	.37	.44	.57	.70	.26
3 Confidence	.32	.44	**.97**	.64	.24	.41	.38
4 Joviality	.10	.37	.62	**.96**	.05	.14	.62
5 Fear	.57	.45	.13	.06	**.96**	.40	.09
6 Hostility	.44	.56	.31	.18	.38	**.95**	.27
7 Amusement	.23	.30	.41	.67	.19	.36	**.94**

N = 387 (Student Sample). N = 344 (ILPP Sample).

Note: Convergent correlations are in bold.

Table 17. Correlations Between the ISEE and Factor Scores from 1-Factor and Promax-Rotated 2-, 3-, 4-, 5-Factor PFAs in the Student and ILPP Samples

Student Sample	1-Factor	2-Factor		3-Factor			4-Factor				5-Factor				
Scales	1	1	2	1	2	3	1	2	3	4	1	2	3	4	5
Sadness	.71	**.79**	.18	**.78**	.26	.03	.48	.32	**.68**	-.11	.38	.20	.00	**.87**	.25
Hostility	.59	**.61**	.20	**.61**	-.04	.42	.16	.04	**.88**	.21	.21	.08	.17	.32	**.90**
Guilt/Shame	.68	**.73**	.20	**.73**	.13	.21	**.87**	.08	.12	.28	**.86**	.06	.28	.17	.05
Fear	.62	**.79**	.03	**.78**	.18	-.11	**.80**	.17	.25	-.10	**.80**	.19	-.10	.24	.18
Joviality	.67	.19	**.84**	.14	**.87**	.31	.12	**.87**	.10	.33	.06	**.80**	.42	.33	-.09
Confidence	.56	.09	**.78**	.07	.30	**.91**	.09	.26	.13	**.92**	.09	.22	**.93**	.02	.15
Amusement	.54	.19	**.63**	.15	**.78**	.07	.11	**.79**	.09	.07	.15	**.86**	.07	-.03	.17
ILPP Sample	1-Factor	2-Factor		3-Factor			4-Factor				5-Factor				
Scales	1	1	2	1	2	3	1	2	3	4	1	2	3	4	5
Sadness	.82	**.76**	.41	**.75**	.40	.22	.52	.34	.22	**.65**	.48	.21	.14	.48	**.69**
Hostility	.69	**.72**	.26	**.72**	.26	.14	.31	.11	.15	**.90**	.34	.16	.14	**.91**	.15
Guilt/Shame	.63	**.84**	.06	**.83**	.04	.07	**.89**	.10	.06	.23	**.88**	.06	.07	.16	.23
Fear	.60	**.81**	.03	**.81**	-.05	.13	**.85**	.01	.13	.23	**.86**	.13	.04	.22	.02
Joviality	.73	.14	**.91**	.13	**.77**	.56	.09	**.79**	.54	.13	.05	.59	**.61**	.00	.55
Confidence	.71	.23	**.80**	.19	.29	**.93**	.13	.30	**.93**	.18	.14	**.95**	.22	.18	.10
Amusement	.53	.09	**.67**	.09	**.82**	.13	.05	**.83**	.11	.12	.07	.20	**.90**	.13	.07

N = 387 (Student Sample). N = 344 (ILPP Sample).

Note: Within each factor score and discrete affect scale correlation matrix, the strongest correlation for the discrete affect scales are in bold and underlined.

Table 18. Correlations Between the ISEE and Factor Scores from Promax-Rotated 6-, 7-Factor PFAs in the Student and ILPP Samples

<u>**Student Sample**</u>	6-Factor						7-Factor						
Scales	1	2	3	4	5	6	1	2	3	4	5	6	7
Sadness	.19	.01	.21	<u>**.89**</u>	.25	.30	.01	.21	.20	<u>**.90**</u>	.24	.33	.09
Hostility	.08	.15	.17	.32	<u>**.92**</u>	.08	.14	.05	.16	.32	<u>**.93**</u>	.12	.08
Guilt/Shame	.11	.17	<u>**.93**</u>	.20	.10	.24	.17	.11	<u>**.92**</u>	.21	.10	.29	.06
Fear	.11	-.01	.42	.31	.14	<u>**.85**</u>	.00	.06	.37	.25	.16	<u>**.91**</u>	.12
Joviality	<u>**.84**</u>	.35	.13	.30	-.06	-.06	.28	<u>**.93**</u>	.09	.15	.05	.08	.31
Confidence	.23	<u>**.96**</u>	.11	.03	.14	.04	<u>**.96**</u>	.25	.11	.04	.14	.01	.11
Amusement	<u>**.82**</u>	.08	.01	-.04	.17	.29	.14	.33	.06	.11	.06	.11	<u>**.94**</u>

<u>**ILPP Sample**</u>	6-Factor						7-Factor						
Scales	1	2	3	4	5	6	1	2	3	4	5	6	7
Sadness	.19	.45	.12	.52	<u>**.69**</u>	.21	.15	.36	.35	<u>**.81**</u>	.32	.10	.25
Hostility	.16	.29	.14	<u>**.92**</u>	.13	.14	.17	.05	.27	.41	<u>**.89**</u>	.14	.16
Guilt/Shame	.09	<u>**.96**</u>	.07	.19	.14	.17	.08	.09	<u>**.95**</u>	.25	.17	.04	.24
Fear	.10	.58	.03	.26	.09	<u>**.81**</u>	.13	-.02	.49	.29	.18	.07	<u>**.85**</u>
Joviality	.57	.05	<u>**.59**</u>	.01	<u>**.59**</u>	.03	.40	<u>**.86**</u>	.07	.20	.06	.38	.02
Confidence	<u>**.95**</u>	.11	.22	.18	.12	.08	<u>**.91**</u>	.38	.10	.09	.17	.16	.09
Amusement	.20	.06	<u>**.90**</u>	.13	.09	.02	.16	.36	.05	.06	.12	<u>**.90**</u>	.03

<u>N</u> = 387 (Student Sample). <u>N</u> = 344 (ILPP Sample).

Note: Within each factor score and discrete affect scale correlation matrix, the strongest correlation for the discrete affect scales are in bold and underlined.

Table 19. Intercorrelations Among the ISEE Discrete Affects and Among the Factors in the Student and ILPP Samples

Scales	1	2	3	4	5	6	7
1 Sadness	--	.61	.55	.53	.47	.39	.29
2 Hostility	.46	--	.45	.45	.26	.34	.24
3 Guilt/Shame	.41	.29	--	.61	.20	.23	.14
4 Fear	.49	.31	.51	--	.15	.24	.12
5 Joviality	.32	.18	.25	.18	--	.63	.56
6 Confidence	.13	.26	.26	.09	.44	--	.37
7 Amusement	.23	.17	.18	.21	.51	.27	--

Factors	1	2	3	4	5	6	7
1 Sadness	--	.45	.52	.35	.27	.39	.29
2 Hostility	.37	--	.37	.30	.10	.29	.22
3 Guilt/Shame	.45	.31	--	.52	.02	.21	.14
4 Fear	.42	.22	.45	--	-.02	.18	.05
5 Joviality	.30	.14	.21	.13	--	.46	.47
6 Confidence	.15	.18	.22	-.02	.44	--	.35
7 Amusement	.19	.13	.17	.18	.47	.25	--

N = 387 (Student Sample below the diagonals). N = 344 (ILPP sample above the diagonals).

Table 20. 1-Factor and Promax-Rotated 2-Factor PFAs of the Intercorrelation Matrix of the Discrete Affect Factors in the Student and ILPP Samples

1-Factor PFA Discrete Affect Scales	Student Factor 1		ILPP Factor 1
Sadness	**.63**	Sadness	**.72**
Guilt/Shame	**.60**	Guilt/Shame	**.60**
Joviality	**.55**	Confidence	**.57**
Fear	**.49**	Hostility	**.54**
Amusement	**.45**	Fear	**.47**
Hostility	**.44**	Amusement	**.46**
Confidence	**.40**	Joviality	**.42**

Promax 2-Factor PFA	Student			ILPP	
Discrete Affect Scales	Factor 1	Factor 2		Factor 1	Factor 2
Fear	**.65**	-.12	Guilt/Shame	**.74**	-.07
Sadness	**.63**	.08	Fear	**.65**	-.13
Guilt/Shame	**.62**	.05	Sadness	**.57**	.28
Hostility	**.43**	.07	Hostility	**.49**	.14
Joviality	.02	**.68**	Joviality	-.14	**.72**
Confidence	-.04	**.56**	Amusement	.01	**.59**
Amusement	.05	**.52**	Confidence	.17	**.56**

N = 387 (Student Sample). N = 344 (ILPP Sample).

Note: Factor 1 and Factor 2 correlated .43 in the Student Sample and .36 in the ILPP Sample. Factor loadings ≥ .30 are in bold.

Table 21. Promax-Rotated 2-Factor and 1-Factor PFAs of the ISEE Positive and Negative Expressivity in the Student Sample

Positive and Negative Expressivity Items	Promax 2-Factor PFA Factor 1	Promax 2-Factor PFA Factor 2	1-Factor PFA Factor 1
If I'm happy then my face will show it.	**.71**	.00	**.59**
My enthusiasm shows.	**.65**	-.02	**.52**
People can tell when I'm feeling happy.	**.63**	.08	**.58**
If I'm feeling cheerful, people can see it in my smile.	**.60**	-.04	**.47**
If I'm feeling confident, it shows.	**.60**	-.05	**.45**
If I'm feeling strong, people can hear it in my voice.	**.59**	-.03	**.46**
If I find something amusing, it shows.	**.55**	.07	**.51**
If I'm feeling sure of myself then others can tell.	**.53**	-.03	**.41**
My voice reveals my confidence.	**.49**	-.05	**.36**
If I find something funny, I laugh out loud.	**.46**	.04	**.41**
I'll laugh out loud when I'm watching TV.	**.44**	.05	**.40**
The slightest joke can make me laugh.	**.37**	.13	**.41**
If I'm feeling down, it shows.	-.01	**.66**	**.51**
The slightest bit of fear seems to come through in my voice.	-.11	**.63**	**.40**
I think my nervousness shows through.	-.19	**.61**	**.32**
If I'm feeling guilty, it shows.	.08	**.60**	**.54**
If I'm feeling sad, other people can see it in my eyes.	.11	**.58**	**.55**
When I'm irritable, people can see it in my face.	.04	**.54**	**.46**
I get jittery when I'm afraid.	-.01	**.54**	**.42**
If I'm feeling ashamed I won't make eye contact.	.04	**.53**	**.45**
Even if I'm only a little bit irritated, it still comes out in my voice.	.02	**.52**	**.42**
I cry when I'm feeling sad.	-.02	**.51**	**.38**
If I'm ashamed of myself, it will show in my face.	.13	**.47**	**.48**
My anger comes out in my tone of voice.	.06	**.37**	**.34**

N = 387 (Student Sample).

Note: Factor 1 and Factor 2 correlated .30 in the Student Sample. Factor loadings ≥ .30 are in bold.

Table 22. Promax-Rotated 2-Factor and 1-Factor PFAs of the ISEE Positive and Negative Expressivity in the ILPP Sample

	Promax 2-Factor PFA		1-Factor PFA
Positive and Negative Expressivity Items	Factor 1	Factor 2	Factor 1
If I'm happy then my face will show it.	**.80**	-.03	**.67**
People can tell when I'm feeling happy.	**.78**	-.03	**.66**
If I'm feeling sure of myself then others can tell.	**.72**	.00	**.62**
My voice reveals my confidence.	**.69**	-.05	**.56**
If I'm feeling confident, it shows.	**.68**	.01	**.60**
If I find something amusing, it shows.	**.68**	.03	**.61**
My enthusiasm shows.	**.67**	-.03	**.56**
If I'm feeling cheerful, people can see it in my smile.	**.67**	-.04	**.55**
If I'm feeling strong, people can hear it in my voice.	**.60**	.13	**.62**
If I find something funny, I laugh out loud.	**.52**	.00	**.45**
I'll laugh out loud when I'm watching TV.	**.45**	-.06	**.35**
The slightest joke can make me laugh.	**.31**	.03	.29
If I'm feeling down, it shows.	.13	**.71**	**.65**
I think my nervousness shows through.	-.13	**.71**	**.42**
If I'm feeling guilty, it shows.	-.01	**.70**	**.52**
When I'm irritable, people can see it in my face.	.08	**.65**	**.56**
Even if I'm only a little bit irritated, it still comes out in my voice.	.02	**.65**	**.50**
If I'm ashamed of myself, it will show in my face.	-.05	**.64**	**.44**
If I'm feeling sad, other people can see it in my eyes.	.20	**.61**	**.64**
The slightest bit of fear seems to come through in my voice.	-.13	**.59**	**.33**
If I'm feeling ashamed I won't make eye contact.	-.08	**.59**	**.37**
I get jittery when I'm afraid.	-.11	**.53**	**.30**
My anger comes out in my tone of voice.	.26	**.45**	**.56**
I cry when I'm feeling sad.	.20	**.41**	**.48**

N = 344 (ILPP Sample).

Note: Factor 1 and Factor 2 correlated .32 in the ILPP Sample. Factor loadings ≥ .30 are in bold.

Table 23. Comparability Coefficients from Promax-Rotated 2-Factor PFAs of the ISEE Positive and Negative Expressivity in the Student and ILPP Samples

Score	Student Factor 1	ILPP Factor 1	Student Factor 2	ILPP Factor 2
Student Data				
Student Factor 1	--			
ILPP Factor 1	**.994**	--		
Student Factor 2	.316	.317		
ILPP Factor 2	.341	.341	**.994**	--

Score	Student Factor 1	ILPP Factor 1	Student Factor 2	ILPP Factor 2
ILPP Data				
Student Factor 1	--			
ILPP Factor 1	**.996**	--		
Student Factor 2	.387	.380		
ILPP Factor 2	.361	.353	**.996**	--

N = 387 (Student Sample). N = 344 (ILPP Sample).

Note: Convergent correlations are in bold.

Table 24. 1-Factor PFAs for the ISEE General Scale in the Student and ILPP Samples

Student Sample	Factor 1
I am emotionally expressive.	.73
People can usually tell what kind of mood I'm in.	.60
My behavior usually shows how I'm feeling.	.58
If I'm experiencing a strong emotion, it shows.	.50
People have a hard time reading how I feel. **(R)**	-.64
I don't express my emotions. **(R)**	-.68

ILPP Sample	Factor 1
I am emotionally expressive.	.82
If I'm experiencing a strong emotion, it shows.	.66
People can usually tell what kind of mood I'm in.	.65
My behavior usually shows how I'm feeling.	.55
People have a hard time reading how I feel. **(R)**	-.66
I don't express my emotions. **(R)**	-.70

N = 387 (Student Sample). N = 344 (ILPP Sample).

(R) = Reverse-keyed items.

Table 25. Descriptive Statistics for the ISEE in the Student and ILPP Samples

Scale (# of items)	α	AIC	*M*	*SD*	Female *M*	Female *SD*	Male *M*	Male *SD*	*d*
Student Sample									
ISEE General (6)	.80	.40	3.82	.69	3.97_a	.67	3.62_b	.68	**0.52**
Negative Expressivity (12)	.83	.29	3.15	.65	3.33_a	.62	2.91_b	.61	**0.68**
Sadness (6)	.81	.42	3.30	.77	3.53_a	.72	2.97_b	.72	**0.78**
Hostility (6)	.77	.36	3.33	.74	3.41_a	.74	3.23_b	.73	**0.24**
Guilt/Shame (6)	.82	.43	2.96	.79	3.00_a	.81	2.92_a	.77	**0.10**
Fear (6)	.78	.37	3.27	.78	3.50_a	.69	2.97_b	.79	**0.71**
Positive Expressivity (12)	.83	.29	4.08	.50	4.13_a	.50	4.02_b	.51	**0.21**
Joviality (7)	.83	.41	4.32	.57	4.44_a	.52	4.17_b	.60	**0.48**
Confidence (7)	.83	.41	3.77	.65	3.71_b	.66	3.86_a	.62	**-0.23**
Amusement (6)	.74	.32	4.14	.62	4.27_a	.56	3.97_b	.67	**0.49**
ILPP Sample									
ISEE General (6)	.84	.47	3.87	.77	4.04_a	.68	3.41_b	.83	**0.83**
Negative Expressivity (12)	.87	.36	2.99	.76	3.18_a	.67	2.45_b	.74	**1.03**
Sadness (6)	.84	.47	3.28	.87	3.51_a	.76	2.63_b	.85	**1.09**
Hostility (6)	.84	.47	3.24	.89	3.36_a	.87	2.90_b	.87	**0.53**
Guilt/Shame (6)	.85	.49	2.69	.84	2.81_a	.80	2.37_b	.88	**0.52**
Fear (6)	.81	.42	2.86	.87	3.03_a	.82	2.40_b	.83	**0.76**
Positive Expressivity (12)	.87	.36	4.10	.54	4.15_a	.51	3.95_b	.58	**0.37**
Joviality (7)	.88	.51	4.26	.61	4.35_a	.55	4.00_b	.70	**0.56**
Confidence (7)	.88	.51	3.76	.69	3.81_a	.65	3.61_b	.77	**0.28**
Amusement (6)	.75	.33	4.12	.63	4.19_a	.60	3.94_b	.69	**0.39**

N = 387 (Student Sample; Female = 222, Male = 165). N = 344 (ILPP Sample; Female = 253 Male = 91).

Note: Different subscripts within a row indicate mean level sex differences at $p < .05$.

Table 26. Descriptive Statistics for the EES, BEQ and 3-Item BEQ in the Student Sample

Scale (# of items)	α	AIC	*M*	*SD*	Female *M*	Female *SD*	Male *M*	Male *SD*	*d*
EES (17)	.93	.44	3.26	.74	3.48_a	.74	2.97_b	.95	**0.60**
BEQ Total (16)	.87	.30	3.35	.64	3.60_a	.59	3.00_b	.53	**1.07**
Negative Expressivity (6)	.75	.34	2.89	.73	3.13_a	.72	2.57_b	.62	**0.83**
Positive Expressivity (4)	.71	.38	4.01	.64	4.15_a	.59	3.81_b	.66	**0.54**
Impulse Strength (6)	.81	.42	3.36	.86	3.71_a	.75	2.89_b	.77	**1.08**
Revised BEQ Scales									
3-Item General (3)	.76	.51	3.18	.85	3.40_a	.83	2.88_b	.79	**0.64**
3-Item Neg Exp (3)	.62	.35	2.75	.83	3.02_a	.80	2.38_b	.71	**0.85**
3-Item Pos Exp (3)	.67	.40	4.19	.61	4.31_a	.56	4.04_b	.65	**0.49**

N = 387 (Student Sample; Female = 222, Male = 165).

3-Item General = 3-Item General Expressivity. 3-Item Pos Exp = 3-Item Positive Expressivity. 3-Item Neg Exp = 3-Item Negative Expressivity.

Note: Different subscripts within a row indicate mean level sex differences at $p < .05$.

Table 27. 1-Factor PFA of the EES in the Student Sample

EES Items	Factor 1
I am not very emotionally expressive. **(R)**	.81
I hold my feelings in. **(R)**	.79
I keep my feelings to myself. **(R)**	.78
I don't express my emotions to other people. **(R)**	.75
Other people aren't easily able to observe what I'm feeling. **(R)**	.74
Even when I'm experiencing strong feelings, I don't express them outwardly. **(R)**	.73
Even if I am feeling very emotional, I don't let others see my feelings. **(R)**	.72
I don't like to let other people see how I'm feeling. **(R)**	.65
People think of me as an unemotional person. **(R)**	.61
The way I feel is different from how others think I feel. **(R)**	.51
I am often considered indifferent by others. **(R)**	.43
I am able to cry in front of other people.	-.52
I can't hide the way I'm feeling.	-.56
Other people believe me to be very emotional.	-.61
People can read my emotions.	-.70
I display my emotions to other people.	-.75
I think of myself as emotionally expressive.	-.79

N = 387.

EES = Emotional Expressivity Scale.

(R) = Reverse-keyed items.

Note: EES items Copyright © 1994 by the American Psychological Association. Adapted with permission. The official citation that should be used in referencing this material is Kring, A. M., Smith, D. A., & Neale, J.M. (1994). Individual differences in dispositional expressiveness: Development and validation of the Emotional Expressivity Scale. *Journal of Personality and Social Psychology, 66,* 934-949.

Table 28. 1-Factor PFA of the BEQ in the Student Sample

BEQ Items	Factor 1
I am an emotionally expressive person.	.77
What I'm feeling is written all over my face.	.73
I experience my emotions very strongly.	.73
My body reacts very strongly to emotional situations.	.68
I have strong emotions.	.66
Whenever I feel negative emotions, people can easily see exactly what I am feelings.	.64
There have been times when I have not been able to stop crying even though I tried to stop.	.58
I am sometimes unable to hide my feelings even though I would like to.	.58
I sometimes cry during sad movies.	.56
Whenever I feel positive emotions, people can easily see exactly what I am feeling.	.53
It is difficult for me to hide my fear.	.49
When I'm happy, my feelings show.	.46
I laugh out loud when someone tells me a joke that I think is funny.	.29
I've learned it is better to suppress my anger than to show it. **(R)**	-.31
No matter how nervous or upset I am, I tend to keep a calm exterior. **(R)**	-.41
People often do not know what I am feeling. **(R)**	-.57

N = 387.

BEQ = Berkeley Expressivity Questionnaire.

(R) = Reverse-keyed items.

Note: BEQ items reprinted from *Personality and Individual Differences*, v. 19, J.J. Gross and O.P. John, "Facets of emotional expressivity: Three self-report factors and their correlates," pp. 555-568, Copyright 1995, with permission from Elsevier.

Table 29. Promax-Rotated 3-Factor PFA of the BEQ in the Student Sample

BEQ Items	Factor 1	Factor 2	Factor 3
I experience my emotions very strongly.	**.83**	-.08	.07
I have strong emotions.	**.79**	-.22	.20
My body reacts very strongly to emotional situations.	**.55**	.12	.14
There have been times when I have not been able to stop crying even though I tried to stop.	**.54**	.27	-.17
I am sometimes unable to hide my feelings even though I would like to.	**.53**	.17	-.05
I sometimes cry during sad movies.	**.35**	**.33**	.00
Whenever I feel negative emotions, people can easily see exactly what I am feelings.	**.34**	**.56**	-.15
It is difficult for me to hide my fear.	.18	**.48**	-.06
What I'm feeling is written all over my face.	**.37**	**.46**	.08
I've learned it is better to suppress my anger than to show it. **(R)**	.23	**-.45**	-.27
People often do not know what I am feeling. **(R)**	.01	**-.52**	-.27
No matter how nervous or upset I am, I tend to keep a calm exterior. **(R)**	-.01	**-.58**	.08
When I'm happy, my feelings show.	.00	-.01	**.73**
Whenever I feel positive emotions, people can easily see exactly what I am feeling.	.01	.15	**.62**
I laugh out loud when someone tells me a joke that I think is funny.	.09	-.11	**.46**
I am an emotionally expressive person.	**.34**	.29	**.37**

N = 387.

BEQ = Berkeley Expressivity Questionnaire.

(R) = Reverse-keyed items.

Note: Factor loadings ≥ .30 are in bold.

Note: BEQ items reprinted from *Personality and Individual Differences*, v. 19, J.J. Gross and O.P. John, "Facets of emotional expressivity: Three self-report factors and their correlates," pp. 555-568, Copyright 1995, with permission from Elsevier.

Table 30. 1-Factor and Promax-Rotated 2-Factor PFAs of the Items Included in the Revised 3-Item BEQ Scales in the Student Sample

BEQ Items	Factor 1
I am an emotionally expressive person.	**.75**
What I'm feeling is written all over my face.	**.75**
Whenever I feel negative emotions, people can easily see exactly what I am feeling.	**.64**
Whenever I feel positive emotions, people can easily see exactly what I am feeling.	**.58**
When I'm happy, my feelings show.	**.51**
It is difficult for me to hide my fear.	**.48**
I laugh out loud when someone tells me a joke that I think is funny.	.28
No matter how nervous or upset I am, I tend to keep a calm exterior. **(R)**	**-.42**
People often do not know what I am feeling. **(R)**	**-.63**

BEQ Items	Factor 1	Factor 2
When I'm happy, my feelings show.	**.74**	-.02
Whenever I feel positive emotions, people can easily see exactly what I am feeling.	**.60**	.12
I laugh out loud when someone tells me a joke that I think is funny.	**.50**	-.10
Whenever I feel negative emotions, people can easily see exactly what I am feeling.	-.02	**.66**
It is difficult for me to hide my fear.	.05	**.54**
No matter how nervous or upset I am, I tend to keep a calm exterior. **(R)**	.06	**-.50**

N = 387.

BEQ = Berkeley Expressivity Questionnaire.

(R) = Reverse-keyed items.

Note: Factor loadings ≥ .30 are in bold.

Note: BEQ items reprinted from *Personality and Individual Differences*, v. 19, J.J. Gross and O.P. John, “Facets of emotional expressivity: Three self-report factors and their correlates,” pp. 555-568, Copyright 1995, with permission from Elsevier.

Table 31. Comparability Coefficients from Promax-Rotated 3-Factor PFAs of the BEQ in the Student and CAP Samples

Score	Student Factor 1	CAP Factor 1	Student Factor 2	CAP Factor 2	Student Factor 3	CAP Factor 3
Student Data						
Student Factor 1	--					
CAP Factor 1	**.99**	--				
Student Factor 2	.66	.67	--			
CAP Factor 2	.62	.61	**.95**	--		
Student Factor 3	.63	.59	.49	.61	--	
CAP Factor 3	.62	.57	.30	.42	**.96**	--

Score	Student Factor 1	CAP Factor 1	Student Factor 2	CAP Factor 2	Student Factor 3	CAP Factor 3
CAP Data						
Student Factor 1	--					
CAP Factor 1	**.99**	--				
Student Factor 2	.63	.66	--			
CAP Factor 2	.61	.63	**.97**	--		
Student Factor 3	.63	.63	.56	.68	--	
CAP Factor 3	.69	.66	.42	.54	**.94**	--

N = 387 (Student Sample). N = 394 (CAP Sample).

Note: Convergent correlations are in bold.

Table 32. Comparability Coefficients from Promax-Rotated 2-Factor PFA of the Items Included in the Revised 3-Item BEQ Scales in the Student and CAP Samples

Score	Student Factor 1	CAP Factor 1	Student Factor 2	CAP Factor 2
Student Data				
Student Factor 1	--			
CAP Factor 1	**.99**	--		
Student Factor 2	.55	.46	--	
CAP Factor 2	.52	.43	**.99**	--

Score	Student Factor 1	CAP Factor 1	Student Factor 2	CAP Factor 2
CAP Data				
Student Factor 1	--			
CAP Factor 1	**.99**	--		
Student Factor 2	.56	.63	--	
CAP Factor 2	.60	.66	**.99**	--

N = 387 (Student Sample). N = 394 (CAP Sample).

Note: Convergent correlations are in bold.

CHAPTER 6. NOMOLOGICAL NETWORK

"'Learning more about' a theoretical construct is a matter of elaborating the nomological network in which it occurs, or of increasing the definiteness of the components" (Cronbach & Meehl, 1955, p. 290). Cronbach and Meehl (1955) acknowledged that early attempts at conceptualizing a construct exist without an elaborate nomological network; however they emphasized the importance of understanding the conceptualization within the nomological network as construct validity accumulated. The previous chapter focused on the increasing the definiteness of the components of emotional expressivity within a hierarchical model and this chapter focuses on elaborating the nomological network of emotional expressivity, that is, establishing connections with the universal hierarchical model of personality (Markon et al., 2005) and the three-level hierarchical model of affectivity (Tellegen et al., 1999a, 1999b).

The universal hierarchical model of personality (Markon et al., 2005) outlines the broadest framework within which to locate and understand emotional expressivity. Chapter 1 briefly covered how the basic factors of personality emerged from the two-factor model to the five-factor model. Importantly, Negative Emotionality remained very similar across the three-, four- and five-factor levels of abstraction and led to the expectation of similar nomological connections between emotional expressivity and Negative Emotionality, or N. Positive Emotionality remained very similar across the three- and four-factor levels of abstraction; however, at the five-factor level, O emerged from a broader Positive Emotionality factor. In this regard, it should be noted that positive expressivity and positive discrete affects previously have demonstrated connections with O (Gross & John, 1995; Trierweiler et al., 2002). Therefore the current examination of the nomological connections between emotional expressivity and personality focused at the level of the five-factor model of personality.

As previously mentioned, at the broadest level of emotional expressivity, Riggio and Riggio (2002), in a recent meta-analysis, reported a significant relation between emotional expressivity and E but a nonsignificant relation between emotional expressivity and N. However, the EES and the BEQ Total scale correlated with both E and N (Gross & John, 1995; Kring et al., 1994) and results from the preliminary study also identified significant relations between the general factor of emotional expressivity and both E and N. The nomological connections of the general factor of emotional expressivity need to be clarified to facilitate comparisons within the hierarchical model, that is, a general factor of emotional expressivity may lack the nomological connections of the lower order factor of negative expressivity.

In comparison to the general factor, the lower order levels of emotional expressivity clearly aligned with either E or N. The results from the preliminary study and the scales created from the structural analyses of six emotional expressivity questionnaires demonstrated clear differential relations of positive and negative expressivity with E and N, respectively (Gross & John, 1998). With the increased discriminant validity of PE and NE, I re-examined the nomological network of the second-order level. In addition to relations with E, positive expressivity previously revealed some specificity with weak relations with O and A (Gross & John, 1995); however, in the current study no specific predictions were made about the level of relations between emotional expressivity and the remaining Big Five traits.

After extending the hierarchical model of emotional expressivity into the discrete affects level, I examined the nomological connections of these discrete affects. In previous research (Trierweiler et al., 2002), across both self- and peer-ratings, the positive discrete affects correlated with E, A and Intellect and the negative discrete affects correlated with N; however, in addition, the peer-ratings of anger correlated negatively with A and shame correlated positively with A. The current examination of the nomological network of the discrete affects focuses on replicating the differential relations with E and N and demonstrating specificity in relation to the other Big Five traits.

The nomological network of emotional expressivity also needs to be examined in relation to the three-level hierarchical model of affect—a general bipolar factor of happiness at the highest level, two relatively independent dimensions of PA and NA at the next level, and multiple discrete affects at the lowest level (Tellegen et al., 1999a, 1999b; Watson & Tellegen, 1999). In particular, PA and NA served as models for the development of PE and NE and the potential existed for these second-order levels of trait affect and emotional expressivity to be strongly related. However, in the preliminary study, both emotional expressivity and trait affect aligned more closely with E and N than with each other. The moderate magnitude of the correlations suggested quite a bit of differentiation between emotional expressivity (i.e., the tendency to express affective states through nonverbal means) and trait affectivity (i.e., in general, what one subjectively experiences or feels). With increased discriminant validity for PE and NE it is important to re-examine the nomological connections with PA and NA.

Furthermore, extending the hierarchical model of emotional expressivity into the discrete affects level allowed the nomological network of the ISEE and PANAS-X discrete affect scales (i.e., fear, hostility, guilt, sadness, joviality, and confidence/self-assurance) to be examined in an intercorrelation matrix similar to a MTMM (Campbell & Fiske, 1959). The monoaffect-heteroscale correlations (i.e., ISEE fear correlated with PANAS-X fear) forms the validity diagonal of the heteroaffect-heteroscale block. Comparisons of monoaffect-heteroscale correlations with other correlations in the heteroaffect-heteroscale block, allows for an examination of the distinctiveness of the connections between the same discrete affect across domains. However, the influence of the higher order dimensions of PE and NE and PA and NA leads to the expectation of stronger correlations within the monoscale triangles compared to the monoaffect-heteroscale correlations.

Conceptualizing emotional expressivity within the nomological network facilitates generating tests of the relations between multiple constructs. The hierarchical models of emotional expressivity and trait affect exist at similar levels of abstraction; however, both emotional expressivity and trait affect exist at narrower levels of abstraction compared to the universal model of personality; the broader factors of personality encompass related components of affect and emotional expressivity. To examine relations between personality, trait affect and emotional expressivity, hierarchical regressions were used to test the unique predictive validity of emotional expressivity and trait affect in predicting personality. In the preliminary study, regression analyses provided initial support for the incremental predictive validity of emotional expressivity and trait affect and these analyses are repeated in these data.

The first step of these regression analyses predicts E and N from PE and NE, respectively. The next step adds PA as a predictor of E and added NA as a predictor of N. Then, the reversal of the order of entry for the predictors examines the differential predictive validity of expressivity and affectivity. A parallel set of regressions analyses tested the differential predictive validity of the discrete affect scales of the ISEE and PANAS-X as predictors of E and N.

Method

Participants and Procedures

Student and Retest Sample

Chapter 5 provided details about the initial assessment of the Student Sample (*N* = 387). At the initial assessment, participants provided their name and email address to indicate interest in participating in an optional online 2-week retest assessment. On the morning of the 14th day after the initial assessment, participants received email invitations containing a link to the online assessment. The Retest Sample consisted of 352 participants, which is 91% of the initial sample. Participants received additional credit towards the research requirement or additional extra credit in exchange for participating.

The participants completed the retest assessment at the time and place of their choosing; however, at the initial assessment I asked participants to set aside the same time of day as the initial assessment to complete the retest assessment. When participants completed the online retest assessment, a time date stamp determined the test retest interval. Approximately 70% of the Retest Sample ($N = 246$) submitted the retest assessment within a day of receiving the email invitation and roughly 95% of the Retest Sample ($N = 330$) submitted the retest assessment within a week of receiving the email invitation.

ILPP and Informant Samples

Chapter 5 also provided a description of the ILPP Sample ($N = 344$). As an optional component to the ILPP assessment, participants provided contact information for two people as potential sources for other-ratings. An email and a mailed postcard invited informants to complete an online assessment or to wait for a mailed packet of questionnaires. Additionally, mailed reminder postcards and emailed reminder invitations served as a final step in the recruitment process. ILPP informants completed questionnaires rating the personality, trait affect, and emotional expressivity of the original ILPP participant. Completing the questionnaires took approximately 30 minutes and informants received $15 in exchange for participating.

The Informant Sample initially consisted of 238 other-ratings; however, as described below in the initial analyses, seven other-ratings were removed. The remaining 231 other-ratings matched with 174 self-ratings—117 cases with only a single other-rating and 57 cases with two other-ratings. In the 57 cases with two other-ratings, the average of the scale scores of the two other-ratings formed a composite rating. Thus, for most analyses the Informant Sample ($N = 174$) was based on the 57 composite and 117 single other-ratings. The Informant Sample respondents consisted of 48% romantic relationships (e.g., husband, wife, fiancé, partner, etc.), 36% family (e.g., mother, sister, father, etc.) and 16% friends and co-workers. On average the 231 informants reported having known the participant for 18 years. The Informant Sample consisted of 83 online ratings and 148 mailed paper questionnaires.

Measures

Student and Retest Samples

The Student and Retest Samples completed the ISEE, the EES (Kring et al., 1994), the BEQ (Gross & John, 1995), the BFI (Benet-Martinez & John, 1998; John & Srivastava, 1999) and the PANAS-X (Watson & Clark, 1999).

ILPP and Informant Samples

The ILPP Sample and Informant Sample completed the ISEE, the BFI (Benet-Martinez & John, 1998; John & Srivastava, 1999) and the PANAS-X (Watson & Clark, 1999). The measures completed by the informants were adapted to an other-rating format. The ISEE other-rating version changed the statements from first person to third person (e.g., "I am emotionally expressive" became "X is emotionally expressive") with the directions to substitute the ILPP participant's name (i.e., the person who referred him or her) for the "X". The BFI other-rating version changed the statement "I see myself as someone who…" to "I see this person as someone who…" and then listed the same characteristics (e.g., "Is talkative"). The directions for the PANAS-X directed the respondent to rate "the person who referred you."

Results

Initial Analyses

Retest Sample

The Retest Sample consisted of 352 participants; however, missing data resulted in the removal of six participants' data. In the rest of the Retest Sample, 171 participants completed all items, 114 participants omitted only one or two items, 49 participants omitted between three to six items, and 12 participants omitted between 7 to 17 items. A total of 495 missing items—imputed from the average of the five most strongly correlated items in the rest of the sample—equaled 0.7% of the 70,930 total item responses.

The online retest assessment contained more missing data than the initial assessment; however, roughly a third of the Retest Sample only missed one or two items that likely resulted from "missed clicks." In the online assessment, participants clicked a response option, and *missed-clicks* resulted from clicking near the response without actually clicking on the response. Additionally, the questionnaire format consisted of about five or six questions on a page (i.e., computer screen). After completing the questions on each screen, participants advanced to the next set of questions by clicking on a *next-screen* button at the bottom on the screen. Double clicking on *next-screen* button likely resulted in a few cases with nearly complete data that omitted five or six questions in a row.

Based on the initial assessment scale mean scores, comparisons between the Retest Sample ($N = 346$) and the initial assessment only participants ($N = 35$) examined the representativeness of the Retest Sample (at the initial assessment one participant failed to complete the BFI and another failed to complete the PANAS-X). Limiting the comparisons to the 15 higher order scales of the ISEE (3), BEQ (4), EES (1) BFI (5), and PANAS-X (2) minimized the number of comparisons, and thus minimized Type I error. One significant difference existed; the Retest Sample reported higher levels of conscientiousness ($M = 32.7$, $SD = 5.9$) compared to initial only participants ($M = 29.4$, $SD = 5.8$; $t(378) = 3.2$, $p < .05$). This mean level difference on conscientiousness between participants completing versus not completing the retest assessment replicates previous longitudinal research (Watson & Humrichouse, 2006). Overall, the samples appeared to be similar and highly biased attrition therefore seemed unlikely.

<u>Informant Sample</u>

As noted earlier, the Informant Sample in ILPP consisted of 238 participants; however, three participants' data were removed due to missing data. In the rest of the sample, 188 participants had complete data and 47 participants had some missing data; 35 participants were missing only one or two items and 12 participants were missing between three to nine items. There was a far greater percentage of participants with complete data with the paper format (95%) versus the online format (52%); however, more than 70% of the participants with missing online data were missing just one or two items. Therefore, similar to the ILPP self-ratings, there is a greater potential to accidentally miss an online item or two by clicking *near* a response but failing actually to click on the response to select it. A total of 98 missing items—imputed from the average of the five most strongly correlated items in the rest of the sample—equaled 0.2% of the 40,322 total item responses. Finally, four participants' data were eliminated due to the fact that the matching self-rating had been removed due to missing data. As noted previously, the remaining 231 other-ratings were matched with 174 self-ratings and for most analyses the Informant Sample (N = 174) was based on the combined 57 composite and 117 single other-ratings.

Nomological Relations

Tables 33 and 34 report the correlations between emotional expressivity, affectivity and the Big Five across the different samples. First, the ISEE General scale correlated most strongly with E and revealed generally weak or nonsignificant relations with remaining BFI scales. This primary relation between E and emotional expressivity replicated the results of the Riggio and Riggio (2002) meta-analysis. Although the ISEE General scale contains only abstract items without any valenced content or discrete affective terms, it consistently demonstrated weak but significant correlations with PA.

The lower order levels of emotional expressivity revealed a strikingly different pattern than the general factor of emotional expressivity. PE and NE demonstrated clear differential relations with E and N, respectively. In all samples, PE and the all the positive discrete affect scales correlated significantly with E and weakly negatively or nonsignificantly with N. Conversely, in all samples, NE and all of the negative discrete affect scales correlated significantly with N and either weakly negatively or nonsignificantly with E. Within the positive and negative discrete affects the convergent relations with either E or N tended to be fairly consistent. Among the negative discrete affects, ISEE Fear tended to be the most strongly correlated with N whereas ISEE Guilt/Shame tended to be the least correlated with N. Among the positive discrete affects, ISEE Joviality consistently correlated most strongly with E.

In addition to the clear differential relations with E and N, the ISEE scales demonstrated a few other significant relations with the remaining Big Five. First, PE and the positive discrete affects of Joviality and Confidence consistently correlated with O; however, these correlations tended to be moderate in magnitude and less strong than the primary correlations with E. Second, relations with A demonstrated the clearest differentiation among the discrete affects. ISEE Hostility consistently negatively correlated with A and at a similar magnitude as the positive correlation between ISEE Hostility and N. Hostility demonstrated both similar convergent relations as the other negative discrete affects, yet also demonstrated differential specificity through convergent relations with A. ISEE Joviality consistently positively correlated with A across all samples; however, the magnitude of this association tended to be weaker than its convergent relations with E.

Similar to the convergent relations with E and N, PE and the positive discrete affects correlated with PA whereas NE and the negative discrete affects correlated with NA. The levels of convergence tended to be somewhat weaker with NA. Across both the Student and ILPP samples, ISEE NE correlated significantly less with NA than with N ($Z = 3.4$, $p < .01$; $Z = 5.4$, $p < .01$, respectively). Moreover, the convergent relations of ISEE NE and the negative discrete affects with NA tended to be weaker than the convergent relations of ISEE PE and the positive discrete affects. Interestingly, although NE and NA shared a substantial amount of affective content, a difference clearly existed between trait levels of experiencing negative affect and trait levels of expressing negative affect. More specifically, within the negative discrete affects scales, ISEE Sadness consistently correlated the least with NA. Conversely, ISEE Hostility tended to be the highest correlated with NA.

Considering the strong structural similarities and overlapping affective content of the ISEE and PANAS-X discrete affects, a variant of a MTMM matrix examined the convergence between affectivity and expressivity in relation to discriminant relations in the Student and Retest Samples (see Table 35) and the ILPP and Informant Samples (see Table 36). First, along the validity diagonal most of the corresponding ISEE and PANAS-X discrete affects correlated significantly, although Sadness failed to converge in three of the four samples. The convergence of the Joviality and Confidence/Self-Assurance scales tended to be stronger than the convergence of the negative discrete affect scales. Moreover, 82% (118 out of 144) and 85% (122 out of 144) of the convergent correlations exceeded the other discriminant correlations in the same row or column of the heteroscale blocks for Table 35 and Table 36, respectively and thus, provided some support for discriminant validity at the level of discrete affects. However, the convergent correlations for Sadness and Guilt/Shame provided only limited support for discriminant validity.

In addition to the cross-scale convergence of the discrete affects, a few other similarities existed between the ISEE and PANAS-X. In Table 35 the ISEE within-valence discrete affect correlations averaged .49, whereas the PANAS-X within-valence discrete affects correlations averaged .54. Similarly, in Table 36 the ISEE within-valence discrete affect correlations averaged .56 and the PANAS-X within-valence discrete affects correlations averaged .57. In other words, the discrete affects of positive and negative expressivity and the discrete affects of positive and negative affect hung together to roughly the same extent. In Table 35 the ISEE cross-valence discrete affect correlations averaged .24 and the PANAS-X cross-valence discrete affect correlations averaged -.22. In Table 36 the ISEE cross-valence discrete affect correlations averaged

.25 and the PANAS-X cross-valence discrete affect correlations averaged -.27. These cross-valence correlations generally fit within the three-level hierarchical structures for emotional expressivity and trait affect. The moderate positive correlation between ISEE positive and negative discrete affects fit with the nonspecific variance of the general factor of emotional expressivity. In parallel fashion, the moderate negative correlation between the PANAS-X positive and negative discrete affects fit with the nonspecific variance of the general bipolar factor of happiness versus unhappiness.

Next, regression analyses tested the incremental predictive validity of PE and NE. In both the upper and lower halves of Table 37, the first step of these regression analyses predicted E and N from PE and NE, respectively. The next step added PA as a predictor of E and NA was added as a predictor of N. Then, reversing the order of entry for the predictors allowed for the comparison of the incremental predictive validity of expressivity and affectivity. In the Student Sample (the left side of the table) comparing the R^2 change from Step 2 for the prediction of E revealed that PA contributed more unique variance than PE (.11 and .07, respectively). Similarly, comparing the R^2 change from Step 2 for the prediction of N revealed that NA contributes more unique variance than NE (.16 and .10, respectively). Moreover, this same pattern of results replicated in the ILPP Sample (the right side of the table), that is, in two large independent samples, PE and NE explained meaningful variance beyond PA and NA.

In Table 38 a similar set of regression analyses tested the incremental predictive validity of the ISEE discrete affects. Across all PANAS-X and ISEE discrete affects and across both samples, the PANAS-X discrete affects contributed more unique variance than the ISEE discrete affects with overall R^2 change averages of .16 and .05, respectively. Importantly, however, both the ISEE and PANAS-X discrete affects added significant incremental validity in predicting both N and E.

Table 33. Correlations Between the ISEE, PANAS-X and BFI in the Student and Retest Samples

Scales	NA	PA	N	E	O	A	C
Student Sample							
ISEE General	.03	.19	.13	**.30**	.06	.10	.10
Negative Expressivity	**.32**	-.09	**.47**	-.05	.03	-.08	-.04
Sadness	.18	-.03	**.35**	.04	.04	.04	.00
Hostility	**.29**	-.06	**.31**	.04	.03	**-.38**	-.02
Guilt/Shame	.19	.00	**.22**	-.02	.04	.04	-.04
Fear	**.25**	-.13	**.47**	-.17	-.03	.08	.00
Positive Expressivity	-.17	**.47**	-.16	**.48**	**.31**	.19	.13
Joviality	-.16	**.37**	-.07	**.44**	**.21**	**.22**	.08
Confidence	-.08	**.45**	-.18	**.33**	**.31**	.04	.13
Amusement	-.09	**.23**	-.06	**.39**	.16	.16	.01

Scales	NA	PA	N	E	O	A	C
Retest Sample							
ISEE General	-.06	**.27**	**.20**	**.38**	.13	.18	.16
Negative Expressivity	**.31**	.03	**.49**	-.01	-.02	-.10	-.03
Sadness	**.21**	.08	**.46**	.09	.02	.04	.04
Hostility	**.25**	-.01	**.33**	.08	-.01	**-.34**	-.06
Guilt/Shame	**.25**	.04	**.28**	-.03	-.01	-.02	-.07
Fear	**.24**	.01	**.50**	-.08	-.02	.00	-.01
Positive Expressivity	**-.30**	**.49**	-.11	**.52**	**.32**	**.32**	**.21**
Joviality	**-.29**	**.43**	-.02	**.52**	**.26**	**.34**	.18
Confidence	-.14	**.51**	-.18	**.39**	**.35**	.15	**.21**
Amusement	**-.29**	**.29**	-.01	**.39**	.19	**.30**	.10

N = 386 (Student Sample). N = 346 (Retest Sample).

PA = Positive Affect. NA = Negative Affect. E = Extraversion. N = Neuroticism.
O = Openness. A = Agreeableness. C = Conscientiousness.

Note: All correlations ≥ |.14| are significant at $p < .01$. Correlations ≥ |.20| in bold.

Table 34. Correlations Between the ISEE, PANAS-X and BFI in the ILPP and Informant Samples

Scales	NA	PA	N	E	O	A	C
ILPP Sample							
ISEE General	-.05	.17	.10	**.36**	.13	.05	.02
Negative Expressivity	**.22**	-.17	**.46**	-.11	-.09	-.12	-.15
Sadness	.10	-.01	**.31**	.07	.01	-.03	-.04
Hostility	**.22**	-.10	**.39**	-.02	-.11	**-.36**	-.15
Guilt/Shame	.17	**-.20**	**.29**	-.13	-.08	-.07	-.13
Fear	**.24**	**-.22**	**.44**	**-.23**	-.16	-.08	**-.21**
Positive Expressivity	-.15	**.31**	-.16	**.40**	**.28**	.16	.12
Joviality	-.16	**.32**	-.12	**.44**	**.29**	**.23**	.09
Confidence	-.05	**.30**	-.07	**.25**	**.27**	.06	.12
Amusement	-.10	.14	-.09	**.30**	.14	.10	.01

Scales	NA	PA	N	E	O	A	C
Informant Sample							
ISEE General	-.01	**.22**	.09	**.37**	.02	.12	.10
Negative Expressivity	**.36**	**-.24**	**.51**	-.15	**-.25**	**-.24**	-.16
Sadness	**.22**	-.08	**.44**	.01	-.11	-.11	-.08
Hostility	**.40**	**-.23**	**.54**	-.06	**-.26**	**-.47**	-.16
Guilt/Shame	**.28**	**-.28**	**.35**	**-.24**	**-.23**	**-.20**	**-.29**
Fear	**.33**	**-.21**	**.39**	-.18	**-.27**	-.06	-.11
Positive Expressivity	**-.20**	**.48**	-.11	**.54**	**.28**	**.21**	**.25**
Joviality	-.16	**.46**	-.10	**.49**	.18	**.26**	**.25**
Confidence	-.13	**.32**	-.07	**.46**	.17	.04	.15
Amusement	-.12	**.37**	-.02	**.35**	**.21**	.17	.16

N = 344 (ILPP Sample); correlations ≥ |.14| are significant at $p < .01$.

N = 174 (ILPP Informant); correlations ≥ |.20| are significant at $p < .01$.

PA = Positive Affect. NA = Negative Affect. E = Extraversion. N = Neuroticism. O = Openness. A = Agreeableness. C = Conscientiousness.

Note: Correlations ≥ .20 in bold.

Table 35. Intercorrelations of the ISEE and PANAS-X Discrete Affects in the Student and Retest Samples

Discrete Affects	ISEE 1	2	3	4	5	6	7	PANAS-X 8	9	10	11	12	13	14
ISEE														
1. Sadness	--	.57	.55	.65	.38	.23	.27	**.12**	.10	.04	.20	.16	-.05	.08
2. Hostility	.46	--	.44	.47	.29	.32	.21	.06	**.33**	.11	.17	.01	.03	-.01
3. Guilt/Shame	.41	.29	--	.60	.26	.33	.16	.08	.14	**.18**	.26	.08	-.07	.05
4. Fear	.49	.31	.51	--	.30	.22	.24	.09	.09	.08	**.30**	.08	-.18	.06
5. Joviality	.32	.18	.24	.18	--	.61	.71	-.26	-.33	-.29	-.20	**.60**	.27	.28
6. Confidence	.13	.26	.26	.09	.44	--	.37	-.21	-.11	-.18	-.10	.46	**.46**	.36
7. Amusement	.23	.17	.18	.21	.52	.27	--	-.19	-.31	-.24	-.24	.42	.14	.19
PANAS-X														
8. Sadness	**.07**	.07	.00	.09	-.20	-.20	-.16	--	.60	.68	.59	-.39	-.31	-.25
9. Hostility	.05	**.36**	-.01	.02	-.20	-.04	-.16	.46	--	.67	.57	-.35	-.07	-.22
10. Guilt	.00	.10	**.10**	.07	-.20	-.16	-.10	.63	.52	--	.61	-.35	-.21	-.27
11. Fear	.24	.16	.25	**.36**	-.09	-.06	-.05	.47	.41	.45	--	-.19	-.20	-.09
12. Joviality	.09	-.09	.08	-.02	**.53**	.31	.35	-.37	-.30	-.34	-.06	--	.63	.53
13. Self-Assurance	-.15	-.02	-.13	-.32	.23	**.47**	.16	-.31	-.06	-.23	-.19	.52	--	.45
14. Attentiveness	-.07	-.06	-.02	-.03	.17	.27	.06	-.19	-.08	-.20	-.05	.40	.36	--

N = 386 (Student Sample below the diagonal). N = 346 (Retest Sample above the diagonal).

Note: All correlations ≥ |.14| are significant at $p < .01$. Monoaffect-heteroscale convergent correlations are in bold.

Table 36. Intercorrelations of the ISEE and PANAS-X Discrete Affects in the ILPP and Informant Samples

	ISEE							PANAS-X						
Discrete Affects	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<u>ISEE</u>														
1. Sadness	--	.59	.59	.62	.39	.35	.26	**.20**	.17	.12	.20	-.04	-.15	-.08
2. Hostility	.61	--	.50	.46	.22	.28	.15	.32	**.49**	.23	.26	-.24	-.12	-.20
3. Guilt/Shame	.55	.45	--	.71	.13	.28	.14	.20	.17	**.25**	.27	-.22	-.28	-.30
4. Fear	.53	.45	.61	--	.18	.29	.16	.15	.13	.25	**.41**	-.13	-.25	-.20
5. Joviality	.47	.26	.20	.15	--	.61	.60	-.10	-.12	-.15	-.09	**.53**	.22	.32
6. Confidence	.39	.34	.23	.24	.63	--	.47	-.10	-.12	-.17	-.06	.32	**.32**	.21
7. Amusement	.29	.24	.14	.12	.56	.37	--	-.10	-.13	-.15	-.04	.34	.23	.30
<u>PANAS-X</u>														
8. Sadness	**.09**	.13	.19	.18	-.17	-.14	-.11	--	.46	.59	.56	-.28	-.11	-.17
9. Hostility	.05	**.29**	.05	.09	-.18	-.02	-.11	.58	--	.52	.50	-.39	-.08	-.23
10. Guilt	.06	.16	**.19**	.14	-.12	-.05	-.11	.66	.67	--	.62	-.28	-.15	-.17
11. Fear	.08	.15	.20	**.28**	-.16	-.06	-.08	.56	.55	.57	--	-.24	-.21	-.19
12. Joviality	.03	-.11	-.12	-.15	**.44**	.29	.29	-.49	-.41	-.39	-.34	--	.52	.54
13. Self-Assurance	-.12	-.09	-.27	-.33	.19	**.28**	.05	-.27	-.14	-.28	-.19	.61	--	.59
14. Attentiveness	.02	-.07	-.21	-.17	.20	.26	.10	-.41	-.29	-.37	-.33	.59	.53	--

<u>N</u> = 344 (ILPP Sample below the diagonal); correlations ≥ |.14| are significant at $p < .01$.

N = 174 (Informant Sample above the diagonal); correlations ≥ |.20| are significant at $p < .01$.

Note: Monoaffect-heteroscale convergent correlations are in bold.

Table 37. Hierarchical Regressions for the BFI, PANAS-X and ISEE in the Student and ILPP Samples

Criterion Neuroticism	Predictor	Student Sample R^2	R^2 Change	β	Beta	ILPP Sample R^2	R^2 Change	β	Beta
Step 1	Neg Exp	.22		.36*	.47*	.21		.34*	.46*
Step 2	Neg Exp	.39	.16	.26*	.33*	.50	.29	.24*	.33*
	NA			.47*	.43*			.73*	.56*
Step 1	NA	.29		.60*	.54*	.40		.83*	.63*
Step 2	NA	.39	.10	.47*	.43*	.50	.10	.73*	.56*
	Neg Exp			.26*	.33*			.24*	.33*

Criterion Extraversion	Predictor	Student Sample R^2	R^2 Change	β	Beta	ILPP Sample R^2	R^2 Change	β	Beta
Step 1	Pos Exp	.23		.52*	.48*	.16		.42 *	.40*
Step 2	Pos Exp	.35	.11	.33*	.30*	.27	.11	.31*	.29*
	PA			.46*	.38*			.41*	.34*
Step 1	PA	.28		.64*	.53*	.19		.52*	.43*
Step 2	PA	.35	.07	.46*	.38*	.27	.08	.41*	.34*
	Pos Exp			.33*	.30*			.31*	.29*

N = 385 (Student Sample). N = 344 (ILPP Sample).

* $p < .001$.

Pos Exp = Positive Expressivity. Neg Exp = Negative Expressivity. PA = Positive Affect. NA = Negative Affect.

Note: For Student Step 1, *df* Model = 1; *df* Error = 383; *df* Corrected Total = 384. For Student Step 2, *df* Model = 2; *df* Error = 382; *df* Corrected Total = 384. For ILPP Step 1, *df* Model = 1; *df* Error = 342; *df* Corrected Total = 343. For ILPP Step 2, *df* Model = 2; *df* Error = 341; *df* Corrected Total = 343.

Table 38. Hierarchical Regressions for the BFI and ISEE and PANAS-X Discrete Affects in the Student and ILPP Samples

		Student Sample				ILPP Sample			
Criterion Neuroticism	Predictor	R^2	R^2 Change	β	Beta	R^2	R^2 Change	β	Beta
Step 1	Sad NE	.12		.45*	.35*	.10		.40*	.31*
Step 2	Sad NE	.30	.18	.41*	.32*	.34	.24	.35*	.27*
	Sad NA			.67*	.42*			1.02*	.49*
Step 1	Sad NA	.20		.71*	.45*	.27		1.08*	.52*
Step 2	Sad NA	.30	.10	.67*	.42*	.34	.07	1.02*	.49*
	Sad NE			.41*	.32*			.35*	.27*

		Student Sample				ILPP Sample			
Criterion Neuroticism	Predictor	R^2	R^2 Change	β	Beta	R^2	R^2 Change	β	Beta
Step 1	Hos NE	.10		.42*	.31*	.15		.49*	.39*
Step 2	Hos NE	.15	.06	.30*	.22*	.32	.17	.33*	.26*
	Hos NA			.41*	.25*			.88*	.43*
Step 1	Hos NA	.11		.54*	.33*	.26		1.03*	.51*
Step 2	Hos NA	.15	.04	.41*	.25*	.32	.06	.88*	.43*
	Hos NE			.30*	.22*			.33*	.26*

		Student Sample				ILPP Sample			
Criterion Neuroticism	Predictor	R^2	R^2 Change	β	Beta	R^2	R^2 Change	β	Beta
Step 1	G/S NE	.05		.27*	.21*	.09		.39*	.29*
Step 2	G/S NE	.19	.14	.22*	.18*	.29	.20	.27*	.20*
	Guilt NA			.55*	.38*			.82*	.46*
Step 1	Guilt NA	.16		.58*	.40*	.25		.89*	.50*
Step 2	Guilt NA	.19	.03	.55*	.38*	.29	.04	.82*	.46*
	G/S NE			.22*	.18*			.27*	.20*

Table 38. Continued

Criterion Neuroticism	Predictor	Student Sample R^2	R^2 Change	β	Beta	ILPP Sample R^2	R^2 Change	β	Beta
Step 1	Fear NE	.22		.59*	.47*	.19		.57*	.44*
Step 2	Fear NE	.33	.11	.43*	.34*	.36	.17	.41*	.32*
	Fear NA			.56*	.35*			.58*	.43*
Step 1	Fear NA	.23		.76*	.48*	.27		1.18*	.52*
Step 2	Fear NA	.33	.10	.56*	.35*	.36	.09	.98*	.43*
	Fear NE			.43*	.34*			.41*	.32*

Criterion Extraversion	Predictor	Student Sample R^2	R^2 Change	β	Beta	ILPP Sample R^2	R^2 Change	β	Beta
Step 1	Jov PE	.20		.73*	.45*	.19		.70*	.44*
Step 2	Jov PE	.38	.18	.29*	.18*	.31	.12	.42*	.39*
	Jov PA			.63*	.50*			.55*	.26*
Step 1	Jov PA	.36		.75*	.60*	.26		.71*	.51*
Step 2	Jov PA	.38	.02	.63*	.50*	.31	.06	.55*	.39*
	Jov PE			.29*	.18*			.42*	.26*

Criterion Extraversion	Predictor	Student Sample R^2	R^2 Change	β	Beta	ILPP Sample R^2	R^2 Change	β	Beta
Step 1	Con PE	.11		.48*	.33*	.06		.36*	.25*
Step 2	Con PE	.28	.17	.16*	.11*	.22	.16	.19*	.13*
	SA PA			.72*	.47*			.76*	.42*
Step 1	SA PA	.27		.80*	.52*	.20		.82*	.46*
Step 2	SA PA	.28	.01	.72*	.47*	.22	.02	.76*	.42*
	Con PE			.16**	.11**			.19**	.13**

N = 385 (Student Sample). N = 344 (ILPP Sample).

* $p < .001$. ** $p < .05$.

Table 38. Continued

Sad NA = Sadness Negative Affect. Sad NE = Sadness Negative Expressivity. Hos NA = Hostility Negative Affect. Hos NE = Hostility Negative Expressivity. Guilt NA = Guilt Negative Affect. G/S NE = Guilt/Shame Negative Expressivity. Fear NA = Fear Negative Affect. Fear NE = Fear Negative Expressivity. Jov PA = Joviality Positive Affect. Jov PE = Joviality Positive Expressivity. SA PA = Self-assurance Positive Affect. Con PE = Confidence Positive Expressivity.

Note: For Student Step 1, *df* Model = 1; *df* Error = 383; *df* Corrected Total = 384. For Student Step 2, *df* Model = 2; *df* Error = 382; *df* Corrected Total = 384. For ILPP Step 1, *df* Model = 1; *df* Error = 342; *df* Corrected Total = 343. For ILPP Step 2, *df* Model = 2; *df* Error = 341; *df* Corrected Total = 343.

CHAPTER 7. CONVERGENT AND DISCRIMINANT VALIDITY

To support construct validity, the process of validating a new scale needs to provide evidence of convergent and discriminant validity (Campbell & Fiske, 1959; Urbina, 2004). To establish support for convergent and discriminant validity, one needs to demonstrate that only the intended construct was assessed and that the construct is distinct from other constructs. The designs of two primary studies specifically incorporated methods to examine convergent and discriminant validity. First, in Study 1 (i.e., Student Sample) participants completed the ISEE, the EES (Kring et al., 1994) and the BEQ (Gross & John, 1995) to examine the convergent validity of scales of emotional expressivity and to examine discriminant validity among lower order factors of emotional expressivity. Second, in the Study 2 (i.e., ILPP Sample) self- and other-ratings served as different methods to examine convergent and discriminant validity within MTMM matrices (Campbell & Fiske, 1959). Demonstrating strong correlations between two measures or two methods purported to assess the same construct supports convergent validity; whereas, demonstrating weak correlations between measures or methods assessing conceptually distinct constructs supports discriminant validity.

Rather than an absolute or all-or-nothing decision, the interpretation of convergent and discriminant correlations is a matter of degree and involves a relative comparison of convergent correlations to discriminant correlations. Although strong convergent correlations support convergent validity, convergent correlations lack established benchmarks for evaluation. If convergent correlations between a new measure and published measures exceed .80, then there exists little need for the new measure; that is, the two measures yield nearly identical information. Generally, convergent correlations need to be significant and exceed discriminant correlations to support convergent validity. Additionally, the distinctiveness of the methods used to obtain convergent correlations (e.g., self-report vs. behavioral observation) influences the magnitude of convergent correlations, that is, more similar methods generally yield stronger convergent correlations.

Within hierarchical models, the interpretation of convergent and discriminant correlations also needs to take into account the interrelations that are represented within these multilevel schemes. More specifically, discriminant correlations between constructs at the same level of a hierarchy need to be greater than zero to support existence of the hierarchical structure. Put differently, significant discriminant correlations reveal the common variance among constructs within a particular level of a hierarchy. Convergent correlations that exceed these significant discriminant correlations reveal the specific variance of a particular construct.

To examine convergent validity, participants completed the ISEE, the EES (Kring et al., 1994) and the BEQ (Gross & John, 1995). Modeled after the EES, the ISEE General scale also contains items assessing general emotional expressivity without explicitly tapping valence content or nonverbal channel specifications. This similarity led to the expectation of strong convergence between the ISEE General scale and the EES (i.e., $r > .70$). The differences in level of abstraction of items and item content between the ISEE General scale and the BEQ led to the expectation of slightly lower levels of convergence (i.e., $r > .50$) between these measures. Loosely modeled after the BEQ Positive and Negative Expressivity scales, the ISEE PE and NE contained a broader range of discrete affects, which led to the expectation of moderate to strong convergence between the corresponding scales (i.e., rs = .30 - .50). Finally, the lack of published emotional expressivity scales at the discrete affect level resulted in examining convergent and discriminant validity of the discrete affect scales in relation to the higher order scales of positive and negative expressivity.

Self-Other Agreement

Self-other agreement—that is, the convergence between self- and other-ratings—served as a test of convergent validity across different methods. Interpreting self-other agreement correlations for emotional expressivity in relation to previous research on self-other agreement for the Big Five and trait affect (Watson & Clark, 1991; Watson et al., 2000) provided established benchmarks for comparison. Comparisons of self-other agreement across emotional expressivity, the Big Five and trait affect needed to take into account the *trait visibility effect*: traits with relevant, observable behaviors yield higher agreement correlations than more internal, subjective traits (Funder & Colvin, 1988; Ready, Clark, Watson, & Westerhouse, 2000; Watson & Clark, 1991; Watson et al., 2000).

Watson et al. (2000) demonstrated the trait visibility effect by comparing the overall mean agreement correlations for the Big Five versus trait affect for married couples (r = .56 vs. .41), dating couples (r = .47 vs. .27), and friendship dyads (r = .41 vs. .26). Trait affect scales consistently showed lower levels of self-other agreement due to the internal, subjective nature of these traits. Therefore, this trait visibility effect led to the expectation of higher levels of agreement for emotional expressivity (i.e., the tendency to express affective states outwardly through nonverbal behaviors) relative to trait affect (i.e., what one is subjectively experiencing or feeling).

Previous research on self-other agreement also supported the *acquaintanceship effect*—that is, agreement across raters (including both self-peer and interjudge agreement) increases with increased levels of acquaintance (Funder & Colvin, 1988; Norman & Goldberg, 1966; Paulhus & Bruce, 1992; Watson et al., 2000). For example, Watson et al. (2000) reported that overall mean self-other agreement correlations in married couples (r = .46) were significantly larger than in either friendship dyads or dating couples (r = .30 and r = .33, respectively).

Kring et al. (1994) reported significant self-other agreement on the EES (r = .49) based on a sample of 37 self-ratings from college students and other-ratings collected from mothers. It should be noted, however, that the design of this self-other agreement study likely influenced the level of agreement. Most notably, based on their "scoring in the top and bottom quartiles of the EES distribution" (Kring et al., 1994, p. 945), researchers selected participants for inclusion in additional studies (e.g., obtaining other-ratings from their mothers) and this selection process likely inflated self-other agreement.

Gross and John (1997) reported significant self-other agreement for the BEQ Total scale (r = .58) and for the individual BEQ facets (rs = .48, .43, .41 for Impulse Strength, Positive Expressivity, and Negative Expressivity, respectively) based on a sample of 44 self-ratings from college students and composite of up to three peer-ratings (primarily friends and roommates). Several aspects of this design likely influenced the level of agreement. First, the friends and roommates providing peer-ratings knew the targets for about 3 years—comparatively higher acquaintanceship than strangers but lower than long-term married couples. Second, creating composite peer-ratings increased the reliability of the composite rating and subsequently resulted in increased self-other agreement (Kenny, 1994; Watson & Clark, 1991). Finally, it is interesting to note that summing the three facets of the BEQ into the BEQ Total scale resulted in higher levels of self-other agreement for the latter.

Based on 482 self-ratings and a composite of two peer-ratings for each target, Trierweiler et al. (2002) reported estimates of self-other agreement correlations for discrete affects in the .20 to .40 range. Interestingly, the positive discrete affects revealed stronger agreement correlations than the negative discrete affects and these researchers interpreted this finding in relation to the trait visibility effect, given that participants reported higher mean levels of expressing positive discrete affects. More broadly, the differing levels of agreement for discrete affects supported a discrete affects model of emotional expressivity (Trierweiler et al., 2002).

Related to comparisons of agreement across the three-level hierarchical model of emotional expressivity, previous researchers have theorized that less abstract constructs should yield higher levels of agreement.

> Broad traits should give individuals more leeway in interpreting those behaviors that are relevant to the trait. Thus personality ratings on broad traits made by different individuals are more likely to be based on somewhat different subsets of past behaviors, which leads to lower interjudge and self-other agreement for broad than for narrow traits (Hampson, John, & Goldberg, 1986, p. 51).

> [S]ince it is easier to obtain consensual agreement about constructs that are narrow, simple, and less abstract, these are also more easily assessed than broader and multifaceted constructs that may have acquired different meanings across diverse contexts, cultures, and historical periods (Urbina, 2004, p. 156).

Multi-Trait Multi-Method Matrix

The concepts of convergent and discriminant validity originated within the context of the MTMM matrix (Campbell & Fiske, 1959). With at least two or more traits assessed by at least two or more methods, the correlations between the same traits assessed by different methods (i.e., convergent correlations) need to exceed the correlations between different traits assessed by the same or different methods (i.e., discriminant correlations). In Study 2, self-other agreement correlations formed the validity diagonal of the heterotrait-heteromethod block. Significant agreement correlations supported convergent validity. Agreement correlations that exceeded other correlations in their row or column of the heterotrait-heteromethod block provided initial support for discriminant validity. A stronger test of discriminant validity necessitated that agreement correlations exceeded other correlations in their row or column of the monomethod triangles.

The ISEE consist of two or more traits at both lower levels of the hierarchy and, therefore, generated two different MTMM matrices. A MTMM matrix at the level of PE and NE and a MTMM matrix at the level of the discrete affects both needed to demonstrate evidence of convergent and discriminant validity. The MTMM matrix at the level of PE and NE addressed the important issue of discriminant validity between these scales/constructs. Similarly, each discrete affect scale needed to demonstrate evidence of meaningful specific variance, that is, significant self-other agreement correlations that exceeded the discriminant coefficients. Additionally, each discrete affect scale also needed to demonstrate evidence of common variance with the other discrete affects, that is, significant discriminant correlations.

The expected level of interrelations within the positive discrete affects and within the negative discrete affects in the monomethod triangles made passing the more stringent test of discriminant validity unlikely (i.e., it is unlikely that agreement correlations would exceed other correlations in monomethod triangles). In other words, the correlation between one's self-rating of ISEE Fear and Sadness was expected to be higher than the self-other agreement correlation for Fear. An intermediate test of discriminant validity consisted of self-other agreement correlations exceeding the cross-valence correlations within the monomethod triangles, that is, self-other agreement for Fear (e.g., $r = .40$) needed to exceed the correlation between self-rated Fear and Joviality (e.g., $r = .25$).

Method

Participants and Measures

Student Sample

Chapter 5 provided a description of the Student Sample ($N = 387$). The Student Sample completed the ISEE, the EES (Kring et al., 1994), the BEQ (Gross & John, 1995), the BFI (Benet-Martinez & John, 1998; John & Srivastava, 1999) and the PANAS-X (Watson & Clark, 1999).

ILPP and Informant Sample

Chapter 5 also provided a description of the ILPP Sample ($N = 344$) and Chapter 6 provided a description of the Informant Sample ($N = 174$). The ILPP Sample and Informant Sample completed the ISEE, the BFI (Benet-Martinez & John, 1998; John & Srivastava, 1999) and the PANAS-X (Watson & Clark, 1999). The measures completed by the informants were adapted to an other-rating format.

Results

Convergent Measures

As shown in Table 39, the ISEE, EES and BEQ all correlated significantly. The strong convergence between these scales supported convergent validity and indicated all scales assessed various aspects of emotional expressivity. At the general factor level, the ISEE General scale demonstrated strong convergence with both the EES and BEQ, (*r*s = .74 and .71, respectively). Tests for significant differences in the comparison of correlated correlation coefficients followed the method outlined by Meng, Rosenthal, and Rubin, (1992). The correlation between the ISEE General scale and the EES ($r = .74$) exceeded the correlations between the ISEE General scale and each of the BEQ facets of Impulse Strength, Negative Expressivity, and Positive Expressivity (*r*s = .54, .64, .64, respectively; *Z*s > 3.6, $p < .01$). Similarly, the correlations between the ISEE General scale and the EES ($r = .74$) exceeded the correlations between EES and any other ISEE scales (i.e., NE and PE and all discrete affect scales; *Z*s > 5.2, $p < .01$). In a similar set of comparisons, the correlations between the ISEE General scale and BEQ Total scale ($r = .71$) exceeded the correlations between the BEQ Total and all other ISEE scales (*Z*s > 2.6, $p > .01$), except for ISEE Sadness ($r = .69$; $Z = 0.6$, *ns*).

At the second-order level, the ISEE and BEQ demonstrated the convergence of negative and positive expressivity. ISEE NE correlated significantly more strongly with BEQ Negative Expressivity than with BEQ Positive Expressivity (*r*s = .57 vs. .33; $Z = 5.6$, $p < .01$). Conversely, ISEE PE correlated more strongly with BEQ Positive Expressivity than with BEQ Negative Expressivity (*r*s = .67 vs. .30; $Z = 8.9$, $p < .01$). Similarly, the ISEE discrete affects of Sadness, Hostility, Fear, Joviality, Confidence, and Amusement correlated more strongly with the corresponding BEQ subscale versus the oppositely valenced BEQ subscale (*Z*s > 3.9, $p < .01$). Only the Guilt/Shame discrete affect scale failed to correlate more strongly with BEQ Negative Expressivity compared to BEQ Positive Expressivity (*r*s .29 vs. .23; $Z = 1.4$, *ns*). The correlations of the ISEE discrete affects with the corresponding BEQ subscale averaged .49 versus .25 with the oppositely valenced BEQ subscale. These relations supported both convergent and discriminant validity for the ISEE and BEQ and again demonstrated that these scales contain both nonspecific variance related to a general factor and specific variance related to either negative or positive expressivity.

Interestingly, the ISEE demonstrated a nearly identical pattern of convergence with BEQ Impulse Strength as with BEQ Negative Expressivity. The ISEE Negative and Positive Expressivity scales, as well as the discrete affects of Sadness, Guilt/Shame, Fear, Joviality, Confidence, and Amusement, all failed to discriminate between correlations with the BEQ Impulse Strength versus the BEQ Negative Expressivity subscale (Zs < |0.8|, *ns*). Only ISEE Hostility had a significantly lower correlation with BEQ Impulse Strength than with BEQ Negative Expressivity (rs .29 vs. .40; $Z = 2.5$, $p < .05$).

In the lower half of Table 39, partialing out both the EES and ISEE General scale from the convergent correlations removed the nonspecific variance related to the general factor of emotional expressivity while leaving the variance related to the lower order factors. The correlations between the ISEE and BEQ Negative and Positive Expressivity revealed a clear pattern of differential convergence; correlations between similarly valenced scales ranged from .16 to .54, whereas correlations between cross-valenced scales ranged from only -.13 to .05. Removal of the nonspecific variance of the general factor of emotional expressivity left little overlap across positive and negative expressivity scales. Again the convergent correlations for BEQ Impulse Strength revealed a similar pattern as seen with the BEQ Negative Expressivity convergent correlations.

Furthermore, removal of the nonspecific variance of the general factor of emotional expressivity revealed that the BEQ Total scale still correlated with ISEE NE and ISEE Sadness and Fear. As previously noted, the BEQ Total scale also failed to discriminate between the ISEE General scale and ISEE Sadness (rs = .71 vs. .69; $Z = 0.6$, *ns*). This lack of discrimination, together with the moderate to strong partial correlations, reveals problems in the balance of positive and negative expressivity in the BEQ Total scale; that is, the BEQ Total scale appears to be tipped towards negative emotional expressivity.

Self-Other Agreement and MTMM Matrices

In Table 40, the self-other agreement correlations for the Big Five and trait affect scales served as benchmarks for comparing the emotional expressivity self-other agreement correlations. First, the BFI self-other agreement correlations ranged from .44 to .65. These agreement correlations compared similarly to previous research reporting BFI self-other agreement (Watson & Humrichouse, 2006; Watson et al., 2000). Second, although still significant, the ISEE and PANAS-X self-other agreement correlations ranged from .29 to .42 and tended to be substantially lower than the BFI. The ISEE self-other agreement correlations exceeded the PANAS-X self-other agreement correlations however, this difference was not statistically significant. Overall, the moderate to strong self-other agreement correlations provided general support for convergent validity across methods, that is, two raters agreed about one's personality, affectivity and emotional expressivity.

Table 40 reports several additional tests for differences in self-other agreement within and across personality, trait affect and emotional expressivity in the current study. Tests for significant differences in agreement utilized the Pearson-Filon test for two correlations from the same sample (Kenny, 1987). Within the BFI, agreement for E (r = .65) exceeded all the other Big Five traits and demonstrated significantly stronger agreement than the corresponding correlations for A and O (rs = .45, .44, respectively; Zs > 2.7, ps < .01; all other comparisons within the BFI scales were nonsignificant; Zs < 1.95, *ns*). E (r = .65) also demonstrated significantly stronger agreement than both PE (r = .37) and PA (r = .32; Zs > 3.7, ps < .01). Similarly, N (r = .56) demonstrated significantly stronger agreement than NE (r = .37) and NA (r = .29; Zs > 2.3, ps < .05). Agreement on the ISEE General scale (r = .42) approached the low-end of the BFI range and only E (*r* = .65; *Z* = 3.08, $p < .01$) demonstrated significantly stronger agreement. Although the ISEE General scale agreement correlation exceeded those for ISEE PE and NE, these differences failed to reach significance (*Z*s < 0.7, *ns*).

Accounting for the trait visibility effect led to the expectation of higher levels of agreement for expressivity compared to trait affect. However, although they were in the expected direction, both comparisons of agreement between PE (*r* = .37) and PA (*r* = .32; *Z* = 0.57, *ns*) and between NE (*r* = .37) and NA (*r* = .29; *Z* = 0.80, *ns*) failed to reach significance. Moreover, as previously reported, both E and N demonstrated significantly stronger agreement than PE and NE, respectively.

The MTMM matrices in Table 41 addressed convergent and discriminant validity of the BFI, ISEE PE and NE and the PANAS-X PA and NA scales. Along the validity diagonals, the significant agreement correlations (in bold) supported convergent validity across methods. The BFI agreement correlations significantly exceeded the discriminant correlations within the same row or column of the hetero-method block (*Z*s > 3.0, $p < .05$). All but one of the BFI agreement correlations exceeded the discriminant correlations in the mono-method triangles—the correlation between other-rated Agreeableness and other-rated Neuroticism (*r* = -.46) was stronger in magnitude than agreement on Agreeableness (*r* = 45)—however, only 14 out of the 20 comparisons reached statistical significance (*Z*s > 1.96, $p < .05$). Half of the comparisons of agreement correlations with the correlations in the other-ratings mono-method triangle failed to reach significance. Related to this, the discriminant correlations in the other-ratings mono-method triangle (mean *r* = .31) tended to be stronger than discriminant correlations in the self-ratings mono-method triangle (mean *r* = .23).

The ISEE PE and NE agreement correlations significantly exceeded the discriminant correlations within the same row or column of the hetero-method block (*Z*s > 2.2, $p < .05$). Although the ISEE PE and NE agreement correlations also exceeded the discriminant correlations in the mono-method triangles, these comparisons failed to reach significance (*Z*s < 1.2, *ns*). The significant correlations between PE and NE fit with the hierarchical model of emotional expressivity; however, PE and NE agreement correlations still exceeded the correlations between PE and NE, thereby establishing discriminant validity. Ignoring the signs of the correlations in the PANAS-X MTMM, only the PA agreement correlation ($r = .32$) significantly exceeded the discriminant correlation between self-rated PA and other-rated NA ($r = |.13|$; $Z = 2.1$, $p < .05$). Although the PA and NA agreement correlations exceeded the other discriminant correlations in the hetero-method block and mono-method triangles, these comparisons failed to reach significance (*Z*s < 1.9, *ns*). Overall, the agreement correlations exceeded any other correlation in its row or column and therefore pass both discriminant validity tests.

Additionally, in Table 42, the discrete affect agreement correlations for the ISEE and PANAS-X averaged to .34 and .33, respectively; however, the median agreement correlations for the ISEE and PANAS-X were .36 and .29, respectively. All differences in agreement correlations between corresponding discrete affects from the ISEE and PANAS-X were non-significant (*Z*s < |1.4|, *ns*). Again, the comparisons between emotional expressivity and trait affect failed to fit with the trait visibility effect. Overall, however, these significant self-other agreement correlations provide general support for convergent validity at the discrete affect level.

Similarly, the MTMM matrix in Table 43 addressed the convergent and discriminant validity of the ISEE discrete affect scales. Along the validity diagonal, the significant agreement correlations (in bold) supported convergent validity across methods. In 95% (80 of the 84) of the comparisons the ISEE discrete affect agreement correlations exceeded the discriminant correlations within the same row or column of the heteromethod block and thus demonstrated discriminant validity. Additionally, 62% (52 of the 84) of the convergent correlations were significantly higher than the discriminant correlations (*Z*s > 1.96, $p < .05$); for each discrete affect scale the number of significant comparisons out of 12 included: Sadness (8), Hostility (10), Guilt/Shame (3), Fear (8), Joviality (6), Confidence (6), and Amusement (11). Most notably, agreement on the Guilt/Shame scale ($r = .23$) failed to demonstrate discriminant validity in relation to the Fear scale, as there were stronger correlations between self-rated Guilt/Shame and other-rated Fear ($r = .24$) and between other-rated Guilt/Shame and self-rated Fear ($r = .29$).

As an intermediate test of discriminant validity, in 88% (21 of the 24) of the comparisons, the discrete affects agreement correlations exceeded the cross-valence correlations within the monomethod triangles; however, only 25% (6 of the 24) of the comparisons reached significance (Zs > 1.96, $p < .05$). The monomethod correlations between the Sadness and Joviality scales ($r = .46$ and $r = .39$ for self- and other-ratings, respectively) exceeded the agreement correlations for Sadness ($r = .36$) and Joviality ($r = .31$). More generally, Sadness demonstrated substantial relations with all other discrete affects. As expected, none of discrete affect agreement correlations exceeded the within-valence correlations within the monomethod triangles.

Similarly, the MTMM matrix in Table 44 addressed the convergent and discriminant validity of the PANAS-X discrete affect scales. Along the validity diagonal, the significant agreement correlations (in bold) supported convergent validity across methods. All comparisons of the PANAS-X discrete affect agreement correlations exceeded the discriminant correlations within the same row or column of the heteromethod block and thus demonstrated discriminant validity. Additionally, 65% (55 of the 84) of the convergent correlations were significantly higher than the discriminant correlations (Zs > 1.96, $p < .05$); for each discrete affect scale the number of significant comparisons out of 12 included: Sadness (12), Hostility (10), Guilt (5), Fear (7), Joviality (11), Confidence (8), and Attentiveness (2).

As an intermediate test of discriminant validity, in 79% (19 of the 24) of the comparisons, the discrete affects agreement correlations exceeded the cross-valence correlations within the monomethod triangles (ignoring the sign of the correlation);

however, only 8% (2 of the 24) of the comparisons reached significance (*Z*s > 1.96, $p <$ .05). Within both the self- and other-rating monomethod triangles, Joviality demonstrated substantial relations with all other discrete affects. As expected, none of discrete affect agreement correlations exceeded the within-valence correlations within the monomethod triangles.

Overall, the ISEE and PANAS-X MTMM matrices for self- and other-ratings yielded comparable levels of support for convergent and discriminant validity. The discrete affect agreement correlations averaged .34 and .33 for the ISEE and the PANAS-X, respectively. Additionally, within the hetero-trait hetero-method blocks, the ISEE and PANAS-X discrete affects demonstrated similar percentages of significant comparisons with agreement correlations exceeding discriminant correlations (62% vs. 65%, respectively); however, it should be noted that two ISEE discrete affect agreement correlations failed to exceed four other discriminant correlations within the hetero-trait hetero-method block. The Joviality scales of the ISEE and PANAS-X revealed the largest difference in the number of significant comparisons (6 vs. 11, respectively). The ISEE Guilt/Shame and PANAS-X Guilt and Attentiveness scales revealed the least number of significant comparisons. Within the mono-method triangles, the ISEE and PANAS-X discrete affects also demonstrated similar percentages for agreement correlations exceeding cross-valence discriminant correlations (88% vs. 79%, respectively).

Tests of potential moderators of self-other agreement utilized moderated multiple regression with (a) the self-ratings as the criteria predicted from other-ratings, and (b) the other-ratings as the criteria predicted from self-ratings. The potential moderating variable, along with the corresponding self- or other-rating, was entered into the equation in the first step; the second step then added the centered interaction term. For example, to test if other-rated PE moderated agreement on PA: (a) self-rated PA served as the criterion, (b) other-rated PA and other-rated PE were entered as predictors in step one, (c) then the centered interaction term—the product of other-rated PA and PE—was entered in step two. If the interaction term added significant incremental predictive validity beyond the two main effects, then PE moderated agreement of PA. Tests of moderators included variables contained in self- and other-ratings (i.e., Big Five, trait affect, and emotional expressivity).

Tests of specific combinations of variables limited the overall number of regressions. Moderated multiple regressions tested (a) each of PE, PA and E as moderators of the other two; (b) each of NE, NA and N as moderators of the other two; and (c) ISEE General and E as moderators of each other. At the discrete affect level, tests for moderators included the six corresponding ISEE and PANAS-X discrete affect scales as moderators for each other. Only 2 of the 52 moderated multiple regressions (i.e., 26 self-rated and 26 other-rated) revealed a significant interaction term. Other-rated NE moderated agreement on NA (self-rated NA = other-rated NA + other-rated NE + (other-rated NA X other-rated NE)); the interaction term added an additional 2% of the variance, but the regression weight was negative. Additionally, other-rated PANAS-X Guilt moderated agreement of ISEE Guilt/Shame expressivity (self-rated ISEE Guilt = other-rated ISEE Guilt + other-rated PANAS-X Guilt + (other-rated ISEE Guilt X other-rated PANAS-X Guilt)); again, the centered interaction term added an additional 2% of the variance. In the first case emotional expressivity moderated agreement on trait affect and in the second trait affect moderated agreement on emotional expressivity; however, 2 significant moderators out of 52 tests is less than would be expected due to chance alone. Overall, consistent with the broader literature in this area (see Watson et al., 2000), there was little support for moderators of self-other agreement through moderated multiple regressions.

Table 39. Convergent Correlations Between the ISEE, EES and BEQ in the Student Sample

ISEE	EES	BEQ	Imp St	Neg Exp	Pos Exp
General	**.74**	**.71**	.54	.64	.64
Negative Expressivity	.43	.62	.58	**.57**	.33
Sadness	.58	.69	.64	**.63**	.38
Hostility	.27	.36	.29	**.40**	.17
Guilt/Shame	.23	.34	.31	**.29**	.23
Fear	.27	.44	.42	**.40**	.21
Positive Expressivity	.42	.44	.29	.30	**.67**
Joviality	.45	.53	.39	.36	**.69**
Confidence	.23	.21	.13	.13	**.37**
Amusement	.36	.39	.27	.26	**.56**

EES & ISEE General Partialed	BEQ	Imp St	Neg Exp	Pos Exp
Negative Expressivity	.47	.42	**.38**	-.03
Sadness	.41	.42	**.32**	-.13
Hostility	.18	.12	**.27**	-.10
Guilt/Shame	.23	.19	**.16**	.05
Fear	.37	.32	**.30**	.01
Positive Expressivity	.12	.00	-.09	**.54**
Joviality	.24	.12	-.03	**.54**
Confidence	-.01	-.05	-.11	**.27**
Amusement	.14	.03	-.05	**.43**

$\underline{N}$ = 387.

Note: All correlations ≥ |.14| are significant at $p < .01$. Convergent correlations are in bold.

EES = Emotional Expressivity Scale. BEQ = Berkeley Expressivity Questionnaire.
Imp St = Impulse Strength. Neg Exp = Negative Expressivity. Pos Exp = Positive Expressivity.

Table 40. Self-Other Agreement Correlations and Differential Agreement *Z*-Scores Differences for the BFI, ISEE and PANAS in the ILPP Sample

Scale	Self-Other Agreement	Differential Agreement *Z*-Score Differences vs. BFI					vs. ISEE
BFI		vs. N	vs. C	vs. A	vs. O		
Extraversion	.65	1.39	1.94	**2.73***	**2.86***		
			vs. C	vs. A	vs. O		
Neuroticism	.56		0.58	1.40	1.48		
				vs. A	vs. O		
Conscientiousness	.51			0.81	0.91		
					vs. O		
Agreeableness	.45				0.11		
Openness	.44						
ISEE		vs. E	vs. N	vs. C	vs. A	vs. O	
General	.42	**-3.08***	-1.67	-1.09	-0.31	-0.20	
		vs. E					vs. ISEE Gen
Positive Expressivity	.37	**-3.74***					-0.65
		vs. N					vs. ISEE Gen
Negative Expressivity	.37	**-2.35****					-0.63
PANAS-X		vs. E					vs. ISEE PE
Positive Affect	.32	**-4.31***					-0.57
		vs. N					vs. ISEE NE
Negative Affect	.29	**-3.33***					-0.80

$N = 174$.

Note: All correlations significant $p < .01$.

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

Note: *Z*s for all comparisons are based on the trait in the left column versus the trait in the subheading (i.e., vs. E). A significant positive or negative *Z* indicates that the trait in the left column or the trait in the subheading is more dependable, respectively.

E = Extraversion. N = Neuroticism. C = Conscientiousness. A = Agreeableness. O = Openness. ISEE Gen = General scale. ISEE NE = Negative Expressivity. ISEE PE = Positive Expressivity.

Table 41. MTMM for Self-Other-Ratings of BFI, ISEE and PANAS-X in ILPP Sample

BFI	Self-Rating					Other-Ratings				
Self-ratings	1	2	3	4	5	6	7	8	9	10
1. Extraversion										
2. Neuroticism	-.22									
3. Conscientiousness	.14	-.35								
4. Agreeableness	.18	-.38	.35							
5. Openness	.32	-.16	.13	.08						
Other-ratings										
6. Extraversion	**.65**	-.25	.23	.15	.18					
7. Neuroticism	-.20	**.56**	-.11	-.15	-.06	-.41				
8. Conscientiousness	.07	-.11	**.51**	.15	-.08	.33	-.26			
9. Agreeableness	.13	-.16	.08	**.45**	-.03	.26	-.46	.37		
10. Openness	.12	-.09	.04	-.05	**.44**	.30	-.24	.30	.21	

ISEE	Self-Rating		Other-Ratings	
Self-ratings	1	2	3	4
1. Positive Expressivity				
2. Negative Expressivity	.26			
Other-ratings				
3. Positive Expressivity	**.37**	.12		
4. Negative Expressivity	.17	**.37**	.27	

PANAS-X	Self-Rating		Other-Ratings	
Self-ratings	1	2	3	4
1. Positive Affect				
2. Negative Affect	-.30			
Other-ratings				
3. Positive Affect	**.32**	-.19		
4. Negative Affect	-.13	**.29**	-.28	

N = 174.

Note: All correlations ≥ |.20| are significant at $p < .01$. Convergent correlations are in bold.

Table 42. Self-Other Agreement Correlations and Differential Agreement *Z*-Scores Differences for the ISEE and PANAS-X Discrete Affects in the ILPP Sample

Scale	Self-Other Agreement	Differential Agreement *Z*-Score Differences						
		Within ISEE						**PANAS-X**
ISEE		vs. H	vs. GS	vs. F	vs. J	vs. C	vs. A	vs. Sadness
Sadness	.37	0.05	1.56	0.11	0.64	0.44	-0.62	-0.50
			vs. GS	vs. F	vs. J	vs. C	vs. A	vs. Hostility
Hostility	.37		1.47	0.06	0.57	0.39	-0.66	-0.08
				vs. F	vs. J	vs. C	vs. A	vs. Guilt
Guilt/Shame	.23			-1.52	-0.80	-0.98	**-2.01****	-0.55
					vs. J	vs. C	vs. A	vs. Fear
Fear	.36				0.51	0.33	-0.71	.81
						vs. C	vs. A	vs. Joviality
Joviality	.31					-0.21	-1.33	-1.36
							vs. A	vs. Self-Assure
Confidence	.33						-1.08	0.49
Amusement	.43							
		Within PANAS-X						
PANAS-X		vs. H	vs. G	vs. F	vs. J	vs. SA	vs. At	
Sadness	.42	0.49	1.53	1.51	-0.15	1.41	**2.17****	
			vs. G	vs. F	vs. J	vs. SA	vs. At	
Hostility	.37		0.99	1.01	-0.63	0.93	1.69	
				vs. F	vs. J	vs. SA	vs. At	
Guilt	.29			0.03	-1.54	0.04	0.78	
					vs. J	vs. SA	vs. At	
Fear	.29				-1.56	0.01	0.75	
						vs. SA	vs. At	
Joviality	.43					1.70	**2.54***	
							vs. At	
Self-Assurance	.28						0.84	
Attentiveness	.21							

N = 174.

Note: All correlations significant $p < .01$.

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

Note: *Z*s for all comparisons are based on the trait in the left column versus the trait in the subheading (i.e., vs. H). A significant positive or negative *Z* indicates that the trait in the left column or the trait in the subheading is more dependable, respectively.

H = Hostility. GS = Guilt/Shame. F = Fear. J = Joviality. C = Confidence. A = Amusement. G = Guilt. SA = Self-Assurance. At = Attentiveness.

Table 43. MTMM for Self- and Other-Ratings for the ISEE Discrete Affects in the ILPP Sample

ISEE	Self-Ratings							Other-Ratings						
Discrete Affects	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<u>Self-Ratings</u>														
1. Sadness	--													
2. Hostility	.60	--												
3. Guilt/Shame	.51	.44	--											
4. Fear	.48	.45	.59	--										
5. Joviality	.46	.23	.17	.09	--									
6. Confidence	.30	.30	.13	.16	.62	--								
7. Amusement	.28	.21	.11	.13	.55	.37	--							
<u>Other-Ratings</u>														
8. Sadness	**.37**	.23	.17	.24	.32	.17	.09	--						
9. Hostility	.13	**.37**	.04	.13	.10	.20	.00	.59	--					
10. Guilt/Shame	.21	.18	**.23**	.29	.15	.17	.08	.59	.50	--				
11. Fear	.28	.23	.24	**.36**	.11	.11	.08	.62	.46	.71	--			
12. Joviality	.15	.08	-.02	.05	**.31**	.21	.26	.39	.22	.13	.18	--		
13. Confidence	.11	.07	.01	.10	.24	**.33**	.19	.35	.28	.28	.29	.61	--	
14. Amusement	.22	.09	.10	.08	.32	.22	**.43**	.26	.15	.14	.16	.60	.47	--

<u>N</u> = 174.

Note: All correlations ≥ .20 are significant at $p < .01$. Convergent correlations are in bold.

Table 44. MTMM for Self- and Other-Ratings for the PANAS-X Discrete Affects in the ILPP Sample

PANAS-X	Self-Ratings							Other-Ratings						
Discrete Affects	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Self-Ratings														
1. Sadness	--													
2. Hostility	.52	--												
3. Guilt	.65	.60	--											
4. Fear	.50	.55	.59	--										
5. Joviality	-.46	-.36	-.31	-.32	--									
6. Self-Assurance	-.24	-.01	-.23	-.15	.59	--								
7. Attentiveness	-.35	-.20	-.32	-.24	.55	.52	--							
Other-Ratings														
8. Sadness	**.42**	.21	.24	.27	-.19	-.05	-.13	--						
9. Hostility	.21	**.37**	.19	.22	-.22	-.01	-.06	.46	--					
10. Guilt	.26	.20	**.29**	.23	-.17	-.09	-.10	.59	.52	--				
11. Fear	.12	.07	.07	**.29**	-.04	-.05	-.06	.56	.50	.62	--			
12. Joviality	-.21	-.32	-.20	-.20	**.43**	.27	.15	-.28	-.39	-.28	-.24	--		
13. Self-Assurance	.00	.04	-.05	-.11	.14	**.28**	.11	-.11	-.08	-.15	-.21	.52	--	
14. Attentiveness	.00	-.09	-.04	-.08	.14	.09	**.21**	-.17	-.23	-.17	-.19	.54	.59	--

N = 174.

Note: All correlations ≥ |.20| are significant at $p < .01$. Convergent correlations are in bold.

CHAPTER 8. TEST-RETEST RELIABILITY

To support construct validity, and more specifically, to support the reliability of a new measure, the scale development process examined test-retest reliability. A sample of participants completed the same measures on two occasions separated by a time interval and the correlations between the scores on the first and second testing—across the sample of participants completing both assessments—estimated test-retest reliability. Although test-retest studies can be designed to examine true change in a trait across time, the current study focused on test-retest reliability across a short time interval to limit the amount of expected true change.

Dependability

Within test-retest reliability, Cattell argued for a distinction between dependability versus stability (Cattell, Eber, & Tatsuoka, 1970; Chmielewski & Watson, 2009; Watson, 2004). Test-retest studies of dependability examine short retest intervals to limit true change; whereas, test-retest studies of stability examine longer retest intervals to allow for true change. By limiting the time interval and minimizing the potential for true change, dependability analyses can directly model the measurement error within the assessment. Alternatively, with longer retest intervals (e.g., a year), stability contains both measurement error and true change; however, dividing stability by dependability can correct for measurement error (Chmielewski & Watson, 2009). "It therefore is important to conduct at least one well-designed study of dependability to validate a measure" (Watson, 2004, p. 328).

Furthermore, Watson (2004) recommended including other measures as benchmarks or points of comparison to facilitate the interpretation of dependability correlations. The design of Study 2 addressed both test-retest reliability and convergent validity and therefore allowed for the dependability of the newly developed measure to be examined in relation to existing convergent measures.

The current study consisted of a 2-week retest interval. In practical terms, the determination of the ideal retest interval needed to take several considerations into account: (a) traits differ in stability, (b) memory effects likely influence very short retest intervals, (c) retest intervals including a major life events need to be avoided (Watson, 2004). First, the stability of affect-based constructs tends to be lower than other constructs, such as personality (Chmielewski & Watson, 2009; Watson, 2004). Second, a 2-week interval reduced the potential for memory effects. Third, undergraduate students, the sample for this study, tend to be less stable (as a function of their younger age) compared to middle aged or older adults (Roberts & DelVecchio, 2000).

Previous research demonstrated strong levels of test-retest reliability for emotional expressivity. The original article on the development of the EES reported test-retest reliability of .90 across a 4-week retest interval (Kring et al., 1994). Similarly, the original article on the development of the BEQ reported test-retest reliability of .86 across a 2-3 month retest interval (Gross & John, 1995). The similarity of test-retest reliabilities across 1 to 3 months fit with the recent literature on the invariance of test-retest reliability across 2-week and 2-month retest intervals (Chmielewski & Watson, 2009). Based on these levels of test-retest reliability, I designed Study 1 with a 2-week retest interval and predicted dependability correlations greater than .85 for most of the ISEE.

Multi-Trait Multi-Occasion Matrices

Additionally, collecting test-retest data on the lower order ISEE allowed for a variant of a MTMM matrix, the multitrait-multioccasion matrix (Longley, Watson, & Noyes, 2005; Watson & Clark, 1991; Watson et al., 2007). To support test-retest reliability, or in other words, convergent validity across occasions, the test-retest correlations needed to be significant. The test-retest correlations formed the diagonal of the hetero-occasion block and needed to be higher than any other correlation in its row or column to demonstrate discriminant validity.

Method

Participants and Measures

Student and Retest Samples

Chapter 5 provided a description of the Student Sample (*N* = 387) and Chapter 6 provided a description of the Retest Sample (*N* = 346). The Student and Retest Samples completed all the same measures including: the ISEE, the EES (Kring et al., 1994), the BEQ (Gross & John, 1995), the BFI (Benet-Martinez & John, 1998; John & Srivastava, 1999), and the PANAS-X (Watson & Clark, 1999).

Results

Dependability

The 2-week test-retest correlations estimated the dependability of higher order scales of emotional expressivity, trait affectivity, and the Big Five (see Table 45). Dependability for the Big Five and trait affect in the current study compared similarly to published dependability estimates (see Table 2, Chmielewski & Watson, 2009), and thus validated the current test-retest methodology (i.e., in-person initial assessment and online retest assessment). Comparisons of the 2-week dependability estimates for E, N, C, A, and O revealed remarkable similarity between the current study and the published results, (.86, .85, .81, .79, .83 vs. .83, .83, .81, .78, .84, respectively). The largest difference existed for E (.03) and indicated a slightly higher estimate of dependability in the current study. Moreover, the comparisons of the dependability estimates of the current study with the average of 2-week and 2-month correlations in Chmielewski and Watson (2009) revealed even greater similarity: three out of the five matched exactly and the other two differed by a summed total of .03. Dependability estimates for PA also demonstrated similarity across the studies (.74 in current vs. .75 in published); however, NA demonstrated lower dependability in the current study ($r = .63$; 95% confidence interval .56 to .69) compared to the published estimate ($r = .73$; 95% confidence interval .68 to .77).

Tests for significant differences in dependability utilized the Pearson-Filon test for two correlations from the same sample (Kenny, 1987). Testing the differences within the BFI dependability estimates yielded another point of comparison with the published study (see Table 4, Chmielewski & Watson, 2009). In the current study, E exhibited greater dependability than A and C ($Z = 3.14$, $p < .01$; $Z = 2.36$, $p < .05$, respectively), and N also exhibited greater dependability than A ($Z = 2.42$, $p < .05$; see Table 45). All other comparisons within the BFI scales revealed nonsignificant differences (*Z*s range from |0.68| - |1.68|, $p > .05$). Importantly, the E versus A and N versus A comparisons replicated previous findings and again supported the comparability of the two studies.

Table 45 reports several additional tests for differences in dependability within and across personality, trait affect and emotional expressivity in the current study. In general, the BFI scales exhibited greater dependability than trait affect or emotional expressivity. E demonstrated significantly higher dependability than PA ($Z = 4.56$, $p < .01$), ISEE PE ($Z = 7.62$, $p < .01$), and BEQ Positive Expressivity ($Z = 6.28$, $p < .01$). Similarly N demonstrated significantly higher dependability than NA ($Z = 6.74$, $p < .01$), ISEE NE ($Z = 5.03$, $p < .01$), and BEQ Negative Expressivity ($Z = 4.14$, $p < .01$). The significant differences between the Big Five and trait affect replicated the findings from the published study (see Table 5, Chmielewski & Watson, 2009); the current study extends these findings into the affectively related domain of emotional expressivity.

The dependability estimates for the measures of emotional expressivity ranged from .83 to .62. The values for the EES and BEQ Total scale fit within the range of dependability estimates of the BFI scales. The ISEE General scale revealed significantly lower dependability than either the EES ($Z = -4.39$, $p < .01$), or the BEQ Total scale ($Z = -4.17$, $p < .01$). Additionally, the second-order level scales (i.e., ISEE PE and NE and BEQ Positive and Negative Expressivity) exhibited dependability estimates lower than the EES and BEQ Total scales.

Comparisons between trait affect and emotional expressivity and within emotional expressivity yielded mixed results. PA was more dependable than ISEE PE ($Z = 3.09$, $p < .01$); however, NA was less dependable than BEQ Negative Expressivity ($Z = -2.59$, $p < .01$). In general, PA and NA and ISEE PE and NE revealed opposite patterns of dependability: PA demonstrated significantly greater dependability than NA ($Z = 2.87$, $p < .01$) and, although nonsignificant, NE demonstrated slightly greater dependability than PE, ($Z = 1.89$, $p > .05$).

In Table 46 the 2-week dependabilities for the ISEE and PANAS-X discrete affect scales also exhibited a slight trend for greater dependability for PA and NE versus NA and PE, respectively. More generally, the dependabilities of discrete affect scales tended to be lower than the BFI scales except for the ISEE Sadness ($r = .75$) and PANAS-X Joviality ($r = .80$) and Self-Assurance ($r = .76$) scales. ISEE Sadness demonstrated the strongest dependability correlation among the ISEE discrete affects. Similarly, PANAS-X Joviality and Self-Assurance demonstrated significantly stronger dependability correlations than all the other PANAS-X discrete affects. Comparisons of discrete affects across the ISEE and PANAS-X scales revealed ISEE Sadness significantly more dependable than PANAS-X Sadness ($Z = 2.92$, $p < .01$) and both PANAS-X Joviality and Self-Assurance significantly more dependable than ISEE Joviality and Confidence ($Z = 3.82$, $p < .01$; $Z = 3.59$, $p < .01$, respectively).

A closer examination of the distributions of scale scores supported the potential influence of variability on dependability correlations (see Table 47 and 48 for descriptive statistics for all measures and samples). Dividing the scale means and standard deviations by the number of items comprising each scale equated scale means and standard deviations on a comparable 1 to 5 metric—the original response scale of the items. Within the BFI scales, for participants completing both the initial and retest assessments, the dependability correlation coefficients correlated .97 ($p < .01$) with the average of the item equated standard deviations across the two assessments. The BFI dependability correlation coefficients demonstrated perfect rank ordering with respect to the item equated standard deviations of these scales, that is, E exhibited the strongest dependability and the highest levels of variability; conversely, A exhibited the weakest dependability and the lowest levels of variability. Moreover, across all fifteen scales listed in Table 45, dependability correlations coefficients correlated .52 ($p < .05$) with the average item equated standard deviations for participants completing both the initial and retest assessments.

The contrasting patterns of dependability—that is, with NE more dependable than PE and PA dependable than NA—also fit with the impact of variability and, more specifically, the skewness of these scales. In particular, PE exhibited the lowest dependability among the higher order scales in Table 45. A large portion of participants reported a high level of PE (i.e., negatively skewed) with an average item equated mean across the two assessments of 4.07 on a 1 to 5 scale. Alternatively, a large portion of participants reported low levels of NA (i.e., positively skewed) with an average item equated mean across the two assessments of 1.97 on a 1 to 5 scale.

Two indices quantified the extremity of the item equated scale means and the skewness of the distributions of the scales. To index means closer to the end points of the 1 to 5 scale, the absolute value of the difference between the midpoint of the scale (i.e., 3) and the average of the item equated scale means across both assessments quantified the extremity of the mean. Additionally, the average across the two assessments of the absolute value of the skewness of the distributions of the scales quantified skewness. For the fifteen higher order scales on Table 45 the extremity and skewness indices correlated .94 ($p < .01$). As the item equated scale means approached the ends of the 1 to 5 scale, the more skewed the distribution. More importantly, across the fifteen higher order scales—for participants completing both the initial and retest assessments—the dependability correlations coefficients correlated negatively with both extremity and skewness ($r = -.59$, $p < .05$, and $r = -.58$, $p < .05$, respectively), that is, scales with item equated means closer to the end points of the 1 to 5 scale and scales with skewed distributions demonstrated lower dependability.

Multi-Trait Multi-Occasion Matrices

The multitrait-multioccasion matrices in Table 49 compared the ISEE PE and NE and the BEQ Positive and Negative Expressivity scales. First, the coefficient alphas (listed in parentheses) for ISEE PE and NE exceeded the coefficient alphas for the BEQ Positive and Negative Expressivity. Along the validity diagonal, the significant dependability correlations (in bold) supported convergent validity across occasions. In relation to the hetero-occasion block and the mono-occasion triangles, the dependability correlations passed the tests of discriminant validity; the ISEE and BEQ dependability correlations all significantly exceeded any other correlations within the same row or column (*Z*s range from 5.93 - 9.84 and 4.72 - 6.62, for the ISEE and BEQ, respectively, all *p*s < .01). Tests for significant differences in the comparison of correlated correlations coefficient followed the method outlined by Meng, Rosenthal, and Rubin (1992).

The discriminant correlations revealed the most striking differences between the top two matrices. More specifically, the discriminant correlation between self-rated BEQ Positive and Negative Expressivity ($r = .48$) significantly exceeded the discriminant correlation between self-rated ISEE PE and NE ($r = .29$; $Z = 3.59$, $p < .01$). Similarly the discriminant correlation between other-rated BEQ Positive and Negative Expressivity ($r = .50$) exceeded the discriminant correlation between other-rated ISEE PE and NE ($r = .32$; $Z = 3.15$, $p < .01$). The potential to improve the discriminant validity of the BEQ Positive and Negative Expressivity scales was further supported by the striking reduction of the discriminant correlations in the matrix of the revised BEQ 3-item Positive and Negative Expressivity scales, which is shown in the lowest third of Table 49.

The significant dependability correlations (in bold) along the validity diagonal in the ISEE discrete affects multitrait-multioccasion matrix (see Table 50) supported convergent validity across occasions. In relation to the hetero-occasion block and the mono-occasion triangles, the dependability correlations passed the tests of discriminant validity. Within the hetero-occasion block all dependability correlations exceeded any other discriminant correlation in the same row or column (*Z*s > 4.2, $p < .01$). Similarly, the dependability correlations exceeded are all discriminant correlations (within the same column) in the initial assessment mono-occasion triangle (*Z*s > 3.1, $p < .01$). However, a few comparisons between the dependability correlations and the discriminant correlations in retest assessment mono-occasion triangle (within the same row) failed to reach statistical significance. More specifically, the difference between the Fear dependability correlation ($r = .68$) and the retest assessment mono-occasion correlation of Fear and Sadness ($r = .65$; $Z = 0.9$, *ns*), the difference between the Confidence dependability correlation ($r = .63$) and the retest assessment mono-occasion correlation of Confidence and Joviality ($r = .61$; $Z = 0.5$, *ns*) and the difference between the Amusement dependability correlation ($r = .67$) and the retest assessment mono-occasion correlation of Amusement and Joviality ($r = .71$; $Z = -1.1$, *ns*) all lacked statistical significance. Although these comparisons lacked statistical significance, only the Amusement dependability correlation actually failed to exceed the discriminant correlations and thus failed the more stringent test of discriminant validity.

In addition to the support for both convergent and discriminant validity, this matrix also supported the existence of common variance due to the general factor and additional specific variance related to PE, NE and the discrete affects. Excluding the dependability correlations, the overall average of .30 for the discriminant correlations supported the existence of nonspecific variance related to the general factor of emotional expressivity. Furthermore, within the discriminant correlations, the within-valence correlations for the negative discrete affects (i.e., Sadness with Fear) averaged .42 and the positive discrete affects averaged .43. These moderate to strong within-valence correlations supported nonspecific variance at the second-order level of PE and NE. Taken together, moving from the average discriminant correlation of .30, to the average of the within-valence correlations of .42, to the average of the dependability correlations of .68, clearly supported additional variance from the general factor, from the PE and NE level, and finally from the discrete affect level.

In summary, these analyses of test-retest reliability provided initial support for the dependability of the ISEE across a 2-week interval; however, the differential levels of dependability need to be examined in future research. The MTMM matrices provided strong support for both convergent and discriminant validity and more generally provided support for the existence of nonspecific variance due to the general factor, as well as nonspecific and specific variance at both of the lower order levels.

Table 45. Dependability Correlations and Differential Dependability *Z*-Scores Differences for the BFI, PANAS, ISEE, BEQ, and EES in the Student Sample

Scale	2-week Dependability	Differential Dependability *Z*-Score Differences vs. BFI				vs. PANAS-X	vs. ISEE
BFI		vs. N	vs. C	vs. A	vs. O		
Extraversion	.86	0.72	**2.36****	**3.14***	1.68		
			vs. C	vs. A	vs. O		
Neuroticism	.85		1.64	**2.42****	0.97		
				vs. A	vs. O		
Conscientiousness	.81			0.78	-0.68		
					vs. O		
Agreeableness	.79				-1.45		
Openness	.83						
PANAS-X		vs. E				vs. NA	
Positive Affect	.74	**-4.56***				**2.87***	
		vs. N					
Negative Affect	.63	**-6.74***					
ISEE							
General	.69						
		vs. E				vs. PA	vs. ISEE NE
Positive Expressivity	.62	**-7.62***				**-3.09***	-1.89
		vs. N				vs. NA	
Negative Expressivity	.70	**-5.03***				1.69	
BEQ							vs. ISEE Gen
Total	.82						**4.17***
		vs. E				vs. PA	vs. ISEE PE
Positive Expressivity	.68	**-6.28***				-1.71	1.40
		vs. N				vs. NA	vs. ISEE NE
Negative Expressivity	.73	**-4.14***				**2.59***	0.92
Impulse Strength	.78						
EES							vs. ISEE Gen
Total	.83						**4.39***

N = 346 (except N = 345 for BFI).

Note: All correlations significant $p < .001$. * $p < .01$. ** $p < .05$.

Note: *Z*s for all comparisons are based on the trait in the left column versus the trait in the subheading (i.e., vs. E). A significant positive or negative *Z* indicates that the trait in the left column or the trait in the subheading is more dependable, respectively.

E = Extraversion. N = Neuroticism. C = Conscientiousness. A = Agreeableness. O = Openness. ISEE Gen = General scale. ISEE NE = Negative Expressivity. ISEE PE = Positive Expressivity. PA = Positive Affect. NA = Negative Affect.

Table 46. Dependability Correlations and Differential Dependability *Z*-Scores Differences for the ISEE and PANAS-X Discrete Affects in the Student Sample

Discrete Affects	2-week Dependability	Differential Dependability *Z*-Scores Differences						
		Within ISEE						**vs. PANAS-X**
ISEE		vs. H	vs. GS	vs. F	vs. J	vs. C	vs. A	vs. Sadness
Sadness	.75	1.42	**2.72***	1.72	**2.04****	**3.02***	**2.12****	**2.92***
			vs. GS	vs. F	vs. J	vs. C	vs. A	vs. Hostility
Hostility	.69		1.30	0.30	0.63	1.61	0.71	0.67
				vs. F	vs. J	vs. C	vs. A	vs. Guilt
Guilt/Shame	.64			-1.01	-0.67	0.31	-0.59	-0.03
					vs. J	vs. C	vs. A	vs. Fear
Fear	.68				0.33	1.30	0.41	1.80
						vs. C	vs. A	vs. Joviality
Joviality	.67					-0.90	0.08	**-3.82***
							vs. A	vs. Self-Assure
Confidence	.63						-0.90	**-3.59***
Amusement	.67							
		Within PANAS-X						
PANAS-X		vs. H	vs. G	vs. F	vs. J	vs. SA	vs. At	
Sadness	.63	-0.85	-0.25	0.59	**-4.69***	**-3.49***	-0.58	
			vs. G	vs. F	vs. J	vs. SA	vs. At	
Hostility	.67		0.60	1.44	**-3.85***	**-2.64***	0.26	
				vs. F	vs. J	vs. SA	vs. At	
Guilt	.64			0.84	**-4.44***	**-3.23***	-0.33	
					vs. J	vs. SA	vs. At	
Fear	.60				**-5.26***	**-4.06***	-1.15	
						vs. SA	vs. At	
Joviality	.80					1.20	**4.12***	
							vs. At	
Self-Assurance	.76						**2.92***	
Attentiveness	.66							

N = 346.

Note: All correlations significant $p < .001$.

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

Note: *Z*s for all comparisons are based on the trait in the left column versus the trait in the subheading (i.e., vs. H). A significant positive or negative *Z* indicates that the trait in the left column or the trait in the subheading is more dependable, respectively.

H = Hostility. GS = Guilt/Shame. F = Fear. J = Joviality. C = Confidence. A = Amusement. G = Guilt. SA = Self-Assurance. At = Attentiveness.

Table 47. Descriptive Statistics for the BFI, ISEE, PANAS-X, BEQ, and EES in the Student, Retest, ILPP and Informant Samples

Scales (# items)	Student Sample				Retest Sample				ILPP Sample				Informant Sample			
	M	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
BFI																
Extraversion (8)	3.57	.81	-.45	-.55	3.52	.79	-.38	-.25	3.40	.86	-.29	-.73	3.67	.79	-.43	-.17
Neuroticism (8)	2.89	.75	.07	-.67	2.91	.73	.08	-.37	2.73	.84	.04	-.59	2.71	.82	-.05	-.89
Conscientiousness (9)	3.61	.66	-.31	-.06	3.60	.62	-.10	-.05	4.00	.59	-.64	.31	4.20	.59	-.95	.98
Agreeableness (9)	3.85	.61	-.67	.69	3.82	.64	-.39	-.18	4.02	.59	-.49	-.29	4.08	.60	-.65	.33
Openness (10)	3.59	.70	-.25	-.09	3.55	.67	-.31	.28	3.55	.70	-.09	-.46	3.74	.64	-.33	.25
ISEE																
General (6)	3.82	.69	-.57	.14	3.63	.78	-.17	-.65	3.87	.77	-.54	-.44	3.95	.73	-.75	.21
Positive Express (12)	4.08	.50	-.72	.94	4.07	.58	-.59	.53	4.10	.54	-.59	.28	4.20	.45	-.25	-.38
Negative Express (12)	3.15	.65	-.15	-.49	3.22	.71	.01	-.31	2.99	.76	-.14	-.03	3.12	.68	-.40	-.02
PANAS-X																
Positive Affect (10)	3.69	.54	-.22	.07	3.56	.61	-.20	-.06	3.43	.57	-.13	-.10	3.64	.55	-.15	.27
Negative Affect (10)	1.93	.54	.63	-.04	2.01	.63	.57	-.24	1.73	.51	.79	.16	1.66	.46	.86	.55
BEQ																
BEQ Total (16)	3.35	.64	-.11	-.38	3.33	.63	.13	-.42	--	--	--	--	--	--	--	--
Positive Express (4)	4.01	.64	-.67	.46	3.87	.70	-.32	-.55	--	--	--	--	--	--	--	--
Negative Express (6)	2.89	.73	.14	-.23	2.93	.70	.09	-.13	--	--	--	--	--	--	--	--
Impulse Strength (6)	3.36	.86	-.30	-.47	3.37	.81	-.05	-.70	--	--	--	--	--	--	--	--
EES																
EES Total (17)	3.26	.74	-.03	-.62	3.23	.67	.06	-.18	--	--	--	--	--	--	--	--

N = 386-387 (Student Sample). N = 346 (Retest Sample). N = 344 (ILPP Sample). N = 174 (Informant Sample).

Table 48. Descriptive Statistics for the ISEE and PANAS-X Discrete Affects in the Student, Retest, ILPP and Informant Samples

	Student Sample				Retest Sample				ILPP Sample				Informant Sample			
Scales (# items)	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
ISEE																
Sadness (6)	3.30	.77	-.29	-.39	3.32	.85	-.19	-.50	3.28	.87	-.32	-.35	3.46	.79	-.42	-.13
Hostility (6)	3.33	.74	-.29	-.23	3.37	.81	-.26	-.18	3.24	.89	-.28	-.40	3.33	.86	-.24	-.59
Guilt/Shame (6)	2.96	.79	-.11	-.26	3.05	.81	-.11	-.17	2.69	.84	.05	-.31	2.76	.71	-.30	.02
Fear (6)	3.27	.78	-.35	-.23	3.34	.80	-.19	-.33	2.86	.87	-.15	-.38	2.78	.77	-.24	-.36
Joviality (7)	4.32	.57	-.91	.69	4.20	.65	-.70	-.03	4.26	.61	-.80	.36	4.36	.47	-.52	-.27
Confidence (7)	3.77	.65	-.26	-.16	3.76	.68	-.32	.34	3.76	.69	-.41	.23	3.87	.61	-.53	.52
Amusement (6)	4.14	.62	-.90	.99	4.21	.64	-.88	.70	4.12	.63	-.93	1.17	4.19	.57	-.59	-.36
PANAS-X																
Sadness (5)	1.92	.76	.82	.04	2.02	.79	.71	-.11	1.66	.65	1.22	1.60	1.66	.59	.93	.65
Hostility (6)	1.84	.62	.71	-.05	1.90	.68	.77	-.02	1.67	.55	.99	.66	1.73	.58	1.11	.91
Guilt (6)	1.72	.69	1.22	1.49	1.82	.74	1.00	.84	1.57	.63	1.47	1.98	1.47	.52	1.62	3.10
Fear (6)	1.90	.63	.75	.17	1.97	.68	.55	-.22	1.59	.49	.94	.56	1.52	.46	.95	.47
Joviality (8)	3.75	.65	-.29	-.16	3.65	.71	-.21	-.17	3.34	.62	-.06	-.14	3.48	.62	-.35	1.19
Self-Assurance (6)	3.33	.71	-.08	-.41	3.25	.71	-.09	-.12	2.90	.64	.18	-.04	3.18	.64	-.02	-.63
Attentiveness (4)	3.61	.61	-.36	.22	3.49	.64	-.15	-.07	3.57	.62	-.08	-.17	3.87	.56	-.48	.19

N = 387 (Student Sample). N = 346 (Retest Sample). N = 344 (ILPP Sample). N = 174 (Informant Sample).

Table 49. MTMM-Occasion for ISEE, BEQ and 3-Item BEQ Positive and Negative Expressivity in the Student Sample

ISEE Scales	Initial-Ratings 1	2	Retest-Ratings 3	4
Initial-Ratings				
1. Positive Expressivity	(.83)			
2. Negative Expressivity	.29	(.83)		
Retest-Ratings				
3. Positive Expressivity	**.62**	.22	(.89)	
4. Negative Expressivity	.15	**.70**	.32	(.88)

BEQ Scales	Initial-Ratings 1	2	Retest-Ratings 3	4
Initial-Ratings				
1. Positive Expressivity	(.73)			
2. Negative Expressivity	.48	(.75)		
Retest-Ratings				
3. Positive Expressivity	**.68**	.48	(.72)	
4. Negative Expressivity	.48	**.73**	.50	(.75)

Revised BEQ Scales	Initial-Ratings 1	2	Retest-Ratings 3	4
Initial-Ratings				
1. 3-Item Pos Express	(.69)			
2. 3-Item Neg Express	.16	(.62)		
Retest-Ratings				
3. 3-Item Pos Express	**.59**	.25	(.70)	
4. 3-Item Neg Express	.22	**.66**	.20	(.59)

N = 346.

Note: All correlations significant at $p < .01$. Test-Retest correlations are in bold.

Table 50. MTMM-Occasion for the ISEE Discrete Affects in the Student Sample

Discrete Affects	Initial Assessment							Retest Assessment						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Initial Assessment														
1. Sadness	--													
2. Hostility	.45	--												
3. Guilt/Shame	.43	.29	--											
4. Fear	.50	.30	.50	--										
5. Joviality	.32	.16	.25	.17	--									
6. Confidence	.14	.27	.27	.09	.46	--								
7. Amusement	.18	.14	.16	.17	.53	.27	--							
Retest Assessment														
8. Sadness	**.75**	.35	.27	.44	.27	.07	.14	--						
9. Hostility	.43	**.69**	.20	.29	.13	.20	.09	.57	--					
10. Guilt/Shame	.44	.27	**.64**	.46	.19	.22	.11	.55	.44	--				
11. Fear	.49	.25	.36	**.68**	.17	.05	.12	.65	.47	.60	--			
12. Joviality	.27	.13	.22	.19	**.67**	.35	.43	.38	.29	.26	.30	--		
13. Confidence	.13	.19	.20	.13	.44	**.63**	.18	.23	.32	.33	.22	.61	--	
14. Amusement	.15	.05	.12	.15	.45	.15	**.67**	.27	.21	.16	.24	.71	.37	--

N = 346.

Note: All correlations ≥ .14 are significant at $p < .01$. Test-retest correlations are in bold.

CHAPTER 9. GENERAL DISCUSSION

The conceptualization of a three-level hierarchical structure of emotional expressivity integrated existing models and led to the development of the ISEE—a hierarchical set of scales that targeted three different levels of abstraction. Following the substantive, structural and external phases of scale development (Clark & Watson, 1995), the previous chapters reported on the initial development of the ISEE. Chapter 4 covered the substantive phase including the development of the three-level hierarchical model and pilot-testing of the initial item pool. In Chapter 5, the structural phase consisted of a series of factor analyses in a bottom-up approach to develop the ISEE. The external phase—started in Chapter 6 and extended into Chapter 7—addressed the nomological network and convergent and discriminant validity. Finally, Chapter 8 focused on test-retest reliability and specifically examined the dependability of the ISEE in relation to other measures. This chapter focuses on a general discussion, examining strengths and limitations of these studies, and exploring future directions for this line of research.

Structural Validity

The structural phase focused on creating scales aligned with the three-level hierarchical model of emotional expressivity. The analytic strategy consisted of first creating the discrete affect scales and then using these scales in a bottom-up approach to create scales assessing the higher order factors. Selecting the strongest markers of the higher order factors maximized the amount of common variance related to the higher order factors and minimized the specific variance of the lower order factors. PE and NE contained a representative subset of items from the seven discrete affect scales; the ISEE General scale contained items at a higher level of abstraction, without valence or discrete affect content (e.g., "I'm emotionally expressive"). With 50 items (44 discrete and 6 general), the ISEE assessed all three levels of the hierarchical model. Nearly all of the coefficient alphas of the ISEE replicated above the target alpha of .80; however, the

Amusement scale was below .80 in both samples and the Fear and Hostility scales were below .80 in the Student Sample.

Incorporating the appropriate level of abstraction of the items into the hierarchical scales influenced the meaningfulness of the relative comparisons of the AICs. As predicted, the relatively narrow discrete affect scales demonstrated AICs in the .40 - .50 range and the AICs of these scales exceeded the AICs of PE and NE. Typically, broader constructs are associated with lower AICs (Clark & Watson, 1995); however, the AIC of the General scale also exceeded the AICs of PE and NE (i.e., .40 vs. .29 and .29 in the Student Sample and .47 vs. .36 and .36 in the ILPP Sample, respectively). Although general emotional expressivity is a broader construct than positive and negative expressivity, the items of the ISEE General scale—without discrete affective content or specific nonverbal behaviors—are more abstract and, in some respects, are actually narrower compared to the items of the PE and NE scales, which are comprised of a subset of the discrete affect items. If different sets of items incorporate different levels of abstraction, then the relative comparison of the AICs based on the breadth of the construct is less certain and may not fit with the typical pattern of broader construct having lower AICs.

More importantly, the level of abstraction in the items can influence structural analyses, that is, the inclusion of more abstract items in analyses testing for lower order structure can inflate the correlations between the factors. For instance, to test this idea, a Promax-rotated two-factor PFA of the 12 PE and 12 NE items also included the six ISEE General items. In the Student Sample, all six general items loaded greater than .30 on either Positive or Negative Expressivity—with an average difference of only .06 between the primary loading and the secondary loading—and three out of the six general items loaded greater than .30 on both factors. As a result of the general items splitting across the two factors, the correlation between the factors increased from .30 (without the general items) to .38 (with them). These analyses confirm that failing to account for the level of abstraction of the items influences the structure.

Several analytic methods supported the three-level hierarchical model: (a) factor analyses of the discrete affect items supported one-, two-, and seven-factor solutions, (b) correlations between successive sets of factor scores and the discrete affect scales clearly aligned with the three levels of the hierarchy, (c) second-order factor analyses of the correlation matrix of the discrete affect factors supported both higher order levels. At the item level, structural analyses replicated across two samples; moreover, comparability coefficients exceeded .90 at the discrete affect level and exceeded .99 at the higher order levels. Additionally, the moderate correlation between positive and negative expressivity fit with the broad general factor of emotional expressivity but also emphasized the relative independence of these two dimensions. Based on second-order factor analyses, the positive and negative expressivity factors correlated .43 and .36 in the Student and ILPP Samples, respectively. Similarly, the PE and NE scales correlated approximately .30 in the Student, Retest, ILPP, and Informant Samples.

Confirming the prediction that the general factor of emotional expressivity would account for at least 40% of the common variance, the first factor accounted for about half the common variance and the second factor accounted for an additional 20% - 30% of the common variance. In comparison, even if the seven discrete affect factors perfectly accounted for all the common variance, on average each discrete affect factor would only account for approximately 15% of the common variance. Clearly, these are broad higher order factors, accounting for a substantial proportion of the common variance among more specific discrete affect factors.

Although the second-order factor analyses supported both a general factor and higher order factors of positive and negative expressivity, the only other possible solution included a three-factor solution. With a minimum of two discrete affect factors needed to mark a higher order factor, three is the maximum number of higher order factors possible from the seven discrete affect factors. For instance, the four negative discrete affects could split to form a third higher order factor, but the three positive discrete affects could not split and have two clear markers on more than one factor. Even so, one of the discrete affects, from either the negative or positive discrete affects, could have split and had its strongest loading on the third factor; however, this was not the case.

The MAP test and parallel analyses—statistically based methods of determining the number of factors to retain—indicated between five to nine factors should be retained. Only the parallel analyses based on the ILPP Sample supported less than seven factors; all other MAP tests and parallel analyses indicated seven or more factors. O'Connor (2000a) encouraged researchers to run both the MAP test and parallel analysis because results may differ and furthermore stated that if these tests err, then MAP tests tend to underextract factors whereas parallel analyses tend to overextract factors. In the current study the results of these tests differed; however, assuming the seven-factor solution at the discrete affect is the correct number of factors to extract, then the MAP tests tended to indicate overextraction of factors whereas the parallel analyses tended to indicate underextraction of factors. Other researchers have also reported that parallel analyses may underextract factors, in particular, with multilevel factor structures (see Schweizer, 1992; Turner, 1998). Although the MAP tests and parallel analyses informed the process of determining the number of factors to retain, ultimately, structural analyses at the discrete affect level aimed to identify the maximum number of possible factors and the factors needed to be well defined (i.e., a minimum of three clear markers) to be retained.

The strong effect sizes for sex differences replicated previous research on expressivity (Gross & John, 1995, 1997; Kring et al., 1994; Kring & Gordon, 1998). On all the ISEE, females reported higher mean levels in both samples, except for the novel finding that males in the Student Sample reported higher mean levels of expressed Confidence. This sex difference on the Confidence scale needs to be replicated because even in the two existing data sets the findings are inconsistent, that is, males in the ILPP

Sample reported lower levels of Confidence relative to females. Although the two samples differ in other respects, one primary difference is the age of the participants. Perhaps college-aged males are more expressive of Confidence than males in their early 30s (i.e., those in the ILPP sample). Overall, younger males (i.e., Student Sample) reported higher levels of emotional expressivity than older males (i.e., ILPP Sample). Additionally, in the Student Sample, sex differences on the BEQ Total scale were twice as large as the ISEE General and EES and likely resulted from the inclusion of the Impulse Strength facet in the BEQ Total scale.

Nomological Network

Establishing the interconnections with personality and trait affect elaborated the nomological network of emotional expressivity; in this regard, striking differences were revealed between the general factor and positive and negative expressivity. Replicating a recent meta-analysis (Riggio & Riggio, 2002), the ISEE General scale correlated moderately with E and weakly or nonsignificantly with other remaining Big Five traits. Alternatively, the relations between the lower order levels of emotional expressivity revealed a strikingly different pattern. Confirming predictions, PE and NE demonstrated clear differential relations with E and N, and with PA and NA, respectively.

The discrete affects of positive and negative expressivity also demonstrated consistent differential relations with either E or N, respectively. Additionally, PE, Joviality, and Confidence correlated with O and both Hostility and Joviality demonstrated specificity in relation to A. Overall, the discrete affect level supported both generality and specificity. These differential relations between PE and NE, together with the specificity at the discrete affect level, highlight the importance of assessing the lower order levels of emotional expressivity.

The magnitude of the relations within the nomological network and the differential patterns of convergence for PE and NE have implications for the conceptualizations of these constructs and for the hierarchical structure of emotional expressivity. Although PE and NE correlated at about .30, PE and NE are both more strongly correlated with E and N at about .45 - .50, respectively. In other words, PE shared more variance with E than with NE. Similarly, NE shared more variance with N than PE. The moderate level of relation between PE and NE and the strong correlations with independent traits of personality strongly supported the relative independence of PE and NE.

Furthermore, both emotional expressivity and trait affect aligned more closely with E and N than with each other. The correlations between the corresponding second-order level factors of emotional expressivity and trait affect scales were less than the correlations of these scales with the corresponding traits of E and N. Although emotional expressivity and trait affect are moderately related, the magnitude of the correlations suggest quite a bit of differentiation between emotional expressivity and trait affectivity; this is particularly noteworthy considering the strong overlap in affective content between the ISEE and PANAS-X.

Hierarchical regressions further demonstrated the incremental predictive validity of PE and PA in predicting E and NE and NA in predicting N. Whereas E and N both have an experiential component—such that E is strongly related to PA and N is strongly related to NE—they also both have an expressive component. Considering that the relation between emotional expressivity and N has not been as consistently supported, it is noteworthy that NE accounted for a greater amount of unique variance in N than PE accounted for unique variance in E.

A variant of traditional MTMM analyses revealed convergence between the ISEE and PANAS-X discrete affects with mixed support for discriminant validity. Positive discrete affects demonstrated stronger convergence than the negative discrete affects; however, widely differing levels of convergence existed within the negative discrete affects (range .07 to .50). In particular, Sadness and Guilt/Shame exhibited the weakest support for convergence between emotional expressivity and trait affect. This suggests that there is little relation between the trait affect and trait expressivity of these particular affective states. More broadly, the varying levels of convergence between the discrete affects of emotional expressivity and trait affect stresses the importance of assessing the discrete affect levels for both emotional expressivity and trait affect.

These MTMM analyses also emphasized the distinctiveness of emotional expressivity and trait affect. Although the convergence of emotional expressivity and trait affect demonstrated systematic, affect-specific variance at the discrete affect level, these correlations were far from being 1.0. These moderate correlations indicated that people are responding to more than just the affective content and are likely accounting for the specific expressive content of the ISEE items.

In spite of the moderate levels of convergence between emotional expressivity and trait affect, strong structural similarities existed. More specifically, the stronger within domain correlations (i.e., mono-scale triangles) compared to the hetero-scale correlations supported the existence of higher order factors of PE and NE and PA and NA. Although there was systematic variance across emotional expressivity and trait affect at the discrete affect level, there was also a substantial amount of nonspecificity within emotional expressivity and trait affect.

<u>Convergent and Discriminant Validity</u>

Moderate to strong convergence between the ISEE, EES (Kring et al., 1994) and BEQ (Gross & John, 1995) provided support for convergent validity; at the same time, however, these coefficients were not so strong to indicate redundancy. The ISEE General scale was modeled after the EES and the convergence (r = .74) between these scales confirmed the prediction that they would correlate very strongly (i.e., r > .70). The strong convergence between the ISEE General scale and the BEQ Total scale (r = .71) confirmed the prediction that these two scales would also correlate strongly (i.e., r > .50). At the second-order level, ISEE and BEQ positive and negative expressivity scales demonstrated strong convergent validity (rs = .67 and .57, respectively); these correlations exceeded the prediction for only moderate to strong correlations (r = .30 - .50). Moreover, the ISEE and BEQ within-valence convergent correlations significantly exceeded the cross-valence correlations and demonstrated discriminant validity.

Correlations between the ISEE and BEQ suggested that the BEQ Total and Impulse Strength scales are tipped towards negative expressivity. The BEQ Total scale correlated similarly with the ISEE General scale and ISEE Sadness scale. The BEQ Impulse Strength scale was more strongly related to ISEE NE than PE—this imbalance would tip the BEQ Total scale toward negative expressivity. After partialing out the nonspecific variance of the general scales, the BEQ Total scale still contained specific variance related to ISEE NE. The relative balance of positive and negative expressivity is important to consider in establishing the connections between the general level and the discrete emotions level.

In addition to examining convergence with existing measures of emotional expressivity, the use of both self- and other-ratings in the ILPP sample supported convergent validity through self-other agreement. All scales demonstrated significant self-other agreement correlations, although the BFI tended to have stronger correlations than the ISEE and PANAS-X. Agreement on E was significantly higher than on A, O, the ISEE General scale, PE and PA. Agreement on N was significantly higher than both NE and NA. However, there were no significant differences in self-other agreement correlations between PE and PA, between NE and NA, and between corresponding discrete affects from the ISEE and PANAS-X scales.

Higher levels of self-other agreement across the Big Five compared to trait affect have been interpreted as the trait visibility effect—traits with relevant, observable behaviors result in higher agreement than more internal, subjective traits (Funder & Colvin, 1988; Ready, Clark, Watson, & Westerhouse, 2000; Watson & Clark, 1991; Watson et al., 2000). Although emotional expressivity is conceptualized as the tendency to display visibly expressive behaviors, the comparisons between emotional expressivity and trait affect fail to fit with the trait visibility effect. Overall, there is little support for the trait visibility effect based on comparisons with emotional expressivity: (a) agreement on emotional expressivity was lower than the Big Five, (b) comparable to trait affect, and (c) there was little differentiation between positive and negative emotional expressivity.

As noted earlier, Trierweiler et al. (2002) attributed stronger agreement correlations for positive discrete affects compared to negative discrete affects to the trait visibility effect. Participants reported expressing positive discrete affects more often (i.e., higher trait visibility) than negative discrete affects. In the current study, agreement correlations for PE and NE were equivalent; however, there was significantly higher agreement for Amusement compared to Guilt/Shame. Similarly, Trierweiler et al. (2002) reported the lowest agreement for Shame. Expressions of guilt and shame likely occur relatively infrequently compared to other discrete affects and may be subtler than expressions of other discrete affects.

Previous authors have theorized that narrower, less abstract constructs should yield higher levels of self-other agreement compared to broader, more abstract constructs (Hampson, John, & Goldberg, 1986; Urbina, 2004); however, the current study revealed stronger agreement for broader constructs. Emotional expressivity and trait affect were conceptualized as narrower components of the broader personality factors, but the Big Five all had stronger self-other agreement coefficients than emotional expressivity and trait affect. In the current study, there are certainly other issues potentially influencing these results. If the ISEE and PANAS-X scales are less dependable or if emotional expressivity and trait affect are actually less stable than the Big Five, both of these factors could contribute to lower levels of self-other agreement. As will be discussed shortly, skewness may also be setting ceilings for the ISEE and PANAS-X agreement correlations.

More broadly, these significant self-other agreement correlations support the basic idea that people have insight into their own levels of emotional expressivity. Agreement between two independent raters about one's level of emotional expressivity is an important piece of evidence supporting construct validity. Additionally, recent research addressed the accuracy of recalling nonverbal behaviors—a related but broader domain than emotional expressivity—and participants demonstrated nonverbal self-accuracy by recalling, without forewarning, nonverbal behaviors (e.g., smiling, gazing and gesturing) that occurred during an earlier interaction (Hall, Murphy, & Schmid Mast, 2007).

Comparing self-other agreement correlations to discriminant correlations through MTMM analyses simultaneously addressed convergent and discriminant validity. Convergent correlations (i.e., self-other agreement) exceeded discriminant correlations even though discriminant correlations were also significant. In other words, significant discriminant correlations revealed the common variance among constructs within a particular level of the hierarchy. For instance, PE and NE correlated significantly, that is, both scales contained common variance related to the general factor of emotional expressivity. However, the self-other agreement correlations for PE and NE exceeded the correlations between PE and NE, thus supporting discriminant validity. At the discrete affect level, significant self-other agreement correlations supported convergent validity and most agreement correlations exceeded the other correlations within the same row or column of the heterotrait-heteromethod block, again demonstrating discriminant validity. As hypothesized, most self-other agreement correlations for the discrete affects also exceeded the cross-valence correlations within the monomethod triangles—an intermediate test of discriminant validity.

More importantly, ISEE PE and NE demonstrated improved discriminant validity relative to the corresponding BEQ scales. Based on the preliminary study utilizing the BEQ and exploratory 3-item scales, the predicted correlation between positive and negative expressivity was in the .30 to .50 range. In the current studies—across self- and other-ratings, and at the initial and retest assessments—correlations between PE and NE, as well as correlations between factors of positive and negative expressivity, all supported the relative independence of positive and negative expressivity, with a moderate correlation near .30. Moreover, a direct comparison of the correlation between ISEE PE and NE and BEQ Positive and Negative Expressivity revealed a significantly lower correlation for the former compared to the latter. The importance of the relative independence of PE and NE is further demonstrated in the clear differential relations exhibited in the nomological network.

There are additional considerations to take into account when examining the level of relation between positive and negative expressivity. First, method variance can inflate correlations that share the same method, that is, the correlation between positive and negative expressivity based solely on self-reports is an overestimate of the true relation due to shared methods. Second, correction for attenuation would serve to increase the correlation between positive and negative expressivity. To address these issues, MTMM analyses using confirmatory factor analysis have been developed to separate trait, method and error variance components (Eid et al., 2003); I will return to this issue in the section on future directions.

Test-Retest Reliability

The current examination of test-retest reliability utilized a brief retest interval, collected data from a large sample and included comparison measures as benchmarks (see Watson, 2004). The 2-week test-retest interval minimized the potential for true change and focused on the dependability of the scales. The sample of more than 300 participants resulted in narrower confidence intervals around the test-retest correlations and allowed for meaningful comparisons between scales. The inclusion of existing scales of emotional expressivity, affectivity and personality served as points of comparison to interpret dependability.

Disconfirming the prediction for most ISEE dependability correlations to be greater than .85, the ISEE 2-week dependability correlations ranged from .60 to .75. The two possible explanations for these lower than expected levels of dependability are (a) that emotional expressivity is actually less trait-like than hypothesized, or (b) that the ISEE are less dependable than they should be due to the increased influence of measurement error. Based on comparisons with previous research and the inclusion of alternative scales of emotional expressivity in the current study, it appears that emotional expressivity is relatively stable; consequently, the more likely explanation is that the ISEE exhibited less than optimal dependability.

Compared to the ISEE General scale, both the EES and BEQ demonstrated stronger dependability correlations in the current study and in previous research. Based on 102 undergraduate students across a 4-week retest interval, Kring et al. (1994) reported the test-retest reliability of the EES was .90 (95% confidence interval of .86 - .93). In the current study, the EES dependability was slightly lower than the previously reported estimate (r = .83; 95% confidence interval .78 - .85). In a sample of 68 students across a 2- to 3-month retest interval, Gross and John (1995) reported the test-retest reliability of the BEQ was .86 (95% confidence interval of .78 - .91). In the current study, the BEQ dependability (r = .82; 95% confidence interval .79 - .86) was within the 95% confidence interval of the previously reported estimate. Although the differences in the retest interval across the studies may have slightly influenced the retest correlations, recent research has demonstrated highly similar levels of test-retest correlations for 2-week and 2-month retest intervals (Chmielewski & Watson, 2009).

The lower dependability of the ISEE may be related to the length of the scale, and subtle differences in response formats (i.e., extremity of wording or scale anchors, or the number of response options). Both the EES and BEQ Total scale contain more than two and a half times as many items as the six-item ISEE General scale. In addition to the length of the scales, differences in response formats could also be contributing to lower dependability. The ISEE response scale contains the following descriptors: *disagree, slightly disagree, neutral, slightly agree, agree*. By making the anchor points less extreme (i.e., *disagree* and *agree* vs. *strongly disagree* and *strongly agree*) the intent was to increase the use the full range of the response scale, and thus maximize the amount of variability. Alternatively, in hindsight, making the end points less extreme may have disproportionately influenced the agree end of the response scale, that is, most items were negatively skewed.

The differences in dependability for the ISEE and PANAS-X are difficult to interpret as a function of item response format (i.e., adjectives vs. sentences), instructional differences, or affective content of the scales. Although, in general, the ISEE dependability correlations were within the same range as the PANAS-X, PE was less dependable than PA and NE was more dependable than NA, although this second comparison failed to reach statistical significance. At the discrete affect level, ISEE Joviality was significantly less dependable than PANAS-X Joviality; conversely, ISEE Sadness was significantly more dependable than PANAS-X Sadness. Therefore, differences in item response format and instructional differences across the ISEE and PANAS-X are unable to account for these differences in dependability. Additionally, ISEE Sadness was significantly more dependable than ISEE Joviality although these scales shared identical item format, response scales and instructions. Moreover, affective content is an unlikely explanation for these differences in dependability because the affective content of the ISEE and PANAS-X scales are very similar, that is, the discrete affect terms of the PANAS-X served as a starting point for the creation of the ISEE.

Although the ISEE and PANAS-X dependability correlations were within the same range, the Big Five were more stable than either emotional expressivity or trait affect. E was significantly more dependable than PE and PA; similarly, N was significantly more dependable than NE and NA. These significant differences between the Big Five and trait affect replicated previous findings (Chmielewski & Watson, 2009; Watson, 2004) and extended these findings into the affectively related domain of emotional expressivity. Even within the Big Five, E was significantly more dependable than A or O; thus, the current research replicated the previous findings of E and A demonstrating the strongest and weakest dependabilities, respectively, within the Big Five (Chmielewski & Watson, 2009). Dependability correlations, like all correlations, depend on variability and after equating the BFI scales by the number of items contained in each measure, there was perfect rank ordering of the dependability correlations with respect to the variability in these scales.

The most striking comparisons on dependability for conceptually related scales are between E and PE (.86 vs. .62) and between N and NA (.85 vs. .63). Examining the distribution of the scale scores revealed that both PE and NA are more skewed than E and N, respectively. Based on the limited number of higher order scales, skewness correlated negatively with dependability correlations. In previous research, skewed variables reduced the magnitude of test-retest correlations based on Pearson product-moment correlations (Dunlap, Chen, & Greer, 1994). Test-retest correlations—which necessarily involve assessing the same scale twice—likely compound the influence of skewness, that is, correlating two skewed variables would likely create a reduced ceiling for dependability. For instance, the correlation between PE and E may be relatively strong even if PE is negatively skewed and E is skewed in the same direction but to a lesser extent; however, the test-retest correlation for PE, based on two variables with skewed distributions, may be significantly less than the test-retest correlation for E.

Although the current research focused on dependability correlations, clearly the potential influence of limited variability and skewness also exists in long-term stability correlations. For instance, N was significantly more stable than NA across a 2.5-year interval, even though N and NA were strongly correlated (Vaidya et al., 2002, Watson, 2004). If in this previous research NA was more skewed than N, as was the case in the current dependability study, then the lower stability of NA is likely partially a result of skewness. Moreover, long-term stability correlations systematically decrease with increasing time intervals (Watson, 2004) and weaker correlations are even more influenced by skew than are stronger correlations (Greer et al., 2006; Dunlap, Chen, & Greer, 1994).

To reduce the negative effects of skewness, Greer et al. (2006) recommended using more response options (with potential benefits from including more than five options) or data transformation through ranking. A reasonable first step prior to resorting to data transformation would be to attempt to minimize the skewness of the scales through increased response options in future data collection; that is, if less skewed scales can be created, then data transformations would be unnecessary. Additionally, preliminary exploratory analyses examining the influence of skew revealed only slight differences between test-retest correlations based on Pearson product-moment correlations versus Spearman rank-order correlations.

Relative to existing research on skewness (Dunlap, Chen, & Greer, 1994; Dunlap, Burke, & Greer, 1995), the skewness levels in the current research appear to be weak (i.e., less than |1.5|) but this author is unaware of any known benchmarks for skewness. Transforming *strongly* skewed variables increased the correlation between the variables; however, results from transforming *weakly* skewed variables appeared inconsistent, increasing some correlations and decreasing others (Dunlap, Chen, & Greer, 1994). Even if transforming weakly skewed variables yielded consistent small increases in correlations, this benefit must be weighed against potential complications of interpreting results based on transformed variables (e.g., power or log transformations; Dunlap, Chen, & Greer, 1994; Dunlap, Burke, & Greer, 1995).

Watson asked a fundamental question that must be addressed in the development of any trait measure: "'Is this the optimal level of stability that can be achieved?'" (2004, p. 344). Clearly, the ISEE are less than optimal compared to existing measures of the same construct and in relation to conceptually related measures. Fortunately, a short-term retest or dependability study was conducted in the early stages of scale development; hopefully, the stability of the ISEE can be improved in subsequent stages of the scale development process. Future research should continue to examine distributions and skewness of the scale scores and items and the potential subtle influences of number of response options and the extremity of the wording of scale anchors. It is possible that changing the response scale (i.e., number of response options or extremity of the wording of scale anchors) could decrease skewness of the items and scales and could increase variability. Additionally, incorporating reverse-keyed items potentially could reduce skewness of scales—E and N both contain reverse-keyed items whereas ISEE and PANAS-X scales contain only positive-keyed items.

Strengths and Significance

The three-level hierarchical model of emotional expressivity integrated existing models, linked the general factor level to the discrete affect level of emotional expressivity, and integrated previous research based on measures at different levels of abstraction. The ISEE provide researchers with a brief scale of general emotional expressivity, two brief scales of positive and negative expressivity and multiple discrete affect scales. The PE and NE scales have improved discriminant validity and clearer differential relations within the nomological network in comparison to existing measures of positive and negative expressivity that contain too few items, have significant structural problems and poorer discriminant validity. Additionally, the creation of the ISEE with seven discrete affect scales—Fear, Hostility, Guilt/Shame, Sadness, Joviality, Confidence and Amusement—extends the literature on emotional expressivity. Although a recent structural analysis of emotional expressivity supported the discrete affect level, this research was based in the German language, consisted of rating only four adjectives to assess each discrete affect, contained a fewer number of discrete affects, and lacked any information about test-retest reliability (Trierweiler et al., 2002).

As a hierarchical set of scales, the lower level scales of the ISEE demonstrated both common variance and specific variance. The common variance among the lower level scales supported the factor at the next highest level; however, without also demonstrating specific variance at the lower levels then only the next higher level of the hierarchy is needed to assess the common variance. The recognition of the importance of demonstrating the common and specific variance at each level of the hierarchy shaped the development of this hierarchical assessment. If the lower level scales were simply added into composites then this would combine the common and specific variances of the lower level scales and thus add noise to the assessment of the common variance of the higher order factors. Alternatively, the items of the ISEE specifically target different levels of the hierarchy and brief scales were created to assess each level of the hierarchy.

To this author's knowledge, this is the first systematic attempt to account for the level of abstraction of the items to target specific constructs at different levels of the hierarchy. Even if the intention is only to create scales at the lowest order level, one must still carefully consider the level of abstraction of the items and match the level of abstraction of the item content with the level of the intended target construct. In other words, the failure to consider the level of abstraction of the items will create issues in the identification of the lower order structures. For instance, recent research has demonstrated that impulsivity is a multidimensional construct consisting of four factors: lack of planning, lack of perseverance, sensation seeking and urgency (Smith et al., 2007). Therefore assessments of impulsivity containing more abstract item content (e.g., "Sometimes I act impulsively") likely will obscure structural analyses aimed at identifying the lower order structure. More broadly, as the field continues to extend the lower order structure of personality below the Big Five, there will be an increased need to carefully consider the level of abstraction of the item content to create optimally valid assessments at lower order levels.

The ISEE allow for comparisons of various levels of abstraction of emotional expressivity. Examinations of the nomological network revealed striking differences between the relations of ISEE General scale compared to the relations of PE and NE. With a hierarchical set of scales, the relations between the various levels of the hierarchy can be examined with observed measures even if more sophisticated statistical techniques (e.g., structural equation modeling) for modeling latent higher order factors are unavailable to researchers. Alternatively, latent higher order factors could be modeled from discrete affect scales and then compared to results based on observed higher order scales.

Another strength of the current research was that the constructs were conceptualized and the scales were developed within a clear, well-articulated nomological network. As more evidence accumulates to support the higher order structure of personality, researchers should be sure to make connections within the existing nomological network to further elaborate the interconnections. Moreover, working in relation to existing structures further enhances the conceptualizations of the proposed constructs; that is, the recently proposed three-level hierarchical model of affect (Tellegen et al., 1999a, 1999b) was the impetus for creating the three-level hierarchical model of emotional expressivity. Additionally, comparing the various levels of abstraction of emotional expressivity in relation to the universal model of personality emphasized the relative independence of positive and negative expressivity.

Another strength of the current studies was that all samples consisted of more than a single monomethod assessment of college students. First, the sample used in the preliminary study contained nearly 400 newlyweds and incorporated both self- and spouse-ratings. Second, the pilot-testing of over 200 participants consisted of a combination of a small student sample and an online adult sample (e.g., relatives, friends, etc.) that was well represented across multiple age ranges. Third, the Student Sample incorporated a retest design with over 300 participants reassessed on the second occasion. Fourth, the ILPP Sample assessed more than 300 young adults and incorporated other-ratings from more than 200 knowledgeable informants (e.g., spouses, relatives, and friends). Altogether, this research represents responses from more than 1,500 participants.

Limitations

Although the inclusion of multiple discrete affects was a strength of the three-level hierarchical model and extended the assessment of emotional expressivity into the lowest level, the circumscribed number of discrete affects included in this model was a general limitation of this research. The proposed model did not attempt to model every possible discrete affect exhaustively. The primary rationale was to focus first on providing initial support for the proposed seven discrete affects because currently there is not a general consensus among researchers about the lower order of affect (Watson & Vaidya, 2003). The four negative discrete affects of fear, hostility, guilt, and sadness are included in other models of affect (Diener et al., 1995; Trierweiler et al., 2002). Other potential negative discrete affects could include: disgust, fatigue, frustration, mistrust/skepticism, self-doubt, shyness, and embarrassment; in addition, shame and guilt perhaps could be modeled and assessed separately (Watson, Chmielewski, Humrichouse, Naragon-Gainey, & Stasik, 2010).

There tends to be less consensus about the positive discrete affects. The three Positive Emotion scales of the PANAS-X (Watson & Clark, 1994) served as a starting point for the positive discrete affect scales in the current model: joviality, self-assurance, and attentiveness. However, in pilot-testing, attentiveness failed to form a distinct factor. Items containing the word *interest* demonstrated some potential for hanging together (AIC = .36) and this should be examined in future attempts to extend the number of discrete affects. Trierweiler et al. (2002) included a discrete affect scale of love that consisted of participants rating single word adjectives (i.e., love, affection, intimacy, and caring). Capturing the nonverbal expression of love in written items containing specific behaviors needs to be balanced against reducing the range of applicability of the scale by including more intimate behaviors (e.g., kissing, touch, etc). Other potential positive discrete affect could include: awe, surprise, hope, pride, gratitude, challenge/determination, and contentment/serenity (Tugade & Morrow, 2010; Ahrens & McIntosh, 2010).

In addition to the multiple discrete affects being both a strength and a limitation, incorporating the level of abstraction of the item into assessments at different levels of the hierarchical model can be considered both a strength and a limitation. More specifically, the creation of PE and NE as composites based on a subset of the discrete affect items could be criticized as not matching the items to the level of appropriate abstraction. Positive and negative expressivity could be assessed more directly with items containing such content as: positive, negative, good, bad, pleasant, unpleasant, up, down, on top of the world, uneasy, etc. However, in an attempt to model emotional expressivity after trait affect, a representative subset of the complete discrete affect items were selected to capture the second-order level of the hierarchy.

Another limitation is that the relations between scale scores on PE, NE and the discrete affect scales cannot be examined due to the presence of overlapping items; for instance, the correlation between PE and Joviality is artificially inflated due to the item, "If I'm feeling cheerful, people can see it in my smile," being scored for both scales. Future studies could test more abstract items for the second-order level or create a large enough discrete affect item pool to be able to have non-overlapping discrete affect items for both the second-order level and the discrete affect level.

Additionally, there are other potential limitations related to the item content. First, some item content is left ambiguous to incorporate expressions of less intense affective states, to capture less prototypical expressions, and to fit with the individual's conceptualizations of emotional expressivity (e.g., "When I'm excited, it shows"). In general, the current set of items included a mixture of more abstract items "If I'm feeling guilty, it shows" and items containing more specific nonverbal behaviors such as "If I'm feeling ashamed, I won't make eye contact." Alternatively, each item could incorporate a specific nonverbal expressive behavior and future research can continue to examine the validity of items as a function of different levels of abstractness.

Second, some item content attempted to emphasize the communicative function of emotional expressions and included ambiguous item content about other people being able to *tell* or *know* how one is feeling (e.g., "People can tell when I'm feeing sad"). These items became more ambiguous when changed to an other-rating format (e.g., "People can tell when X is feeling sad"). These types of items potentially could be interpreted to focus on other people's level of perceiving emotions in others; however, in the context of several other items focused on the target's level of emotional expressivity, it seems reasonable to interpret these items as tapping emotional expressivity. In future research, the validity of these types of items will continue to be examined.

Third, another potential limitation of the discrete affect items is the lack of reverse-keyed items. As the item content explicitly contained specific affect terms there is potential for the reverse-keyed items to be interpreted as emotional suppression or masking rather than a lack of emotional expressivity. For instance the item, "I rarely show others how I feel," is less likely to be interpreted as emotional suppression than an item such as "If I am feeling angry, it does not show." Based on the results from the preliminary study in Chapter 3, the item containing *suppress* had very low factor loadings. Additionally, in the structural analyses of several expressivity related questionnaires, the masking factor was only weakly related to the other factors of emotional expressivity (Gross & John, 1998). The potential lack of strong relations of reverse-keyed items needs to be balanced against the possibility that reverse-keyed items may be particularly beneficial for increasing variability and decreasing skewness.

Additionally, the instructions for the scale need to be modified slightly to more specifically indicate that the scale is assessing nonverbal emotional expressivity. The more abstract general expressivity items (e.g., "I am emotionally expressive") allow for the potential interpretation of more verbally mediated means of letting another know how one is feeling—such as telling someone you're feeling sad unaccompanied by any clear nonverbal behavior. More specifically, the current instructions for the scale (i.e., "The following statements relate to experiencing and expressing your feelings and emotions") could be adapted to emphasize nonverbal expressivity (i.e., "The following statements relate to experiencing and expressing your feelings and emotions through facial expressions, vocal characteristics (e.g., volume, tone, and rate of speech) and body language (e.g., posture and gestures)").

Across the various samples, the data collected through online questionnaires contained more missing data compared to the paper questionnaire packets. Several online participants had nearly complete data with one or two missing items that potentially resulted from clicking near the response without actually selecting the response. Future online data collection methods could incorporate a *back-button* to allow participants to review their answers or a *missed-item* screen at the end of the questionnaire containing all questions that the participant failed to answer. Although participants are free to skip any questions and requiring participants to respond to all questions would be unethical, providing participants with a second chance to respond to accidently skipped questions could reduce the amount of missing data in online data collection.

Future Directions

The development of the ISEE was based on analyses in two large samples; however, new and independent samples need to be collected for structural validation and to reassess various psychometric properties (e.g., coefficient alpha and AIC). The fit of this hierarchical model also needs to be tested through confirmatory factor analysis; however, both of the current samples were used to select items and therefore, the quantitative fit of the hierarchical model on either of these samples would be biased towards supporting the model. Additionally, future examinations of the fit of the hierarchical model could include various confirmatory factor models with increasing levels of difficulty for achieving an adequate fit. Confirmatory factor models based on (a) the discrete affect scales, (b) item parcels of the discrete affect scales, and (c) all discrete affect items could be tested. Furthermore, adding observed variables for the higher order factors could make it more difficult to obtain an adequate fit.

Additionally, MTMM analyses using confirmatory factor analysis have been developed to separate trait, method and error variance components (Eid et al., 2003). Although these models require identifying one method that the other methods then are compared against, the added benefit of estimating the effects of method and error variance allow for a more sophisticated examination of convergent and discriminant validity.

Future studies need to examine the predictive validity of the ISEE in relation to behavioral assessments of emotional expressivity. Existing measures of emotional expressivity have demonstrated predictive validity in relation to behavior assessments. For instance, the BEQ Positive and Negative Expressivity scales were able to differentially predict behavioral assessments of positive and negative expressions in a film-viewing paradigm (Gross & John, 1997). At a broader level of abstraction, the EES was related to an overall composite of expressivity but was not able to predict behavioral assessments of positive and negative expressivity (i.e., viewing film clips; Kring et al., 1994) consistently. Matching the levels of abstraction of the ISEE with the level of the behavioral predictors will need to be carefully considered and demonstrating the predictive validity at the discrete affect level will be particularly important. Generally, narrower traits that are specifically matched to specific criteria will increase the expected level of prediction.

With the improved discriminant validity of ISEE PE and NE and the extension of the ISEE discrete affects, future research can re-examine previous research findings for potential differential relations of positive and negative expressivity and the discrete affects. Emotional expressivity has demonstrated moderate positive relations with self-esteem (Gross & John, 1997; Kring et al., 1994) and emotionally expressive individuals reported higher levels of well-being and life satisfaction (King & Emmons, 1990; Kring et al., 1994). Regarding interpersonal relations, emotionally expressive individuals reported higher levels of social closeness and lower levels of social anhedonia (Kring et al., 1994). Emotional expressivity also has been examined in relation to aging (Coats et al., 2008; Gross et al., 1997), cultural differences (Gross & John, 1995), attachment (Armitage & Harris, 2006; Ducharme et al., 2002; Searle & Meara, 1999), and relationship satisfaction (Lavee & Ben-Ari, 2004).

The ISEE have several potential applications involving research on psychological disorders and medical problems. Emotional expressivity has been examined in relation to psychological disorders such as schizophrenia (Kring et al., 1993; Lee et al., 2008), depression (Kring et al., 1994; Sloan et al., 1997), social anxiety (Kashdan & Breen, 2008; Kashdan et al., 2007), post-traumatic stress disorder (Tull et al., 2007), and borderline personality disorder (Herpertz et al., 2001). Additionally, emotional expressivity has been examined in relation to medical problem such as cancer (Quartana et al., 2006; Zakowski et al., 2003), cardiovascular disease (Mendes et al., 2003; Shaw et al., 2003), Parkinson disease (Mikos et al., 2009; Simons et al., 2004), and asthma (Hollaender & Florin, 1983).

Additionally, to extend the nomological network, emotional expressivity will need to be examined in relation to the domain of affective empathy. The literature on empathy spans several fields and there are various definitions for this general construct (Preston & de Waal, 2002); however, affective empathy refers to vicariously feeling another's affective state. In cognitive psychology, Prinz (1992) proposed a radical view that a common coding system exists for both action and perception, such that the same codes for performing a behavior are involved when perceiving another's similar behavior. This link between action and perception has been mapped onto empathy in the Perception-Action Model (Preston & de Waal, 2002). If empathy is based on the link between action and perception, then this suggests that emotional expressivity (i.e., action) and affective empathy should be strongly related.

To examine the relation between emotional expressivity and affective empathy a hierarchical model of affective empathy could be developed. Similar to the hierarchical model of emotional expressivity, affective empathy could be modeled including a general factor of affective empathy and more specific factors of positive and negative affective empathy. Again, the level of abstraction of the items could be incorporated into a hierarchical assessment. For instance, general affective empathy items would be more abstract without valence content (e.g., “Seeing a person’s emotional expression can trigger the same emotion in me” and “I am generally unaffected by other people’s emotional expressions”). In addition, positive and negative affective empathy items would be more specific (e.g., “Seeing another person’s happiness makes me feel happy” for positive affective empathy and “Seeing fear on another’s face makes me feel a little fearful” for negative affective empathy).

Similar to establishing emotional expressivity within the nomological network, various levels of abstraction of affective empathy would be examined within the nomological network to support the validity of the different levels of abstraction. More specifically, the interrelations between positive expressivity, positive affective empathy, positive affect and extraversion—as well as the interrelations between negative expressivity, negative affective empathy, negative affect and neuroticism—would speak to the validity of the second-order level of positive and negative affective empathy.

Conclusion

The three-level hierarchical model of emotional expressivity integrated existing models within this domain. The development of the ISEE clarified assessments of the second-order level and extended the assessment of emotional expressivity to include the discrete affect level. Second-order factor analyses supported the hierarchical model; moreover, support for all three levels replicated across two large samples. The ISEE were examined within the nomological network of personality and affectivity, and PE and NE demonstrated clear differential relations with E and N, as well as incremental predictive validity beyond PA and NA. The ISEE also demonstrated convergent and discriminant validity with existing scales of emotional expressivity and through MTMM analyses of test-retest and self-other agreement data. In conclusion, this research provided support for the three-level hierarchical model and began an extended process of validating the ISEE; clearly, however, this research represents only a small part of the ongoing process of accumulating evidence for the construct validity of emotional expressivity.

REFERENCES

Ahrens A. H., & McIntosh, E. M. (2010, January). Distinguishing our pleasures: People vary in their differentiation of positive emotions. In M. M. Tugade & L. D. Kirby (Chairs), *The differentiation of positive emotions: Recent advances in theory and research on positive emotions*. Symposium conducted at the 11th annual meeting of the Society for Personality and Social Psychology, Las Vegas, Nevada.

Armitage, C. J., & Harris, P. R. (2006). The influence of adult attachment on symptom reporting: Testing a mediational model in a sample of the general population. *Psychology and Health, 21*, 351-366.

Benet-Martinez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology, 75*, 729-750.

Berry, D. S., & Pennebaker, J. W. (1998). Nonverbal and verbal emotional Expression and health. In G. A. Fava & H. Freyberger (Eds.), *Handbook of Psychosomatic Medicine* (pp. 69-83). Madison CT: International Universities Press, Inc.

Campbell D. T., & Fiske, D. (1959). Convergent and discriminant validity by the multi-trait, multi-method matrix. *Psychological Bulletin, 56*, 81-105.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16PF).* Champaign, IL: Institute for Personality and Ability Testing.

Chmielewski, M., & Watson, D. (2009). What is being assessed and why it matters: The impact of transient error on trait research. *Journal of Personality and Social Psychology, 97*, 186-202.

Clark, L. A. (1993). *Manual for the Schedule for Nonadaptive and Adaptive Personality.* Minneapolis: University of Minnesota Press.

Clark. L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309-319.

Coats, A. H., & Blanchard-Fields, F. (2008). Emotion regulation in interpersonal problems: The role of cognitive-emotional complexity, emotion regulation goals, and expressivity. *Psychology and Aging, 23*, 39-51.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment, 64*, 21–50.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

Diener, E., Smith, H. & Fujita, F. (1995). The personality structure of affect. *Journal of Personality and Social Psychology, 69*, 130-141.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246-1256.

Davidson, R. J. (1994). On emotion, mood and related affective constructs. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 51-55). New York: Oxford University Press.

Dobbs, J. L., Sloan, D. M., & Karpinski, A. (2007). A psychometric investigation of two self-report measures of emotional expressivity. *Personality and Individual Differences, 43*, 693-702.

Ducharme, J., Doyle, A. B., & Markiewicz, D. (2002). Attachment security with mother and father: Associations with adolescents' reports of interpersonal behavior with parents and peers. *Journal of Social and Personal Relationships, 19*, 203-231.

Dunlap, W. P., Burke, M. J., & Greer, T. (1994). The effect of skew on the magnitude of product-moment correlations. *The Journal of General Psychology, 122*, 365-377.

Dunlap, W. P., Chen, R., & Greer, T. (1995). Skew reduces test-test reliability. *The Journal of Applied Psychology, 79*, 310-313.

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effect from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological Methods, 8*, 38-60.

Ekman, P. (1992). Are there basic emotions? *Psychological Review, 99*, 550-553.

Everett, J. E. (1983). Factor comparability as a means of determining the number of factors and their rotation. *Multivariate Behavioral Research, 18*, 197–218.

Friedman, H. S., Prince, L. M., Riggio, R. E., & DiMatteo, M. R. (1980). Understanding and assessing nonverbal expressiveness: The Affect Communication Test. *Journal of Personality and Social Psychology, 39*, 333-351.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286–299.

Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality *Psychological Science, 2*, 31–39.

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55*, 149-158.

Greer, T., Dunlap, W. P., Hunter, S. T., & Berman, M. E. (2006). Skew and internal consistency, *Journal of Applied Psychology, 91*, 1351-1358.

Gross, J. J., Carstensen, L. L., Pasupathi, M., Tsai, J., Skorpen, C. G., & Hsu, A. Y. C. (1997). Emotion and aging: Experience, expression, and control. *Psychology and Aging, 12,* 590-599.

Gross, J. J., & John, O. P. (1995). Facets of emotional expressivity: Three self-report factors and their correlates. *Personality and Individual Differences, 19*, 555-568.

Gross, J. J., & John, O. P. (1997). Revealing Feelings: Facets of emotional expressivity in self-reports, peer ratings and behavior. *Journal of Personality and Social Psychology, 72*, 435-448.

Gross, J. J., & John, O. P. (1998). Mapping the domain of expressivity: Multimethod evidence for a hierarchical model. *Journal of Personality and Social Psychology, 74*, 170-191.

Hall, J. A., Murphy, N. A., & Mast, M. S. (2007). Nonverbal self-accuracy in interpersonal interaction. *Personality and Social Psychology Bulletin, 33*, 1675-1685.

Hampson, S. E., John, O. P., & Goldberg, L. R. (1986). Category breadth and hierarchical structure in personality: Studies of asymmetries in judgments of trait implications. *Journal of Personality and Social Psychology, 51*, 37–54.

Herpertz, S. C., Werth, U., Lukas, G., Qunaibi, M., Schuerkens, A., Kunert, H. J., Freese, R., Flesch, M., Mueller-Isberner, R., Osterheider, M., & Sass, H. (2001). Emotion in criminal offenders with psychopathy and borderline personality disorder. *Archives of General Psychiatry, 58*, 737-745.

Hogan, J., & Roberts, B. W. (1996). Issues and non-issues in the fidelity-bandwidth tradeoff. *Journal of Organizational Behavior, 17*, 627–637.

Hollaender, J., & Florin, I. (1983). Expressed emotion and airway conductance in children with bronchial asthma. *Journal of Psychosomatic Research, 27*, 307-311.

Izard, C. E. (1991). *The psychology of emotions.* New York: Plenum 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 (2nd Ed.)* (pp. 102-138). New York: Guilford.

John, O. P., Hampson, S. E., & Goldberg, L. R. (1991). The basic level in personality-trait hierarchies: Studies of trait use and accessibility in different contexts. *Journal of Personality and Social Psychology, 60*, 348–361.

Kashdan, T. B., & Breen, W. E. (2008). Social anxiety and positive emotions: A prospective examination of a self-regulatory model with tendencies to suppress or express emotions as a moderating variable. *Behavior Therapy, 39*, 1-12.

Kashdan, T. B., Volkmann, J. R., Breen, W. E., & Han, S. (2007). Social anxiety and romantic relationships: The costs and benefits of negative emotion expression are context-dependent. *Journal of Anxiety Disorders, 21*, 475-492.

Keltner, D., & Gross, J. J. (1999). Functional accounts of emotions. *Cognition and Emotion, 13,* 467–480.

Kenny, D. A. (1987). *Statistics for the social and behavioral sciences.* Boston: Little, Brown.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York: Guilford Press.

King, L. A., & Emmons, R. A. (1990). Conflict over emotional expression: Psychological and physical correlates. *Journal of Personality and Social Psychology, 58,* 864-877.

Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: Expression, experience, and physiology. *Journal of Personality and Social Psychology, 74,* 686-703.

Kring, A. M., Kerr, S., Smith, D.A., & Neale, J.M. (1993). Flat affect in schizophrenia does not reflect diminished subjective experience of emotion. *Journal of Abnormal Psychology, 104*, 505-517.

Kring, A. M., Smith, D. A., & Neale, J.M. (1994). Individual differences in dispositional expressiveness: Development and validation of the Emotional Expressivity Scale. *Journal of Personality and Social Psychology, 66*, 934-949.

Larsen, R. J., & Diener, E. (1987). Affect Intensity as an individual difference characteristic: A review. *Journal of Research in Personality, 21*, 1-39.

Lavee, Y., & Ben-Ari, A. (2004). Emotional expressiveness and neuroticism: Do they predict marital quality? *Journal of Family Psychology, 18*, 620-627.

Lee, S. J., Yoo, S. Y., Kang, D. H., Lee, K. J., Ha, T. H., Wee, W., Lee, A. R., Kim, N. S., & Kwon, J. S. (2008). Potential vulnerability markers within the affective domain in subjects at genetic and clinical high risk for schizophrenia. *Psychopathology, 41*, 236-244.

Levenson, R.W. (1994). Human emotion: A functional view. In P. Ekman & R.J. Davidson (Eds.), *The nature of emotion* (pp. 123–126). New York: Oxford University Press.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3*, 635–694.

Longley, S. L., Watson, D., & Noyes, R. Jr. (2005). Assessment of the hypochondriasis domain: The Multidimensional Inventory of Hypochondriacal Traits (MIHT). *Psychological Assessment, 17*, 3-14.

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88*, 139-157.

Mendes, W. B., Reis, H. T., Seery, M. D., & Blascovich, J. (2003). Cardiovascular correlates of emotional expression and suppression: Do content and gender context matter? *Journal of Personality and Social Psychology, 84*, 771-792.

Meng, X., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111*, 172-175.

Mikos, A. E., Springer, U. S., Nisenzon, A. N., Kellison, I. L., Fernandez, H. H., Okun, M. S., & Bowers, D. (2009). Awareness of Expressivity Deficits in Non-Demented Parkinson Disease. *Clinical Neuropsychologist, 23,* 805-817.

Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology, 4*, 681-691.

O'Connor, B. P. (2000a). SPSS and SAS programs for determining the number of components using parallel and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*, 396-402.

O'Connor, B. P. (2000b). map.sas [Computer program]. Retrieved March 12, 2010, from https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html.

O'Connor, B. P. (2000c). parallel.sas [Computer program]. Retrieved March 12, 2010, from https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html.

O'Connor, B. P. (2000d). rawpar.sas [Computer program]. Retrieved March 12, 2010, from https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html.

Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63*, 816-824.

Plutchik, R. (1990). Emotions and psychotherapy: A psychoevolutionary perspective. In R. Plutchik & H. Kellerman (Eds.) *Emotion: Theory, research, and experience Vol. 5* (pp. 3–41). New York: Academic Press.

Preston S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*, 1-72.

Prinz, W. (1992). Why don't we perceive our brain states? *European Journal of Cognitive Psychology, 4*, 1-20.

Quartana, P. J., Laubmeier, K. K., & Zakowski, S. G. (2006). Psychological adjustment following diagnosis and treatment of cancer: An examination of the moderating role of positive and negative emotional expressivity. *Journal of Behavioral Medicine, 29,* 487-498.

Ready, R. E., Clark, L. A., Watson, D., & Westerhouse, K. (2000). Self- and peer-reported personality: Agreement, trait ratability, and the "self-based heuristic". *Journal of Research in Personality, 34*, 208-224.

Riggio, H. R., & Riggio, R. E. (2002). Emotional expressiveness, extraversion, and neuroticism: A meta-analysis. *Journal of Nonverbal Behavior, 26*, 195-218.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 2–25.

Searle, B., & Meara, N. M. (1999). Affective dimensions of attachment styles: Exploring self-reported attachment style, gender, and emotional experience among college students. *Journal of Counseling Psychology, 46*, 147-158.

Seim, R. W., & Spates, C. R. (2010). The prevalence and comorbidity of specific phobias in college students and their interest in receiving treatment. *Journal of College Student Psychotherapy, 24*, 49-58.

Shaw, W. S., Patterson, T. L., Semple, S. J., Dimsdale, J. E., Ziegler, M. G., & Grant, I. (2003). Emotional expressiveness, hostility and blood pressure in a longitudinal cohort of Alzheimer caregivers. *Journal of Psychosomatic Research, 54*, 293-302.

Simons, G., Pasqualini, M. C. S., Reddy, V., & Wood, J. (2004). Emotional and nonemotional facial expressions in people with Parkinson's disease. *Journal of the International Neuropsychological Society, 10*, 521-535.

Sloan, D. M., Strauss, M. E., Quirk, S. W., & Sajatovic, M. (1997). Subjective and expressive emotional responses in depression. *Journal of Affective Disorders, 46*, 135-141.

Smith, G. T., Fischer, S., Cyders, M. A., Annus, A. M., Spillane, N. S., & McCarthy, D. M. (2007). On the validity and utility of discriminating among impulsivity-like traits. *Assessment, 14*, 155-170.

Snyder, M. (1974). The self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30*, 526-537.

Schweizer, K. (1992). A correlation-based decision-rule for determining the number of clusters and its efficiency in uni- and multi-level data. *Multivariate Behavioral Research, 27*, 77-94.

Tellegen, A., Watson, D., & Clark, L. A. (1999a). On the dimensional and hierarchical structure of affect. *Psychological Science, 10*, 297-303.

Tellegen, A., Watson, D., & Clark, L. A. (1999b). Further support for a hierarchical model of affect: Reply to Green and Salovey. *Psychological Science, 10*, 307-309.

Trierweiler, L. I., Eid, M., & Lischetzke, T. (2002). The structure of emotional expressivity each emotion counts. *Journal of Personality and Social Psychology, 82,* 1023-1040.

Tugade, M. M., & Morrow J. (2010, January). Where positive emotions meet and where they diverge: The effects of distinct positive emotions on coping and problem solving. In M. M. Tugade & L. D. Kirby (Chairs), *The differentiation of positive emotions: Recent advances in theory and research on positive emotions*. Symposium conducted at the 11th annual meeting of the Society for Personality and Social Psychology, Las Vegas, Nevada.

Tull, M. T., Jakupcak, M., Paulson, A., & Gratz, K. L. (2007). The role of emotional inexpressivity and experiential avoidance in the relationship between posttraumatic stress disorder symptom severity and aggressive behavior among men exposed to interpersonal violence. *Anxiety Stress and Coping, 20*, 337-351.

Turner, N. E. (1998). The effect of common variance and structure pattern on random data eigenvalues: Implications for the accuracy of parallel analysis. *Educational and Psychological Measurement, 58*, 547-568.

Urbina, S. (2004). *Essentials of psychological testing.* Hoboken, NJ: John Wiley & Sons Inc.

Vaidya, J. G., Gray, E. G., Haig, J. R., Mroczek, D. K., & Watson, D. (2008). Differential stability and individual growth trajectories of big five and affective traits during young adulthood. *Journal of Personality 76*, 267–304.

Vaidya, J. G., Gray, E. G., Haig, J. R., & Watson, D. (2002). On the temporal stability of Personality: Evidence for differential stability and the role of life experiences. *Journal of Personality and Social Psychology, 83*, 1469-1484.

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment* (pp. 41-71). Boston: Kluwer.

Watson, D. (2000). *Mood and temperament.* New York: Guilford Press.

Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality, 38,* 319-350.

Watson, D., Chmielewski, M., Humrichouse, J., Naragon-Gainey, K., & Stasik, S. (2010). [Trait Affect Project]. Unpublished raw data.

Watson, D., & Clark, L. A. (1991). Self- versus peer-ratings of specific emotional traits: Evidence of convergent and discriminant validity. *Journal of Personality and Social Psychology, 60*, 927-940.

Watson, D., & Clark, L. A. (1999). The PANAS-X: Manual for the Positive and Negative Affect Schedule-Expanded Form. Retrieved from University of Iowa, Department of Psychology Web site: http://www.psychology.uiowa.edu/Faculty/Watson/Watson.html.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology,* 54, 1063-1070.

Watson, D., Hubbard, B., & Weise, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology, 78*, 546-558.

Watson, D., & Humrichouse, J. (2006). Personality development in emerging adulthood: Integrating evidence from self-ratings and spouse-ratings. *Journal of Personality and Social Psychology, 91*, 959–974.

Watson, D., O'Hara, M. W., Simms, L. J., Kotov, R., Chmielewski, M., McDade-Montez, E. A., Gamez, W., & Stuart, S. (2007). Development and validation of the Inventory of Depression and Anxiety Symptoms (IDAS). *Psychological Assessment, 19*, 253-268.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219-235.

Watson, D., & Tellegen, A. (1999). Issues in the dimensional structure of affect—effects of descriptors, measurement error and response formats: Comment on Russell and Carroll (1999). *Psychological Bulletin, 125*, 601-610.

Watson, D., & Vaidya, J. (2003). Mood measurement: Current status and future directions. In J. A. Schinka & W. Velicer (Eds.), *Comprehensive handbook of psychology. Volume 2: Research Methods* (pp. 351-375). New York: Wiley.

Watson, D., Wiese, D., Vaidya, J. & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology, 76*, 820-838.

Zakowski, S. G., Harris, C., Krueger, N., Laubmeier, K. K., Garrett, S., Flanigan, R., & Johnson, P. (2003). Social barriers to emotional expression and their relations to distress in male and female cancer patients. *British Journal of Health Psychology, 8,* 271-286.