

Preliminary Development and Psychometric Evaluation of the ACL Scale: Factor Structure,  
Measurement Invariance, Reliability, and Validity

Adam M. Kuczynski

A thesis

submitted in partial fulfillment of the  
requirements for the degree of

Master of Science

University of Washington

2016

Committee:

Robert J. Kohlenberg

Jonathan W. Kanter

Brian P. Flaherty

Program Authorized to Offer Degree:

Psychology

© 2016

Adam M. Kuczynski

University of Washington

**Abstract**

Preliminary Development and Psychometric Evaluation of the ACL Scale: Factor Structure, Measurement Invariance, Reliability, and Validity

Adam M. Kuczynski

Chair of Supervisory Committee:  
Robert J. Kohlenberg, Ph.D.  
Psychology

Humans have evolved such that social connectedness (i.e., interpersonal closeness, belongingness) plays an integral role in the health, happiness, and survival of our species. Intimacy has been the target of much research over the past few decades in the effort to gain a more nuanced, behavioral understanding of social connectedness and, in particular, points amenable to intervention. Despite the importance of this pursuit, however, few self-report measures exist that capture theoretically and empirically grounded aspects of intimacy, and those that do exist have important shortcomings that limit their utility in research. The purpose of this study was to create a self report measure of intimacy that has a strong theoretical and empirical foundation, is applicable to a broad population of individuals across a broad range of contexts, is useful in a clinical setting, and is situated in the behavioral psychological tradition. The authors report on the development and preliminary psychometric evaluation of the ACL Scale, a 20-item, 4-factor self-report measure of intimacy based on Reis and Shaver's (1988) model. Factor structure, measurement invariance, internal and test-retest reliability, and validity of the measure are discussed.

Preliminary Development and Psychometric Evaluation of the ACL Scale: Factor Structure,  
Measurement Invariance, Reliability, and Validity

### **Introduction**

Human beings have evolved such that social connectedness (i.e., interpersonal closeness, belongingness) plays an integral role in the health, happiness, and survival of our species (Baumeister & Leary, 1995; Brewer, 2004). One meta-analytic review, for example, found that strong social relationships are associated with a 50% increase in the odds of survival when compared to major public health concerns such as tobacco use, obesity, and alcohol consumption (Holt-Lunstad, Smith, & Layton, 2010). Moreover, longitudinal and cross-sectional research has shown that social connectedness predicts cardiovascular health (Caspi, Harrington, Moffitt, Milne, & Poulton, 2006; Thurston & Kubzansky, 2009), immune functioning (Lutgendorf et al., 2005), neuroendocrine functioning (Grewen, Girdler, Amico, & Light, 2005), and general physiological functioning (see Hawkley & Cacioppo, 2013 for an extensive review). In addition to impacting physical health, feelings of isolation may also be associated with increases in depressive symptomatology (Cacioppo, Hughes, Waite, Hawkley, & Thisted, 2006), suicidal ideation (Van Orden et al., 2008), negative mood and anxiety (Cacioppo et al., 2006), and general psychological distress (Lee, Draper, & Lee, 2001). Lastly, social connectedness has been shown to predict other important psychological and societal factors such as empathy (Beadle, Brown, Keady, Tranel, & Paradiso, 2012), violence (Thomas & Smith, 2004), acculturative stress (Yeh & Inose, 2003), and prosocial behavior (Pavey, Greitemeyer, & Sparks, 2011).

### **Intimacy**

Social connectedness researchers have focused much attention on the concept of intimacy as it relates to feelings of belongingness, social isolation/loneliness, psychopathology, and happiness. For example, using an early version of ecological momentary assessment methodology (Stone & Shiffman, 1994), Wheeler, Reis, and Nezlek (1983) found that the presence of intimacy in subjects' daily

interactions strongly predicted subsequent levels of loneliness. Furthermore, using structural equation modeling to examine the complex relationships between social factors and well-being, Harvey & Bray (1991) found that peer-to-peer intimacy is strongly associated with lower psychological distress and life stress, as well as more frequent engagement in health-enhancing behaviors (e.g., exercise and fitness, stress control). Research by Buhrmester (1990) suggests that peer-to-peer intimacy may play an important developmental role in socioemotional adjustment as well as interpersonal competency in adolescents.

Nearly two decades ago, Reis and Shaver (1998) introduced a theoretical model for the development of intimacy that has since garnered a good deal of empirical support (e.g., Aron, Melinat, Aron, Vallone, & Bator, 1997; Collins & Miller, 1994; Haworth et al., 2015; Laurenceau, Barrett, & Pietromonaco, 1998; Reis, Sheldon, Gable, Roscoe, & Ryan, 2000). The transactional process begins when one member of a social interaction (i.e., the speaker) engages in a vulnerable self-disclosure – an expression of “personally revealing feelings or information...” (Reis & Shaver, 1988, p. 367) – with another individual (i.e., the listener). The process continues when the listener provides the speaker with an empathically attuned response that functions to make the speaker feel understood, validated, and cared for (referred to as “responsiveness” by the authors; Reis & Shaver, 1988, p. 375). If the listener is met with repeated success on this front, the model suggests that the speaker will continue to engage in vulnerable self-disclosures *ad infinitum*. The model is simplified for the sake of presentation by making the role of each individual static, however as Reis and Shaver (1988) note, the process applies to the frequent situation in which both members of a dyad act as the listener and speaker interchangeably.

### **Functional Analytic Psychotherapy**

Functional Analytic Psychotherapy (FAP; Kohlenberg & Tsai, 1991; Tsai et al., 2009) is a contextual-behavioral (Biglan & Hayes, 1997) therapy that is uniquely equipped to target client interpersonal problems, including deficits in intimacy. Based on research (see, e.g., Skinner, 1958) that

suggests the most effective intervention strategies involve consequating (i.e. reinforcing, punishing) behavior close in time to its emission, FAP therapists focus on client behavior as it occurs in-session (as opposed to many other intervention strategies which focus much attention on out-of-session events). Essential to the theoretical foundation of FAP is that this behavior is inherently embedded within a the client-therapist relationship. Thus, although the FAP therapist can target any behavior as long as it is subject to interpersonal consequences, there is a clear link to targeting deficits in intimacy.

FAP posits a model intimacy (i.e., the ACL model) that that is not unlike the process proposed by Reis and Shaver (1988). Intimacy, according to this model, is composed of three processes that occur within an interaction: (1) awareness, (2) courage, and (3) love. Awareness involves focusing one's attention on *intrapersonal* factors that are relevant to the interaction, including physiological sensations, thoughts, reactions, and values. Awareness also includes attending to the internal experiences of the other individual in the interaction, such as thoughts and sensations they may be experiencing as well as their reactions and values within the interaction. Courage involves taking interpersonal risks in the service of improving the relationship, which includes authentically sharing one's innermost thoughts and feelings as well as encouraging the other individual to do the same. Lastly, love involves responding to the other's courage in a way that is reinforcing (i.e., that increases the likelihood of reemission of the behavior) as well as attending to one's own needs when their courage is not met with love in response. Thus, each of the three components is further divided into a self- and other- component. See "Definition, Empirical Foundations and Hypothesized Factor Structure" section below for discussion of the intersection between the FAP model and that proposed by Reis and Shaver (1988).

### **Extant Measures in the Social Connectedness Literature**

The depth and breadth with which social connectedness affects our daily lives presents unique challenges to the development of content valid (i.e., measuring the full range of the construct),

construct valid (i.e., measuring social connectedness as opposed to related, but distinct, constructs), and reliable (i.e., consistent measurement) assessments tools. Due to these difficulties, researchers and psychometricians have developed a variety of measures, each of which captures a particular component of our complex, multifaceted social lives.

A large source of variation among extant social connectedness assessment tools involves the type of relationship with which each was designed to measure. Some tools (e.g., Grote & Frieze, 1994; Lopez & Rice, 2006), for example, were designed to capture intimacy in the context of romantic relationships while others (e.g., Liang et al., 2002) were designed explicitly for non-romantic relationships. Further variability is introduced with respect to the breadth of each measure: while some scales (e.g., Miller & Lefcourt, 1982; Schafer & Olson, 1987) prompt the respondent to answer based on a singular relationship (e.g., husband/wife, friend, family member), others (e.g., Lee et al., 2001) are concerned only with the respondent's general feelings of connectedness with their broader social environment.

Another source of variation among measures involves the way in which social connectedness is conceptualized and operationalized. Many existing tools (e.g., Berscheid, Snyder, & Omoto, 1989; Leonard et al., 2014) measure social connectedness as the presence or absence of behavioral indicators, suggesting that it is something we *do* rather than *have*. Others (e.g., Lee et al., 2001; Miller & Lefcourt, 1982) take a different approach and measure connectedness as a collection of feelings and attitudes; within this latter group of measures is ambiguity with respect to how the construct is treated – as a state variable that is free to vary across time and context or a trait variable that varies only by individual.

Several self-report measures exist in domains related to social connectedness that researchers may be interested in tapping into. For example, much work has focused on the construct of loneliness and isolation (see, for example, Holt-Lunstad, Smith, Baker, Harris, & Stephenson, 2015 and Masi, Chen, Hawkey, & Cacioppo, 2010 for a review). The most cited self-report measures in this domain

capture general feelings of loneliness as a psychological state in adults (Russell, 1996), feelings of loneliness and social dissatisfaction as both state and trait variables in children (Asher, Hymel, & Renshaw, 1984), and perceived level of social support in four functional domains (Sherbourne & Stewart, 1991) and social groups (Zimet, Dahlem, Zimet, & Farley, 1988). Other measures exist that capture constructs important to social connectedness research such as theory of mind (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), emotional intelligence (Schutte et al., 1998), empathic attunement (Davis, 1980), social self-esteem (Lawson, Marshall, & McGrath, 1979), and trust (Larzelere & Huston, 1980).

### **The Need for a New Measure in the Connectedness Literature**

The heterogeneity with which social connectedness is currently measured (i.e., idiosyncratic measures, differential construct conceptualization) has limited the generalizability of much research in the field. As mentioned above, the breadth and depth with which social functioning affects our lives as human beings has made it difficult for psychometricians to develop a brief self-report measure that captures the multifaceted nature of social connectedness (i.e., that is content valid). As such, much intervention research has focused on narrowly defined subpopulations that, although important, lack relevance to the broader population of individuals. Given the implications of social connectedness on our physical and psychological well-being, it is important for intervention researchers to have access to the larger population of those in need.

In addition, while the variability with which existing measures conceptualize social connectedness (e.g., as a trait versus a state) has allowed researchers to examine more nuanced research questions, it has also contributed to the ambiguity with which the construct as a whole is conceptualized. Thus, it is unclear whether researchers who intend to measure the same construct across studies are actually doing so. Ambiguity in the way social connectedness is conceptualized has also led many (e.g., Ashida & Heaney, 2008; Grieve, Indian, Witteveen, Tolan, & Marrington, 2013) to

tailor existing measures to their specific research question, or create new measures altogether. While this may provide the researcher with more precision, and thus more statistical power, these measures are often included with little-to-no psychometric validation and thus aid in furthering the uncertainty with which social connectedness is understood. Given the replication crisis currently facing the field of psychology as a whole (Open Science Collaboration, 2015), attention to measurement precision is warranted.

Considering the aforementioned measurement deficits, there is a need for an assessment tool that captures aspects of social connectedness relevant across populations/relationship types and that focuses on the mechanisms responsible for feelings of connectedness. The process of intimacy, as described by Reis and Shaver (1988), may provide the theoretical foundation needed for such a measure. Therefore, in this study we aimed to develop a self-report questionnaire that (1) integrates the most robust findings from the intimacy process literature, (2) has clinical utility (i.e., accurately measures intimacy-related outcomes over the course of psychotherapy, (3) contains items that are grounded in a behavioral conceptualization of intimacy, and (4) sensitively tracks the core mechanisms responsible for intimacy vis-à-vis a broad population of individuals. The current paper describes the development and psychometric evaluation of the ACL Scale, including factor structure, measurement invariance, internal and test-retest reliability, and convergent/discriminant validity.

### **Definition, Empirical Foundations and Hypothesized Factor Structure**

Given our focus on creating a tool that is applicable to a broad population of individuals and contexts (i.e. relationship types), and that has strong empirical and theoretical roots, we designed the current measure such that it would be sensitive to each component of Reis and Shaver's (1988) process model. We hypothesized four correlated factors:

#### **Other-Awareness.**

In order for Reis and Shaver's (1988) process to unfold, each member of a dyad must be exquisitely attuned to their partner's intimacy generative behavior as well as to the context in which the interaction is taking place. Most notably, this involves each partner's awareness of moments in which their vulnerable self-disclosure has a high likelihood of being reciprocated with an understanding, validating, and caring response. Moreover, because recent experimental research (Haworth et al., 2015) suggests that vulnerable expression *per se* is insufficient in generating feelings of intimacy – and that the listener's response is a necessary component of the model – we included items that target the respondent's awareness of opportunities to provide their partner with an empathically attuned response as outlined by the model.

#### **Self-Awareness.**

In addition to discriminating *when* it is appropriate to self-disclose in an interaction, it is necessary that the speaker knows *what* to disclose. Using ecological momentary assessment methodology to directly examine Reis and Shaver's (1988) model, Laurenceau et al. (1998) found that self-disclosure of emotion is a stronger predictor of intimacy than self-disclosure of information and facts. In light of these findings, we developed items aimed at capturing the respondent's awareness of what constitutes a *vulnerable* self-disclosure – idiosyncratically defined – versus what does not. It is also important for the each partner, *vis-à-vis* their involvement as both the speaker and listener, to be attuned to their own reactions and feelings as they are evoked by the speaker. Thus, the measure developed in this study includes items that capture the respondent's self-awareness in these domains.

#### **Courage.**

Cordova and Scott (2001) operationalize vulnerability as an interpersonal process involving an increased likelihood of punishment (defined behaviorally; see Skinner, 1953, Chapter 12). Continuing to conceptualize Reis and Shaver's (1988) model through a behavioral lens, they suggest that intimacy is built when one's vulnerability is met not with punishment, but instead with a reinforcing response

(i.e., a response that functions to increase the likelihood of the antecedent behavior; in Reis and Shaver's [1988] terminology, a response that leaves the speaker feeling validated, understood, and cared for). This understanding suggests that engaging in a vulnerable self-disclosure is inherently risky and thus takes courage on behalf of the speaker. Therefore, we developed items aimed at measuring the respondent's behavioral inclination to engage in interpersonal risk-taking. Furthermore, considering the positive association between values-consistent behavior, social functioning, relationship satisfaction, and mental health (Wilson, Sandoz, & Kitchens, 2010), several items were developed to capture risk-taking in the context of the respondent's values (interpersonal or otherwise).

### **Love.**

Although Reis and Shaver's (1988) model is situated in a verbal exchange process, nowhere do the authors claim that it must necessarily be so. In light of this, and consistent with functional-contextual theories of human behavior (Biglan & Hayes, 1997), we designed the items in this factor to capture the respondent's inclination to behave in ways that are *functionally* identical to the listener's role as prescribed by the model. Therefore, some items capture the respondent's tendency to engage in verbal responsiveness while others measure non-verbal behavior that leaves the speaker feeling understood, validated, and cared for.

## **Method**

### **Item Development**

Items for the ACL Scale were written collaboratively by a team of clinical psychologists with years of experience in social connectedness research and clinical psychological research methodology. The item generation process was initially guided by the ACL model and resulted in 96-items that mapped onto the self- and other- components of awareness, courage, and love. After preliminary analyses (see "preliminary analyses" section below for more information), however, we removed the other-courage and self-love items in an effort to ground the model in Reis and Shaver's (1988)

theoretical structure. We also removed items that we deemed unfit for a self-report measure (i.e., double-barreled wording). Thus, the item pool from which final analyses were conducted contained 64 items mapping onto four factors: other-awareness, self-awareness, courage, and love.

### **Participants**

The total study sample consisted of 1,804 subjects recruited from five universities across the globe. Four of the five universities are located in the United States (University of Washington, University of Louisville, Utah State University, and Bowling Green State University) while the fifth (Curtin University) is located in Perth, Australia (these samples are henceforth referred to as Sample 1, 2, 3, 4, and 5 and 6 [see below], respectively). The majority of subjects ( $n = 1,547$ ) were college undergraduates participating in exchange for extra course credit, while the remaining subjects ( $n = 257$ ; all from the Curtin University sample) were members of the general public recruited via online advertisements (Sample 6). The sample size breakdown is as follows: University of Washington ( $n = 687$ ; Sample 1), University of Louisville ( $n = 465$ ; Sample 2), Utah State University ( $n = 150$ ; Sample 3), Bowling Green State University ( $n = 62$ ; Sample 4), Curtin University undergraduates ( $n = 183$ ; Sample 5), and Curtin University general public ( $n = 257$ ; Sample 6). Subjects in Samples 1, 2, and 4 were invited to take a shorter survey one-week after the first for test-retest purposes, resulting in a sample of 687 subjects at follow-up.

In order to participate in the study, subjects had to be a college undergraduate (with the exception of Sample 6), age 18 or older, and fluent in English. The majority of subjects were female (65.19%), caucasian (42.57%), and had a mean age of 22.47 years ( $SD = 8.08$ ). See Table 1 for a detailed overview of demographic characteristics for each sample.

### **Measures**

In order to assess convergent and discriminant validity of the final ACL Scale, we used the following measures in this study (due to limited resources, subjects in Samples 5 and 6 were not

administered these additional measures). Our predictions follow Cohen's (1988) criteria for small ( $r < .3$ ), medium ( $.3 \leq r < .5$ ) and large ( $r \geq .5$ ), effect sizes.

### **Total scale correlations.**

***Social connectedness.*** The Social Connectedness Scale (SCS; Lee & Robbins, 1995) is an 8-item self-report measure of perceived social connectedness. Respondents rate the degree to which they agree with each self-descriptive statement on a Likert scale ranging from 1 ("Agree") to 6 ("Disagree"). All items on the SCS are negatively worded such that higher scores represent greater levels of perceived social connectedness. The SCS had strong internal reliability in the current study (Cronbach's  $\alpha = .96$ ), with total scale scores ranging from 8 to 48 ( $M = 36.84$ ,  $SD = 10.03$ ).

***Loneliness.*** The UCLA Loneliness Scale (UCLALS; Russell, 1996) is a 20-item self-report measure of perceived loneliness and isolation. Respondents rate frequency with which they experience feelings of isolation on a 4-point Likert scale ranging from 1 ("Never") to 4 ("Always"). Nine of the 20 items are reverse coded prior to scoring such that higher scale scores indicate greater levels of loneliness. The UCLALS had strong internal reliability in the current study (Cronbach's  $\alpha_{\text{Time1}} = .93$ ,  $\alpha_{\text{Time2}} = .94$ ), with total scale scores ranging from 20 to 80 at both time points (Time 1:  $M = 45.09$ ,  $SD = 10.91$ ; Time 2:  $M = 42.94$ ,  $SD = 11.08$ ).

***Emotional intelligence.*** The Schutte Self-Report Emotional Intelligence Test (SSEIT; Schutte et al., 1998) is a 33-item self-report measure of emotional intelligence based theoretically on Salovey and Mayer's (1990) conceptualization which emphasizes expression, appraisal, regulation, and use of emotions in oneself and others. Respondents rate the extent to which each item applies to them on a 5-point Likert scale ranging from 1 ("Strongly Disagree") to 5 ("Strongly Agree") Three of the 33 questions

are reverse coded prior to scoring such that higher scale scores indicate greater emotional intelligence. The SSEIT had strong internal reliability in the current study (Cronbach's  $\alpha_{\text{Time1}} = .92$ ,  $\alpha_{\text{Time2}} = .93$ ), with total scale scores ranging from 43.2 to 165 (Time 1;  $M = 120.99$ ,  $SD = 16.32$ ) and 62.0 to 165 (Time 2;  $M = 120.98$ ,  $SD = 16.01$ ).

**Social support.** The MOS Social Support Survey (SSS; Sherbourne & Stewart, 1991, developed for a large, longitudinal assessment of patients with chronic conditions (the Medical Outcomes Study), is a 20-item multidimensional self-report measure that emphasizes emotional/informational support, affectionate support, tangible support, and positive social interaction. Respondents rate the frequency with which they receive support in each domain on a 5-point Likert scale ranging from 1 (“None of the time”) to 5 (“All of the time”) such that higher scores indicate greater levels of social support. The SSS had strong internal reliability in the current study (Total: Cronbach's  $\alpha_{\text{Time1}} = .96$ ,  $\alpha_{\text{Time2}} = .97$ ; Subscales:  $\alpha_{\text{Time1}} = .89-.94$ ,  $\alpha_{\text{Time2}} = .89-.95$ ), with total scale scores ranging from 0 to 100 (Time 1;  $M = 72.28$ ,  $SD = 20.90$ ) and 7.89 to 100 (Time 2;  $M = 70.99$ ,  $SD = 20.75$ ).

**Intimacy-related behaviors.** The Functional Analytic Psychotherapy Intimacy Scale (FAPIS; Leonard et al., 2014) is a 14-item, multidimensional self-report measure of engagement in intimacy-related behaviors vis-à-vis a particular person (chosen by respondent; e.g., friend, family member, romantic partner) within the past week. Respondents rate the degree to which each statement accurately describes their behavior on a 7-point Likert scale ranging from 0 (“Not at all”) to 6 (“Completely”). The FAPIS contains three factors – (1) hidden thoughts and feelings, (2) expression of positive feelings, and (3) honesty and genuineness – that each contributes to a total scale score. Four of the 14 items are reverse coded prior to scoring such that higher scores represent

a greater degree of engagement in intimacy-related behaviors. The FAPIS had strong internal reliability in the current study (Total: Cronbach's  $\alpha = .92$ ; Subscales:  $\alpha = .87-.91$ ), with total scale scores ranging from 0 to 84 ( $M = 61.74$ ,  $SD = 15.18$ ).

**Intimacy.** The Miller Social Intimacy Scale (SIS; Miller & Lefcourt, 1982) is a 17-item self report measure of the "maximum level of intimacy currently experienced" in a particular relationship. In the current study, only those subjects who indicated being in a romantic relationship ( $n = 554$ ; Samples 1-4) were administered the scale, and each responded to the items with respect to their romantic partner. Respondents rate the accuracy with which each item describes their relationship and feelings on a 10-point Likert scale ranging from 1 ("Very Rarely"/"Not Much") to 10 ("Almost Always"/"A Great Deal"). Two of the 17 items are reverse coded prior to scoring such that higher scores indicate a greater level of intimacy within the relationship. The SIS demonstrated strong internal reliability in the current study (Cronbach's  $\alpha = .93$ ), with scores ranging from 59 to 170 ( $M = 143.72$ ,  $SD = 21.44$ ).

#### **Subscale correlations.**

**Empathy.** The Interpersonal Reactivity Index (IRI; Davis, 1983) is a 28-item multidimensional self-report measure of empathy defined by Davis (1983) as "reactions of one individual to the observed experiences of another" (p. 113). The four factors – (1) PT: Perspective taking, (2) FS: Fantasy, (3) EC: Empathic Concern, and (4) PD: Personal Distress – each contain an equal number of items that contribute to a subscale score but not a total scale score. Davis (1983) defines the four factors, respectively as: (1) "the tendency to spontaneously adopt the psychological point of view of others," (2) the tendency to "transpose [oneself] imaginatively into the feelings and actions of fictitious characters...", (3) "other-oriented feelings of sympathy and concern for

unfortunate others,” and (4) “self-oriented feelings of personal anxiety and unease in tense interpersonal settings” (pp. 113-114). Respondents rate the degree to which each item describes them on a 5-point Likert scale ranging from 0 (“Does not describe me”) to 4 (“Describes me very well”). Nine of the 28 items are reverse coded prior to scoring such that higher scores indicate a greater empathy as measured by each subscale. The IRI subscales each demonstrated moderate internal reliability in the current study ( $\alpha_{PT} = .77$ ,  $\alpha_{FS} = .81$ ,  $\alpha_{EC} = .77$ , and  $\alpha_{PD} = .73$ ), with scores ranging from 8 to 32 (PT;  $M = 21.45$ ,  $SD = 4.51$ ), 4 to 28 (FS;  $M = 18.92$ ,  $SD = 5.10$ ), 9 to 34 ( $M = 25.29$ ;  $SD = 4.54$ ), and 4 to 32 (PD;  $M = 16.48$ ,  $SD = 4.56$ ). A strong association between other-awareness and perspective taking, love and empathic concern, and courage and personal distress (relative to the other ACL Scale factors) would lend support for the convergent/discriminant validity of the ACL Scale subscales.

***Social intelligence.*** The Reading the Mind-in-the-Eyes Test (RMET; Baron-Cohen et al., 2001) is a 36-item measure designed to evaluate the degree to which one can accurately appraise the mental state of others. For each item, respondents view a photograph of an individual’s facial expression (limited to the region surrounding the eyes) and select, from a list of four emotion words, which best describes the emotional state of the subject depicted in the photograph. Correct responses are coded dichotomously (Correct = 1, Incorrect = 0) such that higher scores indicate greater accuracy in the appraisal of others’ emotions. The MIES demonstrated moderate internal reliability in the current study (Cronbach’s  $\alpha = .77$ ), with scores ranging from 0 to 36 ( $M = 23.20$ ,  $SD = 5.55$ ). A strong association (relative to the other ACL Scale factors) between other-awareness and social intelligence would lend support for the convergent/discriminant validity of the ACL Scale subscales.

**Dispositional behavioral inhibition and activation.** The Behavioral Inhibition and Activation Scales (BIS/BAS; Carver & White, 1994) are comprised of a set of four measures designed to assess the degree to which the respondent's behavioral inhibition and activation systems regulate aversive and appetitive motives, respectively. The behavioral activation system is measured by three orthogonal subscales: drive (D;  $N_{items} = 4$ ), fun seeking (FS;  $N_{items} = 4$ ), and reward responsiveness (RR;  $N_{items} = 5$ ). The behavioral inhibition system (BIS) is measured unidimensionally ( $N_{items} = 7$ ). Respondents rate their level of agreement with each statement on the BIS/BAS using a 4-point Likert scale ranging from 1 ("Very true for me") to 4 ("Very false for me") such that higher scores indicate a greater motivation to pursue one's goals (D), a greater tendency to respond energetically to rewarding stimuli (RR), a greater degree of impulsivity in the pursuit of rewarding stimuli (FS), and a greater tendency to engage in avoidance behavior (e.g., experience negative affect) in the presence of threatening stimuli (BIS). All but two items are reverse coded prior to scoring such that higher scores indicate a greater degree of behavioral inhibition (BIS) and activation (BAS). The BIS/BAS subscales demonstrated acceptable reliability in the current study (Cronbach's  $\alpha_D = .72$ ,  $\alpha_{FS} = .65$ ,  $\alpha_{RR} = .64$ , and  $\alpha_{BIS} = .75$ ), with subscale scores ranging from 4 to 16 (D;  $M = 11.33$ ,  $SD = 2.12$ ), 5 to 16 (FS;  $M = 12.17$ ,  $SD = 2.10$ ), 10 to 20 (RR;  $M = 17.59$ ,  $SD = 1.91$ ), and 9 to 28 (BIS;  $M = 22.07$ ,  $SD = 3.44$ ). A strong association (relative to the other ACL Scale factors) between courage and the BIS/BAS Drive subscale would lend support for the convergent/discriminant validity of the ACL Scale subscales.

**Fear of intimacy.** The Fear of Intimacy Scale (FIS; Duscutner & Thelen, 1991) is a 35-item self-report measure of "anxiety about close, dating relationships" (p. 223).

Respondents rate the accuracy of each self-descriptive statement – within the context of their romantic partnership (or imagined romantic partnership if not currently dating) – on a 5-point Likert scale ranging from 1 (“Not at all characteristic of me”) to 5 (“Extremely characteristic of me”). Fifteen of the 35 items are reverse coded prior to scoring such that higher scores indicate a greater fear of intimacy. The FIS demonstrated strong internal consistency in the current study (Cronbach’s  $\alpha = .93$ ), with scores ranging from 37 to 160 ( $M = 82.94$ ,  $SD = 21.94$ ). A strong association (relative to the other ACL Scale factors) between courage and fear of intimacy would lend support for the convergent/discriminant validity of the ACL Scale subscales.

***Compassionate love for others.*** The Compassionate Love Scale (stranger-humanity version; CLS; Sprecher & Fehr, 2005) is a 21-item self-report measure of one’s attitude toward strangers/all of humanity “containing feelings, cognitions, and behaviors that are focused on caring, concern, tenderness, and an orientation toward supporting, helping, and understanding the other(s), particularly when the other(s) is (are) perceived to be suffering or in need” (p. 630). Respondents rate the accuracy with which each statement describes them on a 7-point Likert scale ranging from 1 (“Not at all true of me”) to 7 (“Very true of me”). The CLS demonstrated strong internal reliability in the current study (Cronbach’s  $\alpha = .96$ ), with scores ranging from 1 to 10 ( $M = 6.17$ ,  $SD = 1.65$ ). A strong association (relative to the other ACL Scale factors) between love and compassionate love for others would lend support for the convergent/discriminant validity of the ACL Scale subscales.

***Mindfulness.*** The Five Facet Mindfulness Questionnaire (FFMQ; Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006) is a 39-item multidimensional self-report measure of engagement in mindfulness-based behavior. The FFMQ is comprised of five subscales

that are scored individually and do not contribute to an overall scale score: (1) Observing (O) – noticing and/or attending to both internal and external experiences, (2) Describing (D) – verbally labeling one’s internal experiences, (3) Acting with awareness (A) – attending to one’s present-moment experience, (4) Nonjudging of inner experience (Nj) – “taking a non-evaluative stance toward thoughts and feelings,” and (5) Non-reactivity to inner experience (Nr) – allowing one’s thoughts and feelings to occur without affecting behavior (Baer et al., 2008). Respondents rate each item on a 5-point Likert scale ranging from 1 (“Never or very rarely true”) to 5 (“Very often or always true”). Nineteen of the 39 items are reverse coded prior to scoring such that higher scores indicate greater engagement in each respective subscale’s construct. The FFMQ demonstrated moderate to strong reliability in the current study across all five subscales (Cronbach’s  $\alpha_O = .77$ ,  $\alpha_D = .85$ ,  $\alpha_A = .85$ ,  $\alpha_{Nj} = .88$ , and  $\alpha_{Nr} = .76$ ), with scores ranging from 8 to 40 in the first four subscales (Observe:  $M = 26.73$ ,  $SD = 4.85$ ; Describe:  $M = 25.65$ ,  $SD = 5.07$ ; Acting with awareness:  $M = 23.49$ ,  $SD = 5.01$ ; Nonjudging:  $M = 23.86$ ,  $SD = 5.75$ ; and Nonreact:  $M = 21.84$ ,  $SD = 3.78$ ). A strong association between self-awareness and observing as well as courage and non-react would lend support for the convergent/discriminant validity of the ACL Scale subscales.

## **Procedure**

After agreeing to participate in the study, subjects were presented with an online battery of measures (described above) that took roughly one hour to complete. The study was administered online such that each subject could participate on their own time in a comfortable environment. Subjects in Samples 1, 2, and 4 were emailed one week after participating with an invitation to complete a shorter battery of measures that included the ACL Scale for the purpose of assessing test-rest reliability.

## **Data Analytic Strategy**

All analyses will be performed using the R programming language (R Core Team, 2016).

### **Preliminary Analyses**

Preliminary analyses were conducted on the initial pool of 96-items (see “item development” section). First, samples 1 and 2 were split into random halves and submitted to exploratory factor analyses (EFA). Each of the four subsamples yielded a discrepant factor structure. In order to reconcile these differences, we fit confirmatory factor analytic (CFA) models of each proposed structure on all four subsamples, testing a total of 16 models. We then compared the average fit of each model in order to assess whether there was a clear superior factor structure. Results indicated that the fit of the factor structure for the two subsamples from Sample 1 were superior. Only minor differences in mean fit for these two groups were found when tested across the four groups. Due to the large discrepancy between the factor structures proposed by the exploratory analyses on these subsamples, however, we were left with two relatively good fitting models that each suggested a different factor structure. In the attempt to obtain a more reliable factor structure, we therefore abandoned this data-driven approach in exchange for a more theory-driven approach to item selection.

The ACL model continued to guide the next round of preliminary analyses. A series CFAs that specified six factors corresponding to the self- and other- dimensions of awareness, courage, and love were estimated in order to reduce the number of scale items. Ultimately, the other-courage and self-love factors were eliminated for theoretical reasons (see “item development” section above). The majority of this paper thus focuses on the analyses performed on the 64-item, 4-factor structure that is consistent with Reis and Shaver’s (1988) theoretical model.<sup>1</sup>

### **Item reduction.**

---

<sup>1</sup> Results from the preliminary analyses are available upon request

Confirmatory factor analysis (CFA) will be used in lieu of an exploratory approach (i.e., EFA) given that the initial item pool was developed vis-à-vis an *a priori* hypothesized factor structure. We will estimate a series of sequential CFA models using full information maximum likelihood (FIML; Lee, Poon, & Bentler, 1990) on raw data with Huber-White standard errors (Huber, 1967; White, 1982; see “statistical assumptions” section below), eliminating the item with the lowest loading on each factor per iteration until a total of 20 items, equally distributed among the factors, is achieved. Standard errors will be estimated using the observed information matrix, rather than the expected information matrix, as simulation studies have shown that this method yields more accurate estimates under missing at random (MAR) data missingness patterns (Savalei, 2010). Samples 1-6 will be pooled for the item reduction phase of the study and later examined independently to assess for measurement invariance of the final item set (see “measurement invariance” section below). The `lavaan` package (Rosseel, 2012) in R (R Core Team, 2016) will be used to perform these analyses. See Table 2 for the initial item set and hypothesized factor structure.

Once a final set of items is achieved, we will assess overall model fit. There are several extant global fit indices (Kline, 2015), each of which assesses fit based on a different set of criteria. Per the recommendations provided by Kline (2015) and Jackson, Gillaspay, & Purc-Stephenson (2009), we will examine the model Yuan-Bentler  $\chi^2$  statistic (Y-B $\chi^2$ ; Yuan & Bentler, 1999; see “statistical assumptions” section), the Bentler Comparative Fit Index (R-CFI; Bentler, 1990), the Tucker-Lewis Index (R-TLI; Tucker & Lewis, 1973), the Steiger-Ling Root Mean Square Error of Approximation (R-RMSEA; Steiger, 1990), and the Standardized Root Mean Square Residual (R-SRMR). The R-CFI, R-TLI, R-RMSEA, and R-SRMR are prefixed with “R-” to indicate that they are each computed using the robust adjusted Yuan-Bentler  $\chi^2$  statistic. In the current study, we follow Hooper, Coughlan, & Mullen’s (2008)

recommendations for determining model fit:  $\chi^2$  p-value > 0.05, CFI > 0.95, TLI > 0.95, RMSEA < 0.07, and SRMR < 0.08.

### **Measurement invariance.**

Measurement invariance (MI; i.e., measurement equivalence) refers to a consistent, invariant relationship between the observed scale scores and the unobserved latent variable as the assessment tool is applied across different groups and time (e.g., Meade & Lautenschlager, 2004). As Drasgow (1984) points out, MI does not refer to equal means or distributional qualities (i.e., variance, skew, kurtosis) across groups, only that those values are measured equivalently. In order for a measure to be a scientifically valid assessment tool such that it can reliably measure the same construct between multiple groups, then, MI must be demonstrated.

We will use multi-group CFA to test for three types of invariant qualities, each increasingly more restrictive (and thus more difficult to achieve): (1) configural invariance (equivalent factor structure), (2) weak invariance (equivalent factor loadings), and (3) strong invariance (equivalent loadings and intercepts). If configural invariance is achieved, each successive model is evaluated in comparison to the previous (as each level of restriction is nested within the previous). Configural invariance assessment follows the same guidelines as described above (see “item reduction” section). Therefore, if configural invariance is demonstrated in the current study, we will assess each successive model by computing a  $\chi^2$  difference test as well as examining the change in other fit statistics as recommended by Chen (2007). For the purpose of having sufficient power to perform a multi-group CFA, Samples 3 and 4 will be analyzed as one group. The `lavaan` package (Rosseel, 2012) in R (R Core Team, 2016) will be used to perform these analyses.

### **Validity and reliability.**

Two types of reliability will be examined in the current study: internal reliability (i.e., internal consistency) and test-retest reliability. Internal reliability is a measure of the degree to which items within a measure correlate with one another. In the current study, we compute Cronbach's  $\alpha$  (Cronbach, 1951), one of many different methods for assessing internal reliability, which displays the proportion of total scale variance that is shared among the items. Test-retest reliability measures the degree to which scores on a measure are consistent across time when all other variables remain invariant. In the current study, we assess test-retest reliability by computing a Pearson product-moment correlation coefficient ( $r$ ) of total ACL Scale scores and subscale scores for 687 subjects across an average of 7.50 days ( $SD = 2.53$ ). Additionally, test-retest reliability will be assessed by way of a series of a dependent samples t-tests. Family-wise  $\alpha$  inflation will be controlled by use of a Bonferroni adjustment, yielding a pairwise comparison  $\alpha$ -criterion of .01 (.05/5).

Convergent and discriminant validity each refer to the degree to which a measure of one construct is associated with a measure of another construct (i.e., the criterion) with which it is related, and unrelated, respectively. We will examine convergent and discriminant validity at the total scale level as well as the subscale level by calculating Pearson's  $r$  for 15 of the scales and subscales as described above. In order to control for family-wise Type I error inflation, we will compare each correlation coefficient to a Bonferroni-adjusted alpha criterion of .05/75. Because each CFA model estimates oblique factors, convergent validity at the subscale level will be demonstrated if the particular association under examination has the highest value among the set of subscales and the criterion.

**Missing data.**

Missing data will be dealt with in two different ways. First, as mentioned above, we will use full information maximum likelihood (FIML) estimation for all structural equation models

estimated in the current study (i.e., the CFA and invariance models). FIML, unlike traditional maximum likelihood estimation (i.e., maximum Wishart likelihood), estimates a log-likelihood function for each individual subject based on that individual's observed data (Graham & Coffman, 2012). Thus, the algorithm *does not* impute missing data similar to other common missing data methods (e.g., multiple imputation; Rubin, 1987). Monte Carlo simulations comparing FIML with other common missing data methods (listwise deletion, pairwise deletion, and similar response pattern imputation) have demonstrated its superiority with respect to providing unbiased parameter estimates across varying sample sizes, factor loading magnitudes, and missing data rates (Enders & Bandalos, 2001).

Second, for those analyses that do not involve estimating parameters in a structural equation model (i.e. internal reliability analysis, construct validity analysis), we will substitute missing data using multiple imputation by chained equations (MICE; i.e., fully conditional specification, sequential regression multiple imputation; van Buuren, 2007). Monte Carlo simulations have shown a distinct power advantage to imputing missing data at the item level as opposed to the scale level (Gottschall, West, & Enders, 2012). Therefore, in this study we will impute 5 datasets, each estimated by 40 iterations of the MICE algorithm using the remaining scale items (those correspond with the missing endogenous variable) as predictors. We will then take the arithmetic mean for each cell in all 5 datasets to create one dataset from which to perform our analyses. The `mice` package (van Buuren & Groothuis-Oudshoorn, 2011) in R (R Core Team, 2016) will be used to conduct these procedures.

### **Statistical assumptions.**

Maximum likelihood estimation (including FIML) assumes that the joint probability distribution of endogenous variables follows a multivariate normal distribution. Two issues with respect to this assumption arise in the current study. First, in order for the joint probability

distribution to be normally distributed, the variables must be measured on a continuous (i.e., interval or ratio) scale. The scale under development in the current study, however, uses Likert indicators. The second issue concerns the high dimensionality of our models, which initially consists of 64 indicators. To the authors' knowledge, there does not exist a measure of multivariate normality that is accurate at such high dimensions. Procedures for addressing this issue in the current study are twofold. First, given the use of seven anchors on our Likert scale, we will assume that each item approximates a continuous distribution (Kline, 2015; Rasmussen, 1989). Second, to address the issue of potential violations of multivariate normality, we will estimate model parameters with Huber-White standard errors (Huber, 1962; White, 1983). Furthermore, the Yuan-Bentler  $\chi^2$  statistic (Y-B  $\chi^2$ ; Yuan & Bentler, 1999), which corrects for multivariate non-normality by dividing the  $\chi^2$  value by a scaling correction factor, will be calculated and used performing subsequent fit calculations that take the  $\chi^2$  value into account (CFI, RMSEA, and SRMR).

Both the FIML and MICE assume that the data are *missing at random* (MAR; Rubin, 1976). Missing data are considered MAR, as opposed to missing *completely at random* (MCAR), when the pattern of missingness is independent of the missing data (Schafer & Graham, 2002) but dependent upon the observed data. Although it is possible to test whether the data are MCAR (e.g., by comparing the missingness patterns across levels of another variable), it is not possible to test whether they are MAR. For that reason, in the current study we assume our data are MAR.

Lastly, given that the Pearson product-moment correlation coefficient is a quantification of the *linear* relationship between two variables, we checked for linearity prior to each reliability/validity analysis.

## Results

### Factor Structure

A series of 14 sequential confirmatory factor analytic (CFA) models were estimated, each model containing one fewer item per factor than the previous (until only 5 items remained on a factor), resulting in a 20-item, 4-factor solution. The final model (henceforth referred to as simply as “the model”) converged after 95 iterations of full information maximum likelihood estimation (FIML) on a total of 1,804 cases. Huber-White robust standard errors (Huber, 1967; White, 1982) were calculated for the model parameters, and fit indices were adjusted for multivariate non-normality using the Yuan-Bentler  $\chi^2$  statistic (Y-B $\chi^2$ ; Yuan-Bentler, 1999). A correlation matrix of the initial 64-item set can be found in Appendix A.

The model demonstrated adequate overall fit. As expected with large sample sizes (Cheung & Rensvold, 2002), the Y-B $\chi^2$  was statistically significant, Y-B $\chi^2(164, N = 1,804) = 773.76, p < .001$ . The other fit measures, however, generally indicated support for the model, R-RMSEA = .045 (90% CI [.043-.048]), R-CFI = .94, R-TLI = .93, and R-SRMR = .04. Standardized factor loadings provide additional support for the multi-dimensionality of the ACL Scale ( $M = .68, SD = .08$ ; see Table 3). The factors were strongly correlated with each other (see Table 4).

Areas of local misfit were identified upon examination of standardized residual correlations (observed minus model-implied correlations). Of the 210 total residual correlations, 38.57% had a magnitude lower than 1, 22.86% lower than 2 (but greater than 1), 18.10% lower than 3 (but greater than 2), and 20.48% greater than 3 ( $M = 0.23, SD = 3.18$ ). Although there were other large residuals, one stood out in particular with a value of 26.80. See Appendix B for the residual correlation and covariance matrices of this model. Table 3 contains item and factor variances as well as item  $R^2$  values.

### Measurement Invariance

Measurement invariance of the final factor structure was tested across Samples 1-6 (Samples 3 & 4, however, were pooled for sufficient sample size) by estimating a series of hierarchical, multi-

group confirmatory factor analytic (CFA) models. Similar to the CFA models estimated in the item reduction analyses, we estimated the model parameters using robust Hubert-White (Huber, 1967; White, 1982) standard errors and report fit indices using the Yuan-Bentler  $\chi^2$  statistic (Y-B $\chi^2$ ; Yuan-Bentler, 1999) to correct for multivariate non-normality of our data. The results of each test of invariance can be found in Table 5.

First, *configural invariance* was tested. Although the  $\chi^2$  test is statistically significant, the other fit indices suggest acceptable configural invariance across groups. The next model tests *weak invariance* by constraining the factor loadings to equivalence across groups. The weak model did not differ appreciably from the configural model according to the recommendations suggested by Chen's (2007) Monte Carlo simulations: although the  $\Delta$ CFI was .01, it was not accompanied by a  $\Delta$ RMSEA value  $\geq$  .015 or a  $\Delta$ SRMR value  $\geq$  .03. Further support for weak invariance is provided by the  $\Delta$ BIC, however it noteworthy that the  $\Delta$ Y-B $\chi^2$  test and  $\Delta$ AIC suggest non-invariance at this level. Next, *strong invariance* was tested by further constraining the factor intercepts to equivalence across groups. According to Chen (2007), the fit suggests strong invariance across groups was met despite the fact that the  $\Delta$ CFI is greater than .01; the  $\Delta$ RMSEA is  $<$  .015 and the  $\Delta$ SRMR is  $<$  .01. The  $\Delta$ BIC provides further support for strong invariance, however the  $\Delta$ Y-B $\chi^2$  test and  $\Delta$ AIC continue to suggest non-invariance at this level.

It is important to note that the tests performed above are tests of *relative* fit, and not *absolute* fit. While some fit indices (i.e., the RMSEA, SRMR) suggest adequate absolute model fit, others (i.e., the Y-B $\chi^2$  statistic, CFI, TLI) do not.

### **Reliability**

Internal reliability was assessed using Cronbach's alpha ( $\alpha$ ; Cronbach, 1951). Given the high correlation among ACL Scale subscales,  $\alpha$  was calculated for each individual subscale as well as the total scale. Results demonstrated strong internal reliability at the total scale level ( $\alpha = .95$ ), and

moderate-to-strong internal reliability at the subscale level (Other-Awareness:  $\alpha = .79$ ; Self-Awareness:  $\alpha = .78$ ; Courage:  $\alpha = .78$ ; Love:  $\alpha = .89$ ). Internal reliability of the ACL Scale assessed at one-week follow-up yielded similar results (Total:  $\alpha = .95$ ; Other-Awareness:  $\alpha = .82$ ; Self-Awareness:  $\alpha = .82$ ; Courage:  $\alpha = .81$ ; Love:  $\alpha = .89$ ).

Test-retest reliability was assessed via two different modalities. First, we calculated a Pearson product-moment correlation coefficient ( $r$ ) between the total scale scores, as well as subscale scores, at each time point. Pearson's  $r$  for the total scale score fell in the acceptable range, as did the love subscale score (Nunally, 1970). The remaining subscales (other-awareness, self-awareness, and courage) fell below the acceptable range. See Table 6 for all test-retest correlations and ACL Scale means. We also performed a series of five dependent samples t-tests to assess for mean differences across time. Each pairwise comparison was assessed for statistical significance with an alpha criterion of .05. There was a statistically significant difference between total scale scores,  $t(686) = 2.88, p = .004$ , 95% CI [0.02, 0.10], other-awareness subscale scores,  $t(686) = 2.34, p = .01$ , 95% CI [0.02, 0.13], self-awareness subscale scores,  $t(686) = 2.34, p = .02$ , 95% CI [0.01, 0.13], and love subscale scores,  $t(687) = 3.37, p < .001$ , 95% CI [0.04, 0.15]. There was no statistically significant difference between courage subscale scores  $t(686) = -0.58, p = .56$ , 95% CI [-0.07, 0.04].

Given the large sample size used in this study, it is not surprising that 4 of the 5 comparisons were statistically significant. Thus, in order to assess the magnitude of scale score changes over time, we calculated an effect size for each comparison. The effect size calculations all fell under the "small" range (Cohen, 1977; Total  $d = 0.06$ , Other-awareness: 0.10; self-awareness: 0.09; Courage: 0.02; and Love: 0.10), which is also evident upon examination of the 95% confidence interval of the mean difference. Although test-retest difference scores were small, they nevertheless warrant caution on behalf of researchers until the reasons for these differences are better understood.

### **Convergent and Discriminant Correlations**

We predicted that total ACL Scale scores would be strongly correlated with social connectedness, loneliness, emotional intelligence, social support, and frequency of intimacy-related behaviors. As is shown in Table 7, these correlations, while all statistically significant, are not all as strong as we predicted ( $M = .47$ , minimum = .32, maximum = .74). Furthermore, in order to examine convergent and discriminant validity of the subscales, we predicted each to correlate most strongly with a unique set of constructs. Other-awareness was predicted to correlate most strongly with perspective taking and social intelligence, self-awareness with observe skills, courage with anxiety in interpersonal settings, BAS Drive, fear of intimacy, and non-reactivity to inner experience, and love with empathic concern. Five of these predictions (62.5%) were met, while remaining 3 (37.5%) were not providing partial support for convergent/discriminant validity of the ACL Scale subscales. These Pearson correlation coefficients are also reported in Table 7.

### **Discussion**

The purpose of this study was to develop and validate a self-report measure of intimacy that has a strong empirical foundation in the intimacy process literature, is useful in a clinical setting, is based contextual behavioral scientific principles, and is applicable to a broad range of populations. We estimated a series of confirmatory factor analytic (CFA) models that yielded a preliminary scale with 20 indicators (i.e., items) equally distributed among four factors. Next, we examined measurement invariance of this preliminary measure by estimating a series of 3 CFA models, each with an increasingly strict set of restrictions. Finally, reliability (internal and test-retest) and validity (convergent/discriminant) of the ACL Scale were assessed.

The preliminary ACL Scale demonstrated adequate overall model fit. Although the  $\chi^2$  statistic suggested poor fit of the model, it is well known that this test is overly sensitive to large sample sizes. (Kline, 2015). Two other measures of overall fit, the CFI and TLI, were on the border of acceptable levels (Hooper et al., 2008). Finally, the RMSEA and SRMR both suggested strong model fit. It is

possible, given the high correlation of the other-awareness factor with the self-awareness and love factors, that model misspecification occurred at this level. Perhaps, for example, other-awareness and self-awareness are better represented as one factor instead of two. Future research should examine this possibility.

Examination of the discrepancies between model-implied and observed correlations (i.e., residual correlations) presented greater cause for concern. More than half of the standardized residuals had magnitudes greater than 1, with extreme values observed throughout the matrix. Visual inspection of the standardized residuals did not indicate a clear problematic set of items, suggesting that the misspecification is more systemic. Future analyses are needed in order to address local fit misspecification for the final version of the ACL Scale.

The 20-item, four-factor model examined in this study is theoretically consistent with Reis & Shaver's (1988) model for the development of intimacy despite its imperfect fit. The first factor – other-awareness – refers to one's awareness of opportunities to engage in vulnerable self-disclosures as well as to respond to the vulnerability of others in a manner that makes them feel understood, validated, and cared for. Unlike other-awareness which emphasizes *interpersonal* contextual factors involved in intimacy, self-awareness refers to one's awareness of *intrapersonal* factors that augment the interpersonal components of the intimacy generative process. These include awareness of one's values in interpersonal settings, emotional and behavioral reactions to others, and idiosyncratic definitions of what constitutes a vulnerable self-disclosure. Given the inherent interpersonal risk involved in the development of intimacy, the courage factor measures one's inclination to engage in risks as they relate to one's interpersonal values. The courage factor also captures one's inclination to engage in non-interpersonal values-consistent behavior given its strong association with relationship satisfaction. Finally, the love factor measures one's inclination to engage in interpersonal behaviors that serve the

function of making one's interaction partner feel understood, validated, and cared for; these include both verbal and nonverbal behaviors.

Although the  $\Delta\chi^2$  value in each measurement invariance analysis was statistically significant, simulation studies (Cheun & Rensvold, 2002; Meade, Johnson, & Braddy, 2008) suggest that it is overly sensitive to large sample size in tests of invariance. It is possible that significance was found for the first model comparison (weak vs. configural invariance) given its relatively large p-value, however it is worth noting that this does not apply to the second comparison (strong vs. weak invariance). We also examined other measures of model fit according to Chen's (2007) criteria and found support for the ACL Scale's invariant qualities across all samples (Samples 3 and 4 were pooled). It is worth noting that examination of absolute model fit, as opposed to relative fit, was questionable at increasingly strict levels of invariance.  $\Delta\chi^2$  value to large sample sizes, the obtained p-values were Future research is needed to explore the ACL Scale's invariant qualities across a more diverse set of groups given the convenience samples used in this study.

Internal reliability was strong across each of the subscales as well as the total ACL Scale, suggesting that this measure can be used with relatively little concern about measurement error. Results from test-retest analyses, however, were mixed. First, we calculated a Pearson product-moment correlation coefficient for subjects who took the measure at follow-up (an average of 7.5 days after the first). The ACL Scale, as well as the love subscale, showed acceptable reliability across the two time points, however the remaining subscales – other-awareness, self-awareness, and courage – did not. We then performed a series of dependent samples t-tests to compare total scale and subscale scores across both time points. Results from these analyses suggested statistically significant differences for the total scale and the love subscale, but not the remaining three factors. Although the effect sizes for the statistically significant comparisons were small (Cohen, 1977), researchers should exercise caution

when using the ACL Scale in longitudinal designs. Future research is needed to explore test-retest reliability of the measure.

Convergent and discriminant validity analyses provided partial support for the theoretical foundation of the ACL Scale and each of its subscales. Of the six criterion measures used for total scale convergent validity, one was in the “small effect” range, three were in the “medium effect” range, and two were in the “strong effect” range (Cohen, 1988). Six of our nine predictions for convergent/discriminant validity were confirmed, only one of which is considered a “large” effect according to Cohen’s criteria (1988). Based on previous research (e.g., Wheeler, Reis, and Nezlek (1983), we expected a stronger correlation between total ACL Scale scores, social connectedness, loneliness, and frequency of intimacy-related behaviors, and partner intimacy. It is possible that the latter construct (measured by the FAPIS) is subject to a great deal of sampling variability given that each subject responds with respect to a given individual within a narrow time period (the previous week). Further research is needed to test this hypothesis as well as to explore reasons that may underlie weaker correlations than expected (including the possibility that the ACL Scale is not measuring intimacy as it was designed to do).

Although partial support was provided for the other-awareness and self-awareness subscales, confirmed predictions were just marginally larger than those of the criterion will the other subscales (unconfirmed predictions were also marginally smaller). The difference in Pearson  $r$  values for the predicted correlations with the courage and love subscales were larger than those of the other two subscales, however they were still small discrepancies in absolute terms. It is likely that the high correlation among ACL Scale subscales is responsible for these findings. Future research exploring the convergent/discriminant validity of the ACL Scale is warranted.

### **Strengths, Limitations, and Future Directions**

The current study had many strengths as well as limitations that are worth considering. First, our large sample size provided a great deal of statistical power (e.g., Wolf, Harrington, Clark, & Miller, 2013), which is particularly important when estimating large CFA models such as those found in the beginning of our item reduction analyses as well as in tests of measurement invariance. Our test-retest sample was also larger than those typically seen in scale development research. An additional strength of this research is the geographic diversity of our sample. While it is unfeasible to sample from all regions, we had representation in three major regions of North America as well as one internationally (Perth, Australia).

With these strengths also come limitations. First, although we had a geographically diverse sample, the majority of subjects were White university undergraduates. Further research is needed to explore the ACL Scale's psychometric properties in a more representative, ecologically valid sample. Second, despite our aim to create a clinically useful measure of intimacy, we did not collect any data on the performance of the measure in a clinical sample. Thus, researchers using this measure with clinical samples should exercise caution until further research is conducted. An additional limitation concerns the current study's lack of multi-trait multi-method methodology for analyzing discriminant validity. As Campbell and Fiske (1959) point out, it is thus impossible to separate the "relative contributions of trait and method variance" (p. 81). Furthermore, it is possible that the criteria proposed by Chen (2007) for measurement invariance fit are not applicable to the current study given inconsistencies between the data used in the current study and the data used in Chen's (2007) simulations: 5 vs. 2 groups, 5 vs. 8-12 indicators per factor, 4 vs 1 factor, and 1,804 vs. 150, 250, and 500 observations.

Given the limitations described above, more work is necessary in order to derive a psychometrically valid, theoretically consistent measure of intimacy. Future research needs to, first and

foremost, explore areas model misspecification and make the necessary adjustments. Perhaps the item reduction analyses used in the current study are insufficient given their narrow focus. Future research could, for example, also take into consideration changes in factor variances, residual item covariances, and test-retest reliability when performing item reduction.

### References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723.
- Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, 23(4), 363-377.
- Asher, S. R., Hymel, S., & Renshaw, P. D. (1984). Loneliness in children. *Child Development*, 1456-1464.
- Ashida, S., & Heaney, C. A. (2008). Differential associations of social support and social connectedness with structural features of social networks and the health status of older adults. *Journal of Aging and Health*, 20(7), 872-893.
- Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, 13(1), 27-45.
- Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., ... & Williams, M. (2008). Construct validity of the Five Facet Mindfulness Questionnaire in meditating and nonmeditating samples. *Assessment*, 15(3), 329-342.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42(2), 241-251.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497-529.
- Beadle, J. N., Brown, V., Keady, B., Tranel, D., & Paradiso, S. (2012). Trait empathy as a predictor of individual differences in perceived loneliness. *Psychological reports*, 110(1), 3-15.

- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238-246.
- Berscheid, E., Snyder, M., & Omoto, A. M. (1989). The Relationship Closeness Inventory: Assessing the closeness of interpersonal relationships. *Journal of Personality and Social Psychology*, *57*(5), 792-807.
- Biglan, A., & Hayes, S. C. (1997). Should the behavioral sciences become more pragmatic? The case for functional contextualism in research on human behavior. *Applied and Preventive Psychology*, *5*(1), 47-57.
- Brewer, M. B. (2004). Taking the social origins of human nature seriously: Toward a more imperialist social psychology. *Personality and Social Psychology Review*, *8*(2), 107-113.
- Buhrmester, D. (1990). Intimacy of friendship, interpersonal competence, and adjustment during preadolescence and adolescence. *Child Development*, *61*(4), 1101-1111.
- Cacioppo, J. T., Hawkley, L. C., Ernst, J. M., Burleson, M., Berntson, G. G., Nouriani, B., & Spiegel, D. (2006). Loneliness within a nomological net: An evolutionary perspective. *Journal of research in personality*, *40*(6), 1054-1085.
- Cacioppo, J. T., Hughes, M. E., Waite, L. J., Hawkley, L. C., & Thisted, R. A. (2006). Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. *Psychology and Aging*, *21*(1), 140-151.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81-105.
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*(2), 319-333. [http://dx.doi.org/10.1037//0022-3514.67.2.319\\_](http://dx.doi.org/10.1037//0022-3514.67.2.319_)

- Caspi, A., Harrington, H., Moffitt, T. E., Milne, B. J., & Poulton, R. (2006). Socially isolated children 20 years later: risk of cardiovascular disease. *Archives of Pediatrics & Adolescent Medicine, 160*(8), 805-811.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equations Modeling, 14*(3), 464-504.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York, NY: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: a meta-analytic review. *Psychological Bulletin, 116*(3), 457-475.
- Cordova, J. V., & Scott, R. L. (2001). Intimacy: A behavioral interpretation. *The Behavior Analyst, 24*(1), 75-86.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.
- Davis, M. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.
- Descutner, C. J., & Thelen, M. H. (1991). Development and validation of a Fear-of-Intimacy Scale. *Psychological assessment: A journal of consulting and clinical psychology, 3*(2), 218-225.
- Dragow, F. (1984). Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological Bulletin, 95*(1), 134-135.

- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8*(3), 430-457.
- Gottschall, A. C., West, S. G., & Enders, C. K. (2012). A comparison of item-level and scale-level multiple imputation for questionnaire batteries. *Multivariate Behavioral Research, 47*(1), 1-25.
- Graham, J. W., & Coffman, D. L. (2012). Structural equation modeling with missing data. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 277-295). New York, NY: Guildford Press.
- Grewen, K. M., Girdler, S. S., Amico, J., & Light, K. C. (2005). Effects of partner support on resting oxytocin, cortisol, norepinephrine, and blood pressure before and after warm partner contact. *Psychosomatic Medicine, 67*(4), 531-538.
- Grieve, R., Indian, M., Witteveen, K., Tolan, G. A., & Marrington, J. (2013). Face-to-face or Facebook: Can social connectedness be derived online? *Computers in Human Behavior, 29*(3), 604-609.
- Grote, N. K., & Frieze, I. H. (1994). The measurement of Friendship-based Love in intimate relationships. *Personal Relationships, 1*(3), 275-300.
- Hawkley, L. C., & Cacioppo, J. T. (2010). Loneliness matters: a theoretical and empirical review of consequences and mechanisms. *Annals of Behavioral Medicine, 40*(2), 218-227.
- Haworth, K., Kanter, J. W., Tsai, M., Kuczynski, A. M., Rae, J. R., & Kohlenberg, R. J. (2015). Reinforcement matters: A preliminary, laboratory-based component-process analysis of Functional Analytic Psychotherapy's model of social connection. *Journal of Contextual Behavioral Science, 4*(4), 281-291.
- Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality a meta-analytic review. *Perspectives on Psychological Science, 10*(2), 227-237.

- Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS Medicine*, 7(7), e1000316.
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modeling: Guidelines for determining model fit. *Journal of Business Research Methods*, 6, 53-60.
- Huber, P. J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1), pp. 221-233. Berkeley, CA: University of California Press.
- Jackson, D. L., Gillaspay, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: A review and some recommendations. *Psychological Methods*, 14(1), 48-76.
- Kline, R. B. (2015) *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.
- Kohlenberg, R. J., & Tsai, M. (1991). *Functional Analytic Psychotherapy: Creating intense and curative therapeutic relationships*. New York, NY: Plenum Press.
- Larzelere, R. E., & Huston, T. L. (1980). The dyadic trust scale: Toward understanding interpersonal trust in close relationships. *Journal of Marriage and the Family*, 595-604.
- Laurenceau, J. P., Barrett, L. F., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: The importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology*, 74(5), 1238-1251.
- Lawson, J. S., Marshall, W. L., & McGrath, P. (1979). The social self-esteem inventory. *Educational and Psychological Measurement*, 39(4), 803-811.
- Lee, R. M., Draper, M., & Lee, S. (2001). Social connectedness, dysfunctional interpersonal behaviors, and psychological distress: Testing a mediator model. *Journal of Counseling Psychology*, 48(3), 310-318.

- Lee, S., Poon, W. & Bentler, P. M. (1990) A three-stage estimation procedure for structural equation models with polytomous variables. *Psychometrika*, 55(1), 45-51.
- Lee, R. M., & Robbins, S. B. (1995). Measuring belongingness: The Social Connectedness and Social Assurance Scales. *Journal of Counseling Psychology*, 42(2), 232-241.
- Leonard, R. C., Knott, L. E., Lee, E. B., Singh, S., Smith, A. H., Kanter, J. W, ... & Wetterneck, C. T. (2014). The development of the Functional Analytic Psychotherapy Intimacy Scale. *The Psychological Record*, 64(4), 647-657.
- Liang, B., Tracy, A., Taylor, C. A., Williams, L. M., Jordan, J. V., & Miller, J. B. (2002). The relational health indices: A study of women's relationships. *Psychology of Women Quarterly*, 26(1), 25-35.
- Lopez, F. G., & Rice, K. G. (2006). Preliminary development and validation of a measure of relationship authenticity. *Journal of Counseling Psychology*, 53(3), 362-371.
- Lutgendorf, S. K., Sood, A. K., Anderson, B., McGinn, S., Maseri, H., Dao, M., ... & Lubaroff, D. M. (2005). Social support, psychological distress, and natural killer cell activity in ovarian cancer. *Journal of Clinical Oncology*, 23(28), 7105-7113.
- Masi, C. M., Chen, H. Y., Hawkey, L. C., & Cacioppo, J. T. (2010). A meta-analysis of interventions to reduce loneliness. *Personality and Social Psychology Review*, 15(3), 219-266.
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, 93(3), 568-592.
- Meade, A. W., & Lautenschlager, G. J. (2004) A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, 7(4), 361-388.
- Miller, R. S., & Lefcourt, H. M. (1982). The assessment of social intimacy. *Journal of Personality Assessment*, 46(5), 514-518.
- Nunally, J. C. (1970). *Introduction to psychological measurement*. New York, NY: McGraw-Hill.

- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. <http://dx.doi.org/10.1126/science.aac4716>
- Pavey, L., Greitemeyer, T., & Sparks, P. (2011). Highlighting relatedness promotes prosocial motives and behavior. *Personality and Social Psychology Bulletin*, 37(7), 905-917.
- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available from <https://www.R-project.org/>
- Rasmussen, J. L. (1989) Analysis of Likert-scale data: A reinterpretation of Gregoire and Driver. *Psychological Bulletin*, 105(1), 167-170.
- Reis, H. T. & Shaver, P. (1988). Intimacy as an interpersonal process. In S. Duck, D. Hay, S. E. Hobfoll, W. Ickes, & B. M. Montgomery (Eds.), *Handbook of personal relationships: Theory, research and interventions* (pp. 367-389). Chichester, UK: John Wiley & Sons.
- Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin*, 26(4), 419-435.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistics Software*, 48(2), 1-36.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581-592.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: J. Wiley & Sons.
- Russell, D. W. (1996). UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of personality assessment*, 66(1), 20-40.
- Savalei, V. (2010). Expected vs. observed information in SEM with incomplete normal and nonnormal data. *Psychological Methods*, 15(4), 352-367.
- Schafer, J. L, & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147-177.

- Schaefer, M. T., & Olson, D. H. (1981). Assessing intimacy: The PAIR inventory. *Journal of Marital and Family Therapy*, 7(1), 47-60.
- Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, 25(2), 167-177.
- Skinner, B. F. (1953). Punishment. In B. F. Skinner (Ed.), *Science and human behavior* (pp. 182-193). NY, New York: The Free Press.
- Skinner, B. F. (1958). Reinforcement today. *American Psychologist*, 13(3), 94-99.
- Sherbourne, C. D., & Stewart, A. L. (1991). The MOS social support survey. *Social science & medicine*, 32(6), 705-714.
- Sprecher, S., & Fehr, B. (2005). Compassionate love for close others and humanity. *Journal of Social and Personal Relationships*, 22(5), 629-651 .
- Steiger, J. H. (1990). Structural model evaluation and modification: An interbal estimation approach. *Multivariate Behavioral Research*, 25(2), 173-180
- Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine*, 16(3), 199-202
- Thomas, S. P., & Smith, H. (2004). School connectedness, anger behaviors, and relationships of violent and nonviolent American youth. *Perspectives in psychiatric care*, 40(4), 135-148.
- Thurston, R. C., & Kubzansky, L. D. (2009). Women, Loneliness, and Incident Coronary Heart Disease. *Psychosomatic Medicine*, 71(8), 836-842.  
<http://doi.org/10.1097/PSY.0b013e3181b40efc>
- Tsai, M. Kohlenberg, R. J., Kanter, J. W., Kohlenberg, B. Follette, W. C., & Callaghan, G. M. (2009). *A guide to Functional Analytic Psychotherapy: Using awareness, courage, love, and behaviorism*. New York, NY: Springer.

- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1-10.
- van Buuren, S. (2007). Multiple imputation of discrete and continuous data by full conditional specification. *Statistical Methods in Medical Research*, 16(3), 219-242.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1-67.
- Van Orden, K. A., Witte, T. K., James, L. M., Castro, Y., Gordon, K. H., Braithwaite, S. R., ... & Joiner, T. E. (2008). Suicidal ideation in college students varies across semesters: The mediating role of belongingness. *Suicide and Life-Threatening Behavior*, 38(4), 427-435.
- Wheeler, L., Reis, T., & Nezlek, J. (1983). Loneliness, social interaction, and sex roles. *Journal of Personality and Social Psychology*, 45(4), 943-953.
- White, H. (1982). Maximum likelihood estimation of misspecified models, *Econometrika*, 50(1), 1-25.
- Wilson, K. G., Sandoz, E. K., Kitchens, J., & Roberts, M. (2010). The Valued Living Questionnaire: Defining and measuring valued action within a behavioral framework. *The Psychological Record*, 60(2), 249-272.
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models an evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 73(6), 913-934.
- Yeh, C. J., & Inose, M. (2003). International students' reported English fluency, social support satisfaction, and social connectedness as predictors of acculturative stress. *Counselling Psychology Quarterly*, 16(1), 15-28.
- Yuan, K. H. & Bentler, P. M. (1999). On asymptotic distributions of normal theory MLE in covariance structure analysis under some non-normal distributions. *Statistics & Probability Letters*, 42(2), 107-113.

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, 52(1), 30-41.

Table 1.  
Demographic Characteristics by Sample

	Time 1			
	Sample			
	Sample 1 ( <i>n</i> = 687)	Sample 2 ( <i>n</i> = 465)	Sample 3 ( <i>n</i> = 150)	Sample 4 ( <i>n</i> = 62)
Female: <i>n</i> (%)	383 (55.75%)	322 (69.25%)	86 (57.33%)	48 (77.42%)
Age: M (SD)	19.06 (1.41)	20.31 (2.81)	20.53 (3.07)	19.31 (1.37)
Ethnicity: <i>n</i> (%)				
White/Caucasian	230 (33.48%)	346 (74.41%)	138 (92.00%)	54 (87.10%)
Black	19 (2.77%)	61 (13.12%)	0 (0.00%)	3 (4.84%)
Non-white Hispanic	23 (3.35%)	9 (1.94%)	5 (3.33%)	0 (0.00%)
Middle Eastern	15 (2.18%)	91.94%)	0 (0.00%)	1 (1.61%)
Asian	339 (49.34%)	21 (4.52%)	42.67%)	0 (0.00%)
Southeast Asian	25 (3.64%)	3 (0.65%)	0 (0.00%)	1 (1.61%)
Native American	1 (0.15%)	2 (0.43%)	1 (0.67%)	0 (0.00%)
Pacific Islander	5 (0.73%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Other	26 (3.78%)	14 (3.01%)	2 (1.33%)	3 (4.84%)
	Time 2			
	Sample 1 ( <i>n</i> = 598)	Sample 2 ( <i>n</i> = 56)	Sample 4 ( <i>n</i> = 33)	
Female: <i>n</i> (%)	335 (56.02%)	41 (73.21%)	25 (75.76%)	
Age: M (SD)	19.07 (1.46)	20.34 (2.01)	19.17 (1.15)	
Ethnicity: <i>n</i> (%)				
White/Caucasian	205 (34.28%)	44 (78.57%)	29 (87.88%)	
Black	18 (3.01%)	4 (7.14%)	2 (6.06%)	
Non-white Hispanic	19 (3.18%)	0 (0.00%)	0 (0.00%)	
Middle Eastern	15 (2.51%)	1 (1.79%)	0 (0.00%)	
Asian	291 (48.66%)	5 (8.93%)	0 (0.00%)	
Southeast Asian	21 (3.51%)	0 (0.00%)	0 (0.00%)	
Native American	1 (0.17%)	1 (1.79%)	0 (0.00%)	
Pacific Islander	3 (0.50%)	0 (0.00%)	0 (0.00%)	
Other	25 (4.18%)	1 (1.79%)	2 (6.06%)	

	Sample	
	Sample 5 ( <i>n</i> = 183)	Sample 6 ( <i>n</i> = 257)
Gender (Female) <i>n</i> (%)	183 (77.05%)	257 (76.26%)
Age M (SD)	21.27(4.79)	34.98 (12.43)
Ethnicity N (%)		
Australian	144 (78.69%)	177 (68.87%)
Aboriginal	0 (0.00%)	3 (1.17%)
Australian and Torres Strait Islander		
New Zealander	2 (1.09%)	5 (1.95%)
Oceanian	0 (0.00%)	1 (0.39%)
North-Western European	8 (4.37%)	32 (12.45%)
Southern and Eastern European	5 (2.73%)	4 (1.56%)
North African and Middle East	4 (2.19%)	1 (0.39%)
Southeast Asian	11 (6.01%)	3 (1.17%)
North East Asian	1 (0.55%)	5 (1.95%)
Southern and Central Asian	4 (2.19%)	3 (1.17%)
People of the Americas	0 (0.00%)	16 (6.23%)
Sub-Saharan African	1 (0.55%)	0 (0.00%)
Other	3 (1.64%)	7 (2.72%)

Table 2.  
Initial Item Set and Hypothesized Factor Structure

---

**Other-awareness**

1. Even when I am emotional, I can see the other person's perspective.
2. I am able to listen deeply to others.
3. I am aware of actions that I take that others may judge or dislike.
4. I am aware of the times when I could be caring, supportive, and loving towards others.
5. I am aware of times when others are trying to be caring, supportive, or loving toward me.
6. I ask questions of others to help me understand exactly what is happening for them in the moment.
7. I am aware of what someone else needs, even when I am arguing with them.
8. I can anticipate people's wants and needs.
9. I know when other people are taking interpersonal risks.
10. I see the other person's perspective when we are in an intense conversation.
11. If I do not understand someone's view, I will ask for clarification.
12. It is hard for me to listen to others talk about personal problems.
13. People do not seem to understand me.
14. People tell me I am cold and distant even when I do not think I am.
15. People tell me I am an empathic person.

**Self-awareness**

16. I am aware of activities, people, or situations that make me feel uncomfortable.
17. I am aware of important moments to be kind and compassionate to myself.
18. I am aware of my reactions or response to others as they occur.
19. I am aware of what makes me feel vulnerable.
20. I am aware of how my reactions affect how I feel.
21. I feel isolated from humanity.
22. I feel my emotions without being too negative or judgmental about them, even if they are painful.
23. I know when I am doing what matters to me.
24. I notice how other people affect how I feel.
25. I notice how what I feel affects what I do.
26. I often do not know what causes me to feel the way I do.
27. I usually do not know what I am feeling,
28. I want to get rid of my negative thoughts, feelings, and memories.
29. I am aware of my feelings as they happen.
30. I am not aware of my bodily sensations.
31. I am unwilling to experience painful thoughts and feelings.

32. When I feel something, I generally am aware of what caused it.

33. When I feel uncomfortable, I have no idea why.

### **Courage**

34. Even if I feel uncomfortable, I will stay in a situation until I have done what I need to do.

35. I am afraid to tell others what I think and feel.

36. I ask others for support when I need it,

37. I will act for something I believe in even if I feel fear or doubt.

38. I admit I am wrong, even if I feel uncomfortable doing so.

39. I am willing to be vulnerable in relationships.

40. I say and do what I think is right even when I am afraid.

41. I check with others about how I am doing, even if I might get negative feedback.

42. I check with others about my effect on them.

43. I easily give up when I feel there are obstacles to what I want.

44. I persevere when moving forward is difficult.

45. I will not back down from conflict if it leads me towards what I value.

46. I will risk feeling uncomfortable un the service of improving my relationships with others.

47. If there is an important reason to face something that's uncomfortable for me, I will face it.

48. Showing feelings is a sign of weakness.

### **Love**

49. I act in ways that help others trust me.

50. I do not act in caring ways towards others.

51. I engage in compassionate actions towards others when they are in need.

52. I create moments of warmth and connection with others.

53. I express love towards those I care about.

54. I express tenderness and caring towards others.

55. I feel a selfless caring for most of human kind.

56. I let other people know that I understand how they feel when they are struggling.

57. I am able to intimately connect with certain people in my life.

58. I only care about other people when they can do things for me.

59. I really do not love anybody.

60. I support others when they need it.

61. I feel compassion for people even when I do not know them.

62. I am able to express love and caring to others just with my eyes and face in key moments.

63. I am able to say "I love you" to the right people, at the right times, in my life.

64. When people close to me share that they love me, I share my love back to them.

---

Table 3.  
Standardized Loadings and Standard Errors of the Final ACL Scale

ACL Scale Component	ACL Scale Factor			
	Factor Loading (SE)	Factor	Variance	R <sup>2</sup>
Item 1	.75	Other-awareness	0.54	0.56
Item 2	.67 (.03)	Other-awareness	0.77	0.45
Item 3	.62 (.04)	Other-awareness	0.90	0.39
Item 4	.62 (.04)	Other-awareness	0.82	0.38
Item 5	.60 (.04)	Other-awareness	0.86	0.37
Item 6	.68	Self-awareness	0.73	0.46
Item 7	.65 (.05)	Self-awareness	0.83	0.42
Item 8	.66 (.04)	Self-awareness	0.76	0.43
Item 9	.61 (.05)	Self-awareness	0.81	0.38
Item 10	.61 (.05)	Self-awareness	0.95	.38
Item 11	.73	Courage	0.67	0.53
Item 12	.68 (.04)	Courage	0.85	0.46
Item 13	.67 (.04)	Courage	0.88	0.45
Item 14	.60 (.05)	Courage	0.97	0.36
Item 15	.59 (.04)	Courage	1.07	0.31
Item 16	.83	Love	0.49	0.68
Item 17	.79 (.04)	Love	0.52	0.62
Item 18	.78 (.03)	Love	0.66	0.62
Item 19	.78 (.03)	Love	0.58	0.61
Item 20	.74 (.03)	Love	0.56	0.55
Other-awareness	–	–	0.69	–
Self-awareness	–	–	0.62	–
Courage	–	–	0.75	–
Love	–	–	1.06	–

Note: Standardized factor loadings are presented. Huber-White robust standard errors were calculated. All p-values are < .001.

Table 4.  
Correlations Among the Factors of the ACL Scale

ACL Scale Factor and Label	ACL Scale Factor			
	1	2	3	4
Factor 1: Other-Awareness	–			
Factor 2: Self-Awareness	.85	–		
Factor 3: Courage	.58	.60	–	
Factor 4: Love	.80	.67	.59	–

Table 5.  
Results of Measurement Invariance Tests for the ACL Scale

	Overall Fit Indices										Comparative Fit Indices							
	Y-B $\chi^2$	df	RMSEA	90% CI	CFI	TLI	SRMR	AIC	BIC	$\Delta Y-B\chi^2$	$\Delta df$	p	$\Delta RMSEA$	$\Delta CFI$	$\Delta TLI$	$\Delta SRMR$	$\Delta AIC$	$\Delta BIC$
Configural	1,618.90	820	.05	.049-.055	.93	.92	.05	100,336.00	102,150.30	--	--	--	--	--	--	--	--	--
Weak	1,723.66	884	.05	.048-.055	.92	.92	.06	100,340.70	101,803.10	101.75	64	.001	.00	.00	.00	.01	4.74	-374.12
Strong	2,037.99	948	.06	.053-.059	.90	.90	.06	100,572.00	101,682.50	314.34	64	< .001	.01	-.02	-.02	.00	231.23	-120.62

Note: All model comparisons of the Y-B $\chi^2$  statistics are significant at the  $p < .001$  level. Only the  $\Delta Y-B\chi^2$ , CFI, AIC, and BIC are statistically valid model comparison indices - other fit indices can be compared, but informally.  $\Delta$ -values are calculated prior to rounding.  $\Delta Y-B\chi^2$  = Yuan-Bentler robust chi-square statistic; df = degrees of freedom; RMSEA = Root Mean Square Error of Approximation; CI = confidence interval; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; SRMR = Standardized Root Mean Residual; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

Table 6.  
Test-Retest Correlations and Mean ACL Scale Scores Across 1-Week Interval

ACL Scale Component	Mean (SD)		Pearson's <i>r</i>
	Time 1	Time 2	
Total Score	5.19 (0.74)	5.12 (0.74)	.75
Other-Awareness	5.23 (0.86)	5.16 (0.80)	.57
Self-Awareness	5.25 (0.86)	5.19 (0.80)	.53
Courage	4.88 (0.90)	4.86 (0.88)	.66
Love	5.41 (1.01)	5.25 (1.00)	.73

Table 7.  
Convergent and Discriminant Validity Correlations

	ACL Scale Component				
	Total Score	Other-awareness	Self-awareness	Courage	Love
Social Connectedness <sup>A</sup>	<b>.32</b>	0.28	0.19	0.23	0.34
Loneliness <sup>B</sup>	<b>-.40</b>	-.34	-.25	-.31	-.40
Emotional Intelligence <sup>C</sup>	<b>.74</b>	.64	.58	.54	.64
Social Support <sup>D</sup>	<b>.52</b>	.44	.37	.35	.54
Frequency of Intimacy-related Behaviors <sup>E</sup>	<b>.37</b>	.30	.27	.25	.39
Partner Intimacy <sup>F</sup>	<b>.24</b>	.20	.15	.16	.25
Perspective Taking <sup>G</sup>	.47	<b>.43*</b>	.36	.37	.38
Social Intelligence <sup>H</sup>	.20	<b>.17</b>	.20	.11	.18
Observe Skills <sup>I</sup>	.40	.34	<b>.34</b>	.31	.31
Anxiety in Interpersonal Settings <sup>J</sup>	-.22	.27	.20	<b>.32*</b>	.19
Motivation to Pursue Goals <sup>K</sup>	.17	.11	.07	<b>.24*</b>	.14
Fear of Intimacy <sup>L</sup>	-.46	-.37	-.29	<b>-.36</b>	-.46
Non-reactivity to Inner Experience <sup>M</sup>	.29	.27	.20	<b>.32*</b>	.19
Empathic Concern <sup>N</sup>	.53	.43	.39	.32	<b>.56*</b>
Compassionate Love <sup>O</sup>	.46	.39	.29	.34	<b>.48*</b>

Note: Predicted convergent correlations are boldfaced. \* = Confirmed prediction. All correlations are significant at a Bonferonni-adjusted alpha criterion of .00067. A = Social Connectedness Scale; B = UCLA Loneliness Scale; C = Schutte Self-Report Emotional Intelligence Test; D = MOS Social Support Survey; E = Functional Analytic Psychotherapy Intimacy Scale; F = Miller Social Intimacy Scale; G = Interpersonal Reactivity Index – Perspective Taking subscale; H = Reading the Mind-in-the-Eyes Test; I = Five Facet Mindfulness Questionnaire – Observe subscale; J = Interpersonal Reactivity Index – Personal Distress subscale; K = Behavioral Activation Scale – Drive Subscale; L = Fear of Intimacy Scale; M = Five Facet Mindfulness Questionnaire – Nonreactivity subscale; N = Interpersonal Reactivity Index – Empathic Concern subscale.

Appendix A  
Correlations Between ACL Scale Items in the Initial Item Pool

ACL Item	SD	ACL Item																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	1.20	-																				
2	1.18	.42	-																			
3	1.18	.35	.38	-																		
4	1.11	.36	.52	.47	-																	
5	1.21	.29	.36	.31	.49	-																
6	1.26	.28	.43	.41	.33	-																
7	1.16	.43	.40	.38	.42	.36	.4	-														
8	1.15	.39	.40	.36	.41	.35	.38	.55	-													
9	1.17	.31	.30	.36	.36	.35	.33	.42	.48	-												
10	1.16	.48	.42	.36	.37	.33	.34	.46	.44	.38	-											
11	1.26	.36	.39	.30	.37	.28	.45	.36	.35	.31	.45	-										
12	1.37	-.13	-.37	-.15	-.27	-.21	-.25	-.17	-.17	-.09	-.18	-.21	-									
13	1.45	-.09	-.10	.01	-.16	-.31	-.15	-.08	-.08	-.09	-.05	-.10	.27	-								
14	1.55	-.12	-.22	-.03	-.23	-.25	-.18	-.14	-.12	-.06	-.15	-.15	.35	.47	-							
15	1.50	.24	.36	.16	.33	.26	.28	.29	.34	.31	.29	.21	-.25	-.15	-.24	-						
16	1.21	.24	.37	.39	.37	.31	.28	.28	.28	.22	.32	.32	-.15	.00	-.04	.23	-					
17	1.33	.29	.29	.24	.37	.41	.30	.33	.28	.30	.29	.27	-.10	-.29	-.19	.22	.25	-				
18	1.14	.38	.37	.34	.43	.39	.36	.39	.40	.36	.36	.35	-.19	-.19	-.22	.27	.34	.45	-			
19	1.24	.24	.35	.34	.43	.32	.29	.30	.31	.28	.31	.31	-.17	-.07	-.09	.21	.46	.33	.42	-		
20	1.16	.30	.32	.35	.39	.39	.33	.34	.35	.34	.35	.37	-.15	-.14	-.12	.27	.38	.42	.5	.54	-	





## Correlations Between ACL Scale Items in the Initial Item Pool (cont.)

ACL Item	SD	ACL Item			
		61	62	63	64
61	1.28	–			
62	1.36	.46	–		
63	1.58	.32	.49	–	
64	1.32	.38	.47	.64	–

## Appendix B

## Residual Correlation Matrix/Covariance Matrix

ACL Item	1	2	3	4	5	6	7	8	9	10
1	0.00/0.16	1.56	2.53	-4.34	-3.01	-1.42	0.24	-2.25	4.23	2.9
2	1.56	0.00/0.00	-3.45	-1.1	-0.48	-1.45	-0.03	-2.36	1.51	0.09
3	2.53	-3.45	0.00/0.61	-1.66	-0.83	-1.02	4.14	0.85	4.69	-0.32
4	-4.34	-1.10	-1.66	0.00/0.00	26.79	-0.76	-1.06	-2.34	5.62	-0.38
5	-3.01	-0.48	-0.83	26.79	0.00/0.00	-2.87	-3.07	-2.64	4.57	-0.62
6	-1.42	-1.45	-1.02	-0.76	-2.87	0.00/0.41	-0.12	0.00	-4.38	0.78
7	0.24	-0.03	4.14	-1.06	-3.07	-0.12	0.00/0.00	-0.94	-1.26	-2.07
8	-2.25	-2.36	0.85	-2.34	-2.64	0.00	-0.94	0.00/0.00	-4.57	-0.66
9	4.23	1.51	4.69	5.62	4.57	-4.38	-1.26	-4.57	0.00/0.49	3.04
10	2.90	0.09	-0.32	-0.38	-0.62	0.78	-2.07	-0.66	3.04	0.00/0.00

Note: Values below the diagonal are residual correlations; values above the diagonal are residual covariances. Format for the diagonal = residual covariance/residual correlation

ACL Item	11	12	13	14	15	16	17	18	19	20
11	0.00/0.00	1.09	0.28	-1.4	0.95	-2.93	1.93	-0.68	-2.08	0.95
12	1.09	0.00/0.00	-0.79	-3.25	0.55	3.81	6.3	4.89	2.71	3.43
13	0.28	-0.79	0.00/0.3	5.14	-0.55	-5.58	-1.15	-3.17	-1.89	-0.3
14	-1.4	-3.25	5.14	0.00/0.25	-0.26	0.1	2.55	1.08	-1.18	2.22
15	0.95	0.55	-0.55	-0.26	0.00/0.24	-2.39	0.75	-1.89	-0.14	-0.03
16	-2.93	3.81	-5.58	0.1	-2.39	0.00/0.00	-3.14	12.48	0.00	-3.06
17	1.93	6.3	-1.15	2.55	0.75	-3.14	0.00/0.00	-3.35	2.17	9.73
18	-0.68	4.89	-3.17	1.08	-1.89	12.48	-3.35	0.00/0	-1.42	-1.5
19	-2.08	2.71	-1.89	-1.18	-0.14	0.00	2.17	-1.42	0.00/-0.30	-4.11
20	0.95	3.43	-0.3	2.22	-0.03	-3.06	9.73	-1.5	-4.11	0.00/0.42

Note: Values below the diagonal are residual correlations; values above the diagonal are residual covariances. Format for the diagonal = residual covariance/residual correlation.