TPM®

SELF-REPORT EMPATHY SCALES LACK CONSISTENCY: EVIDENCE FROM EXPLORATORY AND CONFIRMATORY FACTOR ANALYSIS

CONRAD BALDNER
SAPIENZA UNIVERSITY OF ROMA

JARED J. MCGINLEY
TOWSON UNIVERSITY

Empathy is a construct with a long history of definitional variability, which is reflected in the variety of scales designed to measure it. A recent investigation involved a series of analyses to explore constructs assessed in the self-report of empathy and to illuminate the inconsistency in measurement across multiple scales. The current investigation was designed to extend this line of inquiry by focusing solely on cognitive and affective subscales of empathy. A sample of 855 undergraduates completed empathy questionnaires that contained both an affective and a cognitive subscale. A confirmatory factor analysis revealed poor fit for affective and cognitive empathy factors while an exploratory factor analysis revealed several factors that are not essential for empathy. To navigate current barriers for building a cohesive body of literature, future empathy researchers are encouraged to clearly define their conceptualization of empathy and carefully select a measure to best reflect their definition.

Key words: Affective empathy; Cognitive empathy; Measurement; Exploratory factor analysis; Confirmatory factor analysis.

Correspondence concerning this article should be addressed to Conrad Baldner, Department of Social and Developmental Psychology, Sapienza University of Roma, Via dei Marsi 78, 00185 Roma (RM), Italy. Email: conrad.baldner @uniroma1.it

Self-report is currently the most common method employed to measure empathy (Ilgunaite, Giromini, & Di Girolamo, 2017). However, there are many self-report measures from which researchers can choose, and these measures can reflect different ways of thinking about empathy. Many researchers have already illuminated the poor conceptual consistency across definitions of empathy (e.g., Baldner & McGinley, 2016; Cuff, Brown, Taylor, & Howat, 2016; Reik, 1948; Wispé, 1987), and a recent study has shown that several of the most commonly-used self-report empathy scales reflect this inconsistency by including items that capture constructs that are separable from empathy (Baldner & McGinley, 2014). These issues can present challenges for navigating the empathy research literature: instead of an agreed upon measure derived from a consensus definition, there is vast literature that contains a multitude of measures that reflect a variety of definitions. Consequently, individuals who follow the literature could reasonably develop a confusing or misleading picture of empathy.

Although the Baldner and McGinley (2014) study clearly demonstrated several problems inherent to the contemporary measurement of empathy, there are limitations in what can be drawn from these findings. Notably, this study included self-report measures that reflected a variety of empathy definitions, which could partially explain the diversity of items. Although we can conclude from these results that there is poor consistency among empathy measures, it is possible that measures could be more consistent when they reflect similar definitions of empathy. The current study sought to extend this line of investigation by solely selecting empathy measures that ostensibly reflect similar underlying definitions. Since empathy definitions frequently include constructs for affective and cognitive empathy — approximately, the feeling and understanding of others' emotions (Gerdes, Segal, & Lietz, 2010; Shamay-Tsoory, Aharon-Peretz, & Perry, 2009) — the current study focused solely on self-report measures with both affective and cognitive empathy subscales. Our objective was to explore, through confirmatory and exploratory factor analyses, the factor structure that underlied our selection of empathy questionnaires. Ideally, we should observe that good fit is provided by cognitive and affective empathy factors, given that our selection of measures were designed to measure these constructs. If we do not observe this pattern of results, we will then explore the other factors that could underlie our data. In the following sections, we will briefly summarize the state of empathy conceptualization and measurement and present the research design for the current study.

THE STATE OF EMPATHY CONCEPTUALIZATION

The conceptual confusion within the empathy literature has been noted by many researchers over the past seven decades (e.g., Baldner & McGinley, 2016; Cuff et al., 2016; Reik, 1948; Wispé, 1987). Recently, Cuff and colleagues (2016) employed a snowball sampling approach to the literature and noted themes for disagreement among empathy definitions. One of the example themes they noted was that empathy could be defined primarily as a cognitive process, an affective process, or both. This inconsistency is reflected in self-report measures. It is common, however, for measures to include subscales for both affective and cognitive empathy, although measures may differ in the labels for these subscales. Dividing empathy into affective and cognitive components is not necessarily the most accurate conception of empathy; we will return to this point in the discussion. Nonetheless, if measures of affective and cognitive empathy use similar kinds of items, then the empathy literature could at least present a useful and consistent operationalization of empathy.

Unfortunately, the issues that plague the conceptualization of empathy as a unitary construct also largely affect the subcomponents of affective and cognitive empathy. The affective empathy literature has long debated issues such as the degree to which the emotion response must be congruent with the target's emotion state, whether the respondent must keep track of the source of the emotion, and whether a self-other distinction is necesssary or must be maintained (Baldner & McGinley, 2016; Batson, 2009; Cuff et al., 2016; Decety & Jackson, 2004). Similarly, the literature on the cognitive components of empathy have been debated on issues such as the necessity for, and the role of, cognition in empathy, as well as its component elements. For example, some theorists have reduced the role of cognitive empathy to mere perception-action coupling, while others view cognitive empathy as primarily involving the perspective taking of others (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Decety, 2005). For perspective taking to take place, one must possess a theory of mind (ToM), which itself, has even been deconstructed into subcomponents of cognitive ToM and affective ToM (Shamay-Tsoory & Aharon-Peretz, 2007). It has also been argued if direct perception of the target stimuli is necessary (e.g., empathy for fictional characters; Cuff et al., 2016; Singer & Lamm, 2009). In addition to the variability in conceptualizations of these subcomponents, it is often challenging to cleanly separate cognitive and affective empathy due to their functional interdependence. For example, perspective taking can often be the vehicle for eliciting some types of affect, and emotions often are paired with — or lead to — associated cognitions (e.g., thoughts and memories). Clearly, the presence of so many disparate views of these subcomponents has led to inconsistencies in how these constructs are measured.

PROBLEMS WITH THE SELF-REPORT MEASUREMENT OF EMPATHY

Although many self-report scales divide empathy into affective and cognitive subscales (e.g., Basic Empathy Scale; Jolliffe & Farrington, 2006), this is not a universal practice. Some of the earliest empathy questionnaires were single-factor measures (Hogan, 1969; Mehrabian & Epstein, 1972). Later measures consisted of two, four, or even five subscales (Davis, 1983; Jolliffe & Farrington, 2006; Lietz et al., 2011). There is even a measure of empathy that is composed of only a single item (Konrath, Meier, & Bushman, 2018). Some measures also include clearly nonessential subscales, such as personal distress (i.e., a self-focused response to others in need; Davis, 1983), emotion regulation, and the societal need to care for others (i.e., the Empathic Assessment Index's empathic attitudes subscale; Lietz et al., 2011). A recent study assessed the degree of consistency across several measures through the exploratory factor analysis of participant responses (Baldner & McGinley, 2014). Results revealed that there was a large proportion of items that did not load on any of several factor models, and several factors emerged that were weakly related to contemporary empathy conceptualizations.

Even though this study was informative, it had notable limitations. Most importantly, it included a broad selection of self-report empathy scales that reflected different conceptualizations. For example, some measures had only a single empathy scale, while others had multiple scales that assessed constructs both essential and nonessential to empathy. For example, the Interpersonal Reactivity Index (Davis, 1983) included a personal distress subscale, even though the construct is inherently self- and not other-focused. It is likely difficult, if not impossible, to find strong relationships between empathy measures that differ in quantity and content of subscales. In other words, the measurement inconsistency in that study could have been the result of conceptual inconsistency. Given that many self-report empathy measures assess affective and cognitive factors, then assessing the consistency of affective and cognitive empathy scales can minimize conceptual inconsistency as an explanation for measurement inconsistency. If there is instead a large degree of measurement inconsistency within affective and cognitive empathy scales, then it is likely that the empathy literature has concerning issues in how empathy is operationally defined in addition to the known issues in its conceptual definitions.

THE CURRENT INVESTIGATION

We used confirmatory factor analysis (CFA) and exploratory factor analysis (EFA) to assess the degree of measurement consistency in self-report affective and cognitive empathy measures. Although factor analysis can be used to propose and support theory (Goldberg & Velicer, 2006), we were not interested in either supporting any existing empathy theory nor in proposing our own. Instead, we reasoned that if a two-factor CFA model (i.e., affective and cognitive empathy) had good fit, then there would be evidence that the selection of empathy scales measured these underlying constructs in a similar way. If there was not good fit, then a two-factor EFA model would be assessed with special attention to poor performing and cross-loading items. If we continued to find poor fit, we then would assess a series of EFA models to investigate other factors that emerged from patterns of participant responding.

METHOD

Participants

Participants in the full sample were undergraduates (N = 855; M_{age} = 19.8; 74.6% female) at a large university in the Mid-Atlantic of the United States who were provided course credit for their participation. For eligibility, they must have been enrolled in at least one psychology course at the time of their participation. Six hundred eighty-seven participants (80.35%) reported their ethnicity as White/Caucasian (Non-Hispanic); 98 (11.46%) as Asian/Pacific Islander; 30 (3.51%) as Black/African-American (Non-Hispanic); 21 (2.46%) as Hispanic/Latino; 1 (0.12%) as Native American/Aleut; and 18 (2.11%) reported their ethnicity as "other." The full sample was divided into two subsamples for the EFA (n = 427) and CFA (n = 428). The EFA sample was 74% female (M_{age} = 19.77), and the CFA sample was 75.2% female (M_{age} = 19.69). *T*-tests revealed no significant differences in any demographic or substantive variables between the two subsamples.

Procedure

All data collection occurred via an online survey. Participants completed a short demographics questionnaire and then completed five questionnaires that assessed affective and cognitive empathy. Questionnaires that were included had subscales for both affective and cognitive empathy. These measures are described below. The order of the empathy scales was randomly presented.

Measures

Interpersonal Reactivity Index (IRI; Davis, 1983). The IRI is a 28-item self-report questionnaire that includes four 7-item subscales: empathic concern (EC), perspective taking (PT), fantasy (FS), and personal distress (PD). Davis (1983) argued that the EC (e.g., "I often have tender, concerned feelings for people less fortunate than me") and PT (e.g., "I try to understand my friends better by imagining how things look from their perspective") subscales reflected the affective and cognitive components of empathy, and the scales are often used as proxies for these constructs (e.g., Williams, Morelli, Ong, & Zaki, 2018). The FS and PD subscales were not presented to the participants. Items are responded to on a scale of 1-5, with higher scores indicating higher agreement. The IRI is perhaps the most commonly-used self-report empathy measure (Pelligra, 2011); however, previous researchers have argued that it does not have acceptable model fit (Alterman, McDermott, Cacciola, & Rutherford, 2003, but see also Gilet, Mella, Studer, Grühn, & Labouvie-Vief, 2013). Additionally, other researchers have argued that the EC scale is confounded with sympathy (Jolliffe & Farrington, 2006). The Cronbach's alphas for the EC and PT subscales were .77 and .78 for both the EFA and CFA subsamples.

Basic Empathy Scale (BES; Jolliffe & Farrington, 2006). The BES is a 20-item self-report measure. It consists of a 9-item cognitive empathy subscale (e.g., "I can understand my friend's happiness when she/he does well at something"), and an 11-item affective empathy subscale (e.g., "I often become sad when watching sad things on TV or films"). Items were responded to on a scale of 1-5, with higher scores indicating higher agreement. Unlike other measures, the BES assesses reactions to positive emotions in addition to negative emotions. The BES is often used to measure affective and cognitive empathy (e.g., Mota, et al., 2019). In the EFA subsample, the Cronbach's alphas for the affective and cognitive subscales were .83 and .76. In the CFA subsample, these values were .81 and .73.

Empathy Assessment Index (EAI; Lietz et al., 2011). The EAI is a 17-item self-report measure that consists of five subscales: affective response (AR; three items), self-other awareness (S-OA; three items), emotion regulation (ER; four items), perspective taking (PT; four items), and empathic attitudes (EA; three items). Items are responded to on a 1-6 scale, with higher scores indicating higher agreement. Lietz and colleagues argued that the AR (e.g., "Watching a happy movie makes me feel happy"), and PT (e.g., "I can imagine what it's like to be in someone else's shoes") subscales reflected the affective and cognitive components of empathy (Gerdes, Lietz, & Segal, 2011; Gerdes et al., 2010; Lietz et al., 2011), and they have been used as proxies for these constructs (Greeno, Ting, & Wade, 2018). All items were presented to participants, but only the AR and PT subscales were used in the analyses. In the EFA subsample, the Cronbach's alphas for the AR and PT subscales were .71 and .65. In the CFA subsample, these values were .73 and .65.

The How I Feel in Different Situations Scale (HIFDS; Bonino, Lo Coco, & Tani, 1998; Feshbach & Feshbach, 1991). The HIFDS is a 12-item self-report scale — originally written in Italian and for adolescents — that consists of a 7-item cognitive empathy subscale (e.g., "I am able to understand how people react to the things that I do") and a 5-item affective/emotional empathy subscale (e.g., "When my friend is disappointed, I feel disappointed too"). Items were responded to on a scale of 1-4, with higher scores indicating higher agreement. One item had to be slightly re-worded for the sample: the item "I am able to recognize, before many other children, that other people's feelings have changed" was re-written as "I am able to recognize, before many other people, that other people's feelings have changed." The HIFDS had acceptable face validity after this change. Previous research has used this scale to assess affective and cognitive empathy (Lonigro, Laghi, Baiocco, & Baumgartner, 2014); however, it has not previously been used for a university-age sample. In the EFA subsample, the Cronbach's alphas for the affective empathy and cognitive empathy subscales were .70 and .80. In the CFA subsample, these values were .73 and .78.

Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers, Corcoran, Drake, Shryane, & Völlm, 2011). The QCAE is a 31-item self-report scale that consists of a 19-item cognitive empathy subscale (e.g., "I try to look at everybody's side of a disagreement before I make a decision"), and a 12-item affective empathy subscale (e.g., "I am inclined to get nervous when others around me seem to be nervous"). However, five items from the cognitive scale were taken from other scales that formed a part of this survey. These items were presented to participants, but were not included in correlational and factor analyses. Items were responded to on a 1-4 scale, with higher scores indicating higher agreement. Previous research has used this scale to assess affective and cognitive empathy (Mul, Stagg, Herbellin, & Aspell, 2018). In the EFA subsample, the Cronbach's alphas for the affective empathy and cognitive empathy subscales were .78 and .87. In the CFA subsample, these values were .76 and .86.

RESULTS

Analytic Plan

Participants responded to a total of 94 items, of which 79 were used in the analyses. In order to capture the relationships between these items, we conducted: (1) a correlational analysis with the full sample (N = 855); (2) a CFA (n = 428); and (3) an EFA (n = 427). The results from each analysis are presented below.

Correlational Analysis

Before assessing the correlations we first examined the normality of the data. The absolute kurtosis and skewness statistics for each item-level variable, when divided by their respective standard errors, were often in excess of three (Brown, 2016). Although we investigated the Pearson correlations with all subscales from the study, given this nonnormality, we also reported rank-order correlations (Table 1). As can be seen in Table 1, the differences between the Pearson and rank-order correlations were modest.

As we assessed scales that ostensibly measured the same constructs, it was important to observe whether there was agreement among the measures. There were several notable patterns of correlations that warranted attention. Since each scale included subscales for both affective and cognitive empathy, we initially assessed the intercorrelations between subscales within each measure. Correlations ranged from .28 (QCAE) to .43 (IRI); the mean of the correlations was .36. We then assessed the intercorrelations between scales within each construct. Among the affective empathy subscales, the correlations ranged from .38 (EAI and QCAE) to .73 (BES and QCAE); the average correlation was .51. Among the cognitive empathy subscales, the correlations ranged from .37 (IRI and HIFDS) to .60 (EAI and IRI; BES and HIFDS; HIFDS and QCAE); the average correlation was .51. Given that the HIFDS was not originally designed for an adult sample, we paid particular attention to how its affective and cognitive subscales correlated with the other scales. The correlations between the affective subscale with the other affective subscales ranged from .45 to .65; the mean was .55. The correlations between the HIFDS cognitive subscale with the other cognitive subscales ranged from .37 to .60; the mean was .52.

TABLE 1
Pearson and rank-order correlations

	1	2	3	4	5	6	7	8	9
1. IRI-empathic concern	-	.43	.55	.43	.48	.36	.40	.34	.48
2. IRI-perspective taking	.43	-	.25	.37	.27	.35	.31	.60	.22
3. BES-affective	.56	.25	-	.33	.64	.29	.40	.22	.75
4. BES-cognitive	.42	.40	.36	-	.26	.57	.35	.47	.28
5. HIFDS-affective	.47	.27	.64	.26	-	.31	.43	.27	.63
6. HIFDS-cognitive	.38	.37	.29	.60	.36	-	.37	.52	.24
7. EAI-affective response	.41	.29	.42	.35	.45	.39	-	.38	.37
8. EAI-perspective taking	.32	.60	.20	.49	.29	.54	.41	-	.21
9. QCAE-affective	.46	.20	.73	.27	.65	.25	.38	.22	-
10. QCAE-cognitive	.33	.46	.22	.56	.26	.60	.27	.52	.28

Note. IRI = Interpersonal Reactivity Index BES = Basic Empathy Scale; HIFDS = How I Feel in Different Situations Scale; EAI = Empathy Assessment Index; QCAE = Questionnaire of Cognitive and Affective Empathy.
All p-values are significant at $p < .001$. Pearson (rank-order) correlations are below (above) the diagonal.

Confirmatory Factor Analysis

The CFA portion of the study included several psychometric analyses. Reliability and validity tests were performed to demonstrate the consistency of individual items, illuminate the factor structure, and aid in determining the quality of model fit. Reliability tests for internal consistency (i.e., Cronbach's alphas) were performed in SPSS (v.23) to examine the inter-item relationships. The CFA was performed using Mplus (v.7.31) to identify a measurement model. The default estimation model, maximum likelihood estimation with robust standard errors (MLR), was used in the analysis. Any missing data were handled with the maximum-likelihood method.

We initially attempted to confirm a two-factor model of empathy (i.e., affective and cognitive empathy). Items from each subscale were assigned to the respective factor, and items did not load on more than one factor. The measures differed on response scale; as such, a completely standardized solution was sought. We assumed that all measurement error was uncorrelated. We allowed the latent variables representing affective and cognitive empathy to be correlated; this is consistent with past research that has found a relationship between these constructs (Baldner & McGinley, 2014). Given the large number of items, the model was overidentified with 3001 degrees of freedom. Since our goal was not to eliminate items for the purpose of improving model fit, but to provide insight to the model fit of the items contained in similar scales, we initially deemed the use of a CFA without respecification via modification indices as adequate and meaningful.

The sample correlation matrix was analyzed with MPlus 7.31 with a MLR function (sample correlations and standard deviations can be provided upon request from corresponding author). Goodness of fit was evaluated by the standardized root-mean-square residual (SRMR), root-mean-square error of approximation (RMSEA), and the comparative fit index (CFI). Although previous researchers have warned of an over-reliance on fit indices (Miles & Shevlin, 2007), they can in some cases serve as guidelines for assessing model fit. Hu and Bentler (1999) recommended the following criteria for fit indices: RSMEA $\leq$ .06; SRMR $\leq$ .08; CFI $\geq$.95. The fit indices — $\chi^2(3001) = 6{,}928.66$, $p < .001$; RMSEA = .055; SRMR = .08; CFI = .61 — revealed an inconsistent pattern of fit. However, each fit index assesses a different aspect of fit. For example, RMSEA and SRMR are absolute indices (e.g., χ^2 corrected for sample size), whereas CFI is a comparative index of fit. Specifically, the CFI compares the fit of the hypothesized model to that of the null model (i.e., a model in which the covariances between the input indicators are set to zero). If the null model has fit that is sufficiently high, then the fit of the proposed model may seem poor by comparison. It has been proven mathematically that if the RMSEA of the null model is better than .158, then the CFI cannot be greater than .90 (Kenny, 2015).

In this sample, the fit of the null model approached adequacy: $\chi^2(3081) = 13{,}392.49$, $p < .001$; RMSEA = .088. Since the null model does not have very poor fit, it is impossible for a CFI to show good fit for the proposed model. A model that is a characterized by a poor CFI is symptomatic of poor correlations among indicator variables (Kenny, 2015). This, consequently, is a sign of poor fit that could be caused by low performing items.

Factor loadings for the cognitive and affective items are provided in Appendix 1. All items significantly loaded on their respective factor (error and factor variances can be provided upon request from corresponding author). The standardized factor loadings ranged from .222 to .621 for the cognitive scale and .260 to .677 for the affective scale. It is probable that these lower factor loadings are at least contributing to some of the poor fit (e.g., as seen in the CFI).

Although our intent was only to evaluate the two-factor model and not to respecify the model, we also analyzed modification indices to identify localized areas of strain that could help explain the poor fit. There were 12 modification indices for factor loadings (Appendix 2); these represent opportunities for increased fit by assigning items to load on the opposite factor (e.g., assigning affective empathy items to the cognitive empathy factor). Although the magnitudes of these modification indices are not large, the mere existence of these indices could indicate issues with the development of the component scales. For instance, scale authors created items that ostensibly assess affective empathy but could be interpreted by respondents as a form of cognitive empathy (e.g., “Friends talk to me about their problems as they say that I am very understanding” from the QCAE). Modification indices also suggested that the errors between items loading on the same factor could be correlated in order to improve fit. To address this possibility, we assessed a two-factor model in which the errors from the items on the same factor were correlated. Although fit indices were slightly improved, it still showed a pattern of acceptable RMSEA and SRMR but poor CFI: $\chi^2(2937) = 6{,}605.02$, $p < .001$; RMSEA = .054; SRMR=.079; CFI=.64. The fit indices of the different CFA models are summarized in Table 2.

Given that low correlations could indicate poorly-performing items and could help explain the poor-fitting CFI, we then analyzed the individual correlations within the CFA subsample. Correlations of even low magnitude could be significant given the large sample size. Thus, we flagged all correlations that had *p*-values greater than .01. Correlations among indicators within the affective and cognitive factors were assessed separately. There were a total of 39 items in the affective empathy factor; consequently, there were 741 unique correlations. Ninety-four (12.6%) of these correlations had *p*-values above .01; 51 (6.8%) were nonsignificant. There were a total of 40 items in the cognitive factor, not including duplicated items; consequently, there were 780 unique correlations. One hundred and two of these items (13.07%) had *p*-values above .01; 59 (7.5%) were nonsignificant. To interpret these results, it is necessary to contrast them to the amount of weak correlations to a two-factor model in which the poor-performing items have been removed. If this model has both fewer weak correlations and better fit then there is evidence that the weak correlations among the poor-performing items is at least partially responsible for the poor model fit. We assessed this possibility through an EFA.

TABLE 2
Model summary (CFA)

Model	Fit
Original model	$\chi^2(3001) = 6{,}928.66$ RMSEA = .055 SRMR = .08 CFI = .61
With correlated errors	$\chi^2(2937) = 6{,}605.02$ RMSEA = .054 SRMR = .079 CFI = .64
With all modification indices	$\chi^2(3001) = 7{,}456.88$ RMSEA = .059 SRMR = .093 CFI = .56
With correlated errors and all modification indices	$\chi^2(2635) = 5{,}520.51$ RMSEA = .051 SRMR = .086 CFI = .72

Note. RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual; CFI = comparative fit index. Hu and Bentler (1999) recommend RMSEA ≤. 06; SRMR ≤ .08, CFI ≥ .95.

Exploratory Factor Analysis

The EFAs were conducted with two objectives: (1) to assess the two-factor (i.e., affective and cognitive empathy) model without low-performing items, and (2) to investigate the other factors that could underlie the data. The EFAs were conducted with MPlus 7.31. We used a MLR estimator with an oblimin rotation, which allowed for correlations between factors. According to the method of Muthén and Muthén (2009), any nonloading item was removed from the analysis, and the model was re-run. All models were re-run until there were no nonloading items. We report the following fit indices: χ^2, RMSEA, SRMR, CFI, and the Tucker-Lewis Index (TLI). The latter two fit indices are typically used for CFA (Bentler, 1990). However, since comparative fit indices (e.g., CFI, TLI) and absolute fit indices (e.g., RMSEA, SRMR) assess fit in different ways, both types of fit indices can provide useful information.

We assessed if a two-factor EFA model conformed to affective and cognitive empathy factors, if the model had improved fit relative to the two-factor CFA model, and if any improved fit was associated with a reduced number of weak correlations among indicators. A parallel analysis (Horn, 1965) indicated that up to a seven-factor model could be supported (Figure 1); however, the six- and seven-factor models did not converge. There is not a consensus on the criteria for acceptable factor loadings, so we followed the criteria of Norman and Streiner (1994) that was subsequently followed by Baldner and McGinley (2014). Item loadings of less than .40 were considered insufficient, loadings between .40 and .60 were considered moderate, and loadings above .60 were considered strong. Baldner and McGinley (2014) also did not consider items that loaded above .30 on more than one factor, so we also used this criterion.

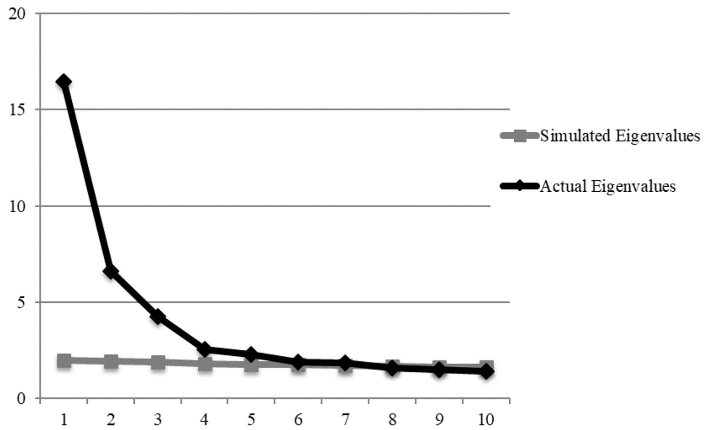

FIGURE 1
Actual eigenvalues superimposed over eigenvalues simulated by parallel analysis.
The *x*-axis represents number of factors. The *y*-axis represents eigenvalues.
The actual eigenvalues for the first seven factors are greater than the corresponding simulated eigenvalues, indicating that up to a seven-factor model is acceptable.

The two-factor model was summarized in Table 3 and the factor loadings can be found in Appendix 3. This model consisted of 46 of the 79 total items (58.2%). In this model, the RMSEA (.054) and SRMR (.052) represent adequate fit. Although stronger, the CFI (.805) and TLI (.786) still indicated poor fit. The first factor from this model explained 24.1% of the variance and consisted of 27 items. Items on this factor solely came from affective subscales. The highest loading items had a focus on being affected by others' emotional states (e.g., "I get caught up in other people's feelings easily"). Although some items explicitly referred to negative emotional states (e.g., "When somebody I care about is sad, I feel sad too"), no items specifically referred to positive emotional states. Items typically referred to participants' direct experiences (e.g., "People I am with have a strong influence on my mood"), but occasionally referred to fiction (e.g., "I often become sad when watching sad things on TV or in films"). We interpreted this broad factor as "affective response."

The second factor from this model explained 11.7% of the variance, and consisted of 19 items. Eighteen items came from cognitive subscales, with an additional item from the QCAE affective subscale (i.e., "Friends talk to me about their problems as they say that I am very understanding"). Many items focused on an individual's ability — or perceived ability — to recognize emotions in others (e.g., "I can tell if someone is masking their true emotion"). The EAI and IRI cognitive subscales — the only scales that had an explicit focus on perspective taking — were *not* represented. We interpreted this factor as "perceived social acuity." The retained factors are clearly derived from the affective and cognitive empathy factors, respectively — albeit excluding perspective taking. The fit, although still poor, was an improvement on the fit in the two-factor CFA model.

TABLE 3
Two-factor EFA: Item distribution

	Factor 1	Factor 2	Loading items (total items)
BES-affective	10 (11)		10 (11)
EAI-affective	0 (3)		0 (3)
HIFDS-affective	4 (5)		4 (5)
IRI-affective	3 (7)		3 (7)
QCAE-affective	10 (12)	1 (12)	11 (12)
BES-cognitive		2 (9)	2 (9)
EAI-cognitive		0 (4)	0 (4)
HIFDS-cognitive		6 (7)	6 (7)
IRI-cognitive		0 (7)	0 (7)
QCAE-cognitive		10 (14)	10 (14)
	27	19	46 (79)

Note. BES = Basic Empathy Scale; EAI = Empathy Assessment Index; HIFDS = How I Feel in Different Situations Scale; IRI = Interpersonal Reactivity Index; QCAE = Questionnaire of Cognitive and Affective Empathy. Factor 1 = affective response; Factor 2 = perceived social acuity.
Table displays number of items which load on each factor, with total number of scale items in parentheses. The row sums represent the total loading items from each scale; the column sums represent the total loading items on each factor.

We next assessed the correlations among the indicators within the affective and cognitive factors. If there were fewer weak correlations relative to results from the CFA model, then there would be evidence that at least some of the poor fit from the CFA was caused by poorly-performing items. Again, we considered any correlation with a p-value greater than .01 as weak. There were 27 items that loaded on the affective response factor; consequently, there were 351 unique correlations. Only one correlation pair — between items on the HFIDS and QCAE — had a p-value greater than .01 ($r = .12$, $p = .012$). Further, there were 19 items that loaded on the perceived social acuity factor. Of the 171 possible correlations, none had p-values worse than .01. As such, there is evidence that the low-performing items, including items that did not load on the two-factor EFA, led to low correlations between items, and was consequently responsible for some — but not all — of the inadequate fit.

Considering the poor fit of the two-factor exploratory model, we explored additional models. As in the two-factor EFA, a summary of each of model can be found on Table 4. The five-factor model was the model with the best performance across fit indices and is reported in further detail (see Table 5 and Appendices 4 and 5). This model consisted of 48 of the 79 total items (60.7%). The RMSEA (.039) and SRMR (.035) represent adequate fit; however, the CFI (.898) and TLI (.872) represent an improved, but still inadequate, fit.

TABLE 4
Model summary (EFA)

Model	Fit	Factors
1	$\chi^2(902) = 3{,}105.996$ RMSEA = .076 CFI = .614 TLI = .596 SRMR = .089	F1: General empathy
2	$\chi^2(944) = 2{,}099.586$ RMSEA = .054 CFI = .805 TLI = .786 SRMR = .052	F1: Affective response F2: Perceived social acuity
3	$\chi^2(1482) = 2{,}911.005$ RMSEA = .048 CFI = .812 TLI = .791 SRMR = .047	F1: Affective response F2: Perspective taking F3: Perceived social acuity
4	$\chi^2(737) = 1{,}244.665$ RMSEA = .040 CFI = .901 TLI = .879 SRMR=.037	F1: Emotional contagion F2: Perspective taking F3: Perceived social acuity F4: Emotional interest in fiction
5	$\chi^2(898) = 1{,}480.003$ RMSEA = .039 CFI = .898 TLI = .872 SRMR = .035	F1: Perspective taking F2: Emotional contagion F3: Perceptiveness to happiness F4: Perceived social acuity F5: Emotional involvement in fiction

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

The first factor in this model explained 20.9% of the variance, and consisted of 12 items. Items on this factor had a focus on perspective taking in neutral contexts (e.g., "I find it easy to put myself in somebody else's shoes"). Thus, we interpreted this factor as "perspective taking." The second factor from this model explained 10.7% of the variance, and consisted of 15 items. Items from this factor solely came from affective subscales. Items from the EAI and the IRI — the scales that had affective scales with an explicit focus on empathic concern — were not represented. Although some items were framed in a neutral context (e.g., “I get caught up in other people’s emotions easily”), others were specifically framed in a negative context (e.g., “Seeing a friend crying makes me feel as if I am crying too.”); no items were solely in a positive context. Loading items had a focus on “catching” emotions without explicit acknowledgment of the self-other distinction. As such, this factor was more consistent with forms of emotional contagion than affective empathy (Hatfield, Cacioppo, & Rapson, 1994). Consequently, we interpreted this factor as “emotional contagion.”

The third factor in this model explained 6.64% of the variance and consisted of only three items. Each item explicitly had a focus on positive emotion (e.g., “I can understand my friend’s happiness when she/he does well at something”). Two of the items began with “I can” — as did several items from the factor we called “perceived social acuity” that appeared on the two-, three-, and four-factor models — however, no items on the current factor loaded on the former factors. This factor is best interpreted as representing “perceptiveness to happiness.”

The fourth factor in this model explained 3.8% of the variance and consisted of 13 items. This factor solely consisted of items from cognitive subscales. These items had a focus on the individual’s belief in their ability to understand others’ behavior in social situations. As with the factor we called “perspective taking,” these items were framed in a neutral context (e.g., “I can easily tell if someone else is interested or bored with what I am saying”). This factor was very similar to the cognitive factor in the two-factor EFA; as such, we called this factor “perceived social acuity.”

The fifth factor explained 3.2% of the variance and consisted of five items. This factor solely consisted of items from affective scales. Each item on this factor explicitly focused on reactions to works of fiction, and were framed in a neutral context — e.g., “I usually stay emotionally detached when watching a film” (reversed). Similar to the factor we called “emotional contagion,” these items focused on “catching” the emotions of others, but only with characters from fiction. As such, this factor was interpreted as “emotional involvement in fiction”; a similar factor was found by Baldner & McGinley (2014).

TABLE 5
Five-factor EFA: Item distribution

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Loading items (total items)
BES-affective		5 (11)			0 (11)	5 (11)
EAI-affective		0 (3)			0 (3)	0 (3)
HIFDS-affective		3 (5)			1 (5)	4 (5)
IRI-affective		0 (7)			0 (7)	0 (7)
QCAE-affective		7 (12)			3 (12)	10 (12)
BES-cognitive	0 (9)		2 (9)	1 (9)		3 (9)
EAI-cognitive	2 (4)		0 (4)	0 (4)	1 (4)	3 (4)
HIFDS-cognitive	0 (7)		1 (7)	2 (7)		3 (7)
IRI-cognitive	7 (7)		0 (7)	0 (7)		7 (7)
QCAE-cognitive	3 (14)		0 (14)	10 (14)		13 (14)
	12	15	3	13	5	48 (79)

Note. BES = Basic Empathy Scale; EAI = Empathy Assessment Index; HIFDS = How I Feel in Different Situations Scale; IRI = Interpersonal Reactivity Index; QCAE = Questionnaire of Cognitive and Affective Empathy; Factor 1= perspective taking; Factor 2 = emotional contagion; Factor 3 = perceptiveness to happiness; Factor 4 = perceived social acuity; Factor 5 = emotional involvement in fiction.
Table displays number of items which load on each factor, with total number of scale items in parentheses. The row sums represent the total loading items from each scale; the column sums represent the total loading items on each factor.

DISCUSSION

Given the poor conceptual consistency in the empathy literature (e.g., Cuff et al., 2016), the objective of the current study was to assess the quality of measurement for existing self-report scales for affective and cognitive empathy. This undertaking was implemented via confirmatory and exploratory factor analyses of the affective and cognitive empathy scales from five frequently used questionnaires. The results demonstrated poor fit for cognitive and affective factor structures in a CFA and the emergence of several nonessential empathy factors from the EFAs. These findings can be interpreted as poor convergent and content validity within existing cognitive and affective empathy scales. Importantly, many of the factors that emerged in the EFA are not consonant with contemporary conceptualizations of empathy (Baldner & McGinley, 2016; Cuff et al., 2016; Decety & Cowell, 2014; Shamay-Tsoory, & Lamm, 2018; Singer & Klimecki, 2014).

Correlational Analysis

Each of the subscales were, unsurprisingly, intercorrelated. However, there were two interesting patterns of results. First, although we made a change to one item in the HFIDS — a reference to "other children" re-written as "other people" — the HFIDS performed similarly to the other subscales, despite being designed for use with adolescents. There is no evidence that this inclusion of this scale disproportionally furthered any observed measurement inconsistency. Second, despite ostensibly assessing the same constructs, the correlations between affective and cognitive subscales were moderate. This trend was particularly pronounced among the cognitive empathy subscales. This was likely due to the broad range of concepts that can be included under the umbrella terms of affective and cognitive empathy. For instance, the EAI and IRI strictly assessed cognitive empathy through a perspective taking subscale, whereas the other subscales included a broader range of attitudes and behaviors in the assessment of cognitive empathy. Researchers who use perspective taking scales (vs. a broader cognitive empathy scale) effectively assess a related yet distinct construct. Although less pronounced, this same trend can be observed among the affective empathy subscales. The EAI and IRI assessed affective empathy through empathic concern subscales, whereas other subscales take a broader approach. Previous works (e.g., Baldner & McGinley, 2016) have concluded that empathic concern is synonymous with sympathy, a construct distinct from empathy. That is, subscales that ostensibly assess the same constructs can instead measure somewhat related constructs, which can explain the moderate intercorrelations in our data. This issue should be troubling for empathy researchers who, after all, must decide on one empathy scale from among a wide selection.

Confirmatory Factor Analysis

The poor fit of the two-factor CFA model also reflects discrepancies in how researchers assess cognitive and affective empathy. Empathy researchers approach the construct from different perspectives, and accordingly, it is expected for this to be reflected in reduced fit. However, we instead found an inconsistent pattern: absolute fit indices (i.e., RMSEA, SRMR) showed acceptable fit, but the CFI showed poor fit. This pattern speaks to several potential issues. First, it is possible that the model could have been improved by assessing other models, including those that adopted recommended modification indices, such as items that could improve fit if they loaded on the opposite factor (e.g., affective items that load on the cognitive factor). If this led to improved fit, then there would be evidence that a simple confusion between by affective and cognitive empathy could explain the relatively poor fit. Additional models also included models that included correlated errors between items that loaded on the same factor. However, none of these additional models were associated with substantially improved fit. There are two additional possibilities that can be informed with EFA: (1) if there are a number of poor fitting items preventing the two-factor CFA from reaching acceptable fit, or if (2) the two-factor model is not a good approach to capture self-report empathy.

Exploratory Factor Analyses

The findings from the EFA models further illuminated several problems with the contemporary measurement of cognitive and affective empathy. The notable construct validity issues are observed in an excessive number of the items that did not load on the two-factor model, and in the emergence of factors, such as perceptiveness to happiness and emotional involvement in fiction, which demonstrate there are many items contained within these scales that assess constructs distinct from contemporary understandings of cognitive and affective empathy. These observations are in line with previous EFAs of existing empathy measures (Baldner & McGinley, 2014). Below we will discuss several insights that emerged from the factor models.

One of the prominent issues to arise in the five-factor model is the separation of empathy with real versus fictitious characters. Although past research has found that reading fictional literature might aid in the development of empathy for *some* individuals under *some* contexts (Bal & Veltkamp, 2013; Mar, Oatley, & Peterson, 2009), and that empathy for real and fictional characters *can* be correlated (Nomura & Akai, 2012), there currently is little evidence that empathy for humans and empathy for fictitious characters relies on the same shared mechanisms. However, it is clear that participants respond to these items in some quantifiably differing way that indicates that there is some categorical separation in how respondents view these situations.

Another measurement issue that was illuminated by the five-factor model is the separation of empathic responding in contexts of positive and negative emotion. Recent work has drawn attention to the distinction — and potentially different outcomes — between individual differences in responses to negative versus positive emotions (Andreychik & Migliaccio, 2015; Morelli, Lieberman, & Zaki, 2015). This distinction is not captured by our included measures of empathy, because most items are focused on neutral or negative states or situations. Additionally, the perceptiveness to happiness factor (which best approximates "positive empathy" among our factors) only accounted for 6.64% of the variance. The current empathy scales allocate a disproportionate number of items for negative affective states, which can result in omitting a large portion of the landscape of empathic responding.

The emergent factor models from the EFA also yield two more important distinctions by the absence of observed factors. Although the research literatures specify that there are separable and distinct processes in affective responding (e.g., affect sharing, emotional contagion, empathic concern; Baldner & McGinley, 2016), and that cognitive empathy is a nuanced and multifaceted construct that can be divisible into separable approaches (e.g., perspective taking and simulation; Baldner & McGinley, 2016; Goldman, 2006), the factors did not reflect these nuances. Instead the additional factors in the five-factor model ultimately captured processes that were not distinctly identifiable as affective or cognitive empathy.

Item-Level Threats to Construct Validity

There are several items across the included measures that threaten basic fundaments of construct validity because of how they are worded. An example of this is "I can usually appreciate the other person's viewpoint, even if I do not agree with it" from the QCAE cognitive scale. In this context, participants could interpret "appreciate" as either *understanding* (as the authors likely intended), or it could be interpreted as *admire* or *being thankful for,* which would draw from a completely different construct. Another clear construct validity issue was the presence of double-barreled questions (i.e., asking two questions in one). An example is "other people tell me I am good at understanding how they are feeling and what they are thinking" which also comes from the QCAE cognitive subscale. This item requires the respondent to assess and respond to feedback they have received from others regarding both their affective and cognitive understanding.

In addition to items that are vague or poorly designed, there were also many items that assessed constructs other than empathy (e.g., sympathy). For example, in the HIFDS' affective subscale, there is the item "I feel sad when something bad happens to a character in a story." The focus of this question is on the action performed against the character and not on the specific affective state that the character is experiencing. The character could be hurt, afraid, or angry, and therefore a response of "sad" would be more in line with sympathy (i.e., empathic concern) or compassion. Another item — from the IRI-PT scale — reads as "I always try to look at everybody's side of a disagreement before I make a decision," which could imply fairness or respect and does not necessarily require perspective taking.

Broader Scale Issues

These findings are somewhat inevitable, given that different measures will assess empathy in different ways, and that these differences will be highlighted by factor analyses. The most important broader issue is that there is a vast selection of available empathy measures, which can reasonably lead to an increase in the number of underlying factors. Furthermore, many of these scales include items that assess constructs distinct from affective and cognitive empathy, despite the measures ostensibly measuring these specific constructs. Although we intentionally selected measures because they included cognitive and affective subscales, we still must acknowledge that even these two constructs do not fully encompass the proposed component constructs of many of the contemporary models of empathy. Largely affirmed by neuroscience literatures, the need for a self-other distinction has been proposed as an essential element of empathy (Decety & Jackson, 2004). Outside of the EAI, no questionnaires distinctly test for this facet of empathy. Additionally, many others have drawn attention to the essential contributing factor of emotion regulation for empathy maintenance. Although this construct is less universally supported as a requisite component of empathy, there is still notable support for it (Decety & Moriguchi, 2007). Once again, the EAI is the only measure that includes a measure of this construct.

Collectively, these self-report measures of empathy suffer from the same problems that most self-report measures encounter. For example, the items included in the current study cannot separate the respondents' belief in their ability from their desire to be empathic. Also, the items do not assess factors such as the respondents' past experience in successfully being empathic, or whether or not the respondents have historically received feedback from others on their empathic success. Worded more succinctly, there is still no inclusion of items to assess one's empathic accuracy. As an extension of the earlier discussion on the functional interdependence of affective and cognitive processes, participant responses to items from scales for these two subcomponents of empathy might co-vary so frequently in experience that they are quantifiably not differentiable — even though, conceptually and physiologically, they are clearly distinct. With the aforementioned problems in the self-report of empathy, it is clear that advances must be made for the continued use of self-report empathy questionnaires.

Conceptualizing Empathy

Affective and cognitive empathy are important because they are often used by empathy researchers, but this does not necessarily mean that the two constructs create the most accurate conception of empathy. Empathy is not an objective phenomenon to be discovered and then have its boundaries clearly defined. It is instead a human-created construct designed to capture a perceived process. The ideal result, therefore, would be for theorists to come to a shared agreement on how to consistently describe empathy. Therefore, we — along with others — have previously argued that it is best to define and understand empathy in its original formulations as posited by Vischer, Titchener, and Lipps (Baldner & McGinley, 2016; Jahoda, 2005; Vincent, 2012; Wispé, 1987). We do, however, also think that it is important to keep it defined in line with research that has preceded it, while still effectively extracting it from similar, yet differentiable, constructs. Consequently, we conceptualize empathy as a process of perceived or actual affect-matching with a maintenance of self-other distinction and knowledge of the source of the emotion. That is, observers ideally match the emotion of a target while realizing that this emotion comes from the target, and not from the self. This state may be "realistic" if the observer actually matches the target's emotion.

This definition pays tribute to the original contexts that it was presented in as a state of affect matching (Lipps, 1903). It also differentiates empathy from the definitions of some developmentalists and ethologists (e.g., de Waal, 2009; Zahn-Waxler, Robinson, & Emde, 1992) that do not include a self-other distinction, and from the affective states such as sympathy that do not require a degree of matching (Eisenberg & Eggum, 2009). Lastly, this definition still fits with many of the more recent, contemporary, and neuroscience-driven conceptualizations (Decety & Cowell, 2014; Singer & Klimecki, 2014).

Although the cognitive components of empathy are clearly important as vehicles for initiating and sustaining empathy, none are consistently present during the state of affective sharing and therefore do not earn pre-ordinate status. Similar to the importance of emotion regulatory strategies that are recruited to maintain other-focused responses and prevent a transition into personal distress, we view cognition as an oft-paired set of processes. However, there are already frameworks to effectively capture these elements (e.g., perspective taking, emotion regulation). Nonetheless, we understand the importance that these processes play in empathy and also value the utility of their measurement.

Recommended Standards for Future Self-Report Assessment of Empathy

The problems with both defining and measuring empathy have been thoroughly explored in this, along with many other, works (Baldner & McGinley, 2016; Batson, 2009; Cuff et al., 2016; Gerdes et al., 2010; Decety & Jackson, 2004; Singer & Lamm, 2009). Of course, studying separable constructs that are nonetheless presented as "empathy" will only further dilute the cohesion within the empathy literature. Given the level of conceptual inconsistency, it is not realistic to expect that all current empathy researchers will agree on the same conception of empathy. Instead, it is important that researchers who include empathy in their models realize the extent of the inconsistency between measures. For instance, empathic concern (as conceptualized by Batson, 2011), empathic accuracy (as conceptualized by Ickes, 1997), empathy for pain (as conceptualized by Singer et al., 2004), and empathic distress (as conceptualized by Hoffman, 2000) each approach empathy and its effects from a different perspective. However, all of these constructs have been described simply as "empathy," without recognizing the many inconsistencies in approach and definition. Empathy researchers should *explicitly state*, at the outset of each manuscript, the approach that they endorse and should not assume that readers will be able to discern which approach was used by a close inspection of the Methods section. Second, we strongly encourage future researchers to carefully select their measure of empathy to reflect their conceptualization of the construct. For example, the IRI-EC subscale, like other measures of empathic concern, is better understood as a measure of sympathy (Baldner & McGinley, 2016; Hawk et al., 2013), which has clear empirical distinctions from empathy (Wispé, 1986). Therefore, continued use of this measure framed as "empathy" will only maintain barriers to progression in the research literature. Following this recommendation should ameliorate, but not eliminate, the confusion surrounding empathy measurement.

Limitations and Future Directions

There are a few limitations that we must address. Participants responded to 79 empathy items, in addition to several demographics questions. Although this is not an excessive number of items, we cannot rule out that fatigue affected participants' responses. This study employed a convenience sample of university students. Although this afforded the necessarily large sample, it lacked diversity in cultural background and age that may have resulted in differences in responding or relationships between constructs (e.g., Beadle, Sheehan, Dahlben, & Gutchess, 2015; Cassels, Chan, Chung, & Birch, 2010). The majority of participants were also women. Even though previous research has found that empathy — measured with a variety of scales that reflect a variety of definitions — is disproportionatley experienced by women (Lennon & Eisenberg, 1987; Rueckert & Naybar, 2008), there is no evidence that the underlying structure of these scales would vary by gender. This study also did not use all available empathy scales with cognitive and affective elements. For instance, the recently developed Empathy Components Questionnaire (ECQ; Batchelder, Brosnan, & Ashwin, 2017) assesses individuals' drive towards, and ability for, affective and cognitive empathy, and differentiates these constructs from other factors, such as social skills and sympathy. The addition of this measure, which was not available during data collection of the present manuscript, could have changed the fit of the CFA, or aided the yield of additional factors in the EFA. Future research can investigate new measures as they become available.

In addition to our recommendations for future researchers to state their definition of empathy and choose their measures accordingly, we have recommendations for next steps in empathy research. For example, we need to assess what these measures actually predict (e.g., do high scores predict actual empathic responding or affect matching?). Also, there is still a dearth of research to assess whether different questionnaires vary in the outcomes that they predict. We also encourage others to consider, but critically evaluate, newly designed measures, such as the Single Item Trait Empathy Scale (SITES; Konrath et al., 2018).

Conclusion

Interest in the self-report of empathy has yielded a multitude of measures with a variety of component scales. The current investigation demonstrated that there is continued variability in constituent constructs, even within cognitive and affective empathy scales. The underlying factors from this study yield factors that extend beyond contemporary conceptualizations and do not reflect the multifaceted understandings of cognitive and affective empathy as they have been presented in broader research literatures. To avoid extending barriers in the consolidation of empathy research findings, we recommend that future researchers clearly define their conceptualization of empathy and carefully select measures that reflect it.

References

Alterman, A. I., McDermott, P. A., Cacciola, J. S., & Rutherford, M. J. (2003). Latent structure of the Davis Interpersonal Reactivity Index in methadone maintenance patients. *Journal of Psychopathology and Behavioral Assessment*, *25*(4), 257-265. doi:10.1023/A:1025936213110

Andreychik, M. R., & Migliaccio, N. (2015). Empathizing with others' pain versus empathizing with others' joy: Examining the separability of positive and negative empathy and their relation to different types of social behaviors and social emotions. *Basic and Applied Social Psychology*, *37*(5), 274-291. doi:10.1080/01973533.2015.1071256

Bal, P. M., & Veltkamp, M. (2013). How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *PloS ONE*, *8*(1), e55341. doi:10.1371/journal.pone.0055341

Baldner, C., & McGinley, J. J. (2014). Correlational and exploratory factor analyses (EFA) of commonly used empathy questionnaires: New insights. *Motivation and Emotion*, *38*(5), 727-744. doi:10.1007/s11031-014-9417-2

Baldner, C., & McGinley, J. J. (2016). Extracting empathy from related constructs: Historical, theoretical, and empirical support. In C. Edwards (Ed.), *Psychology of empathy: New Research* (pp. 69-128). Hauppage, NY: Nova Science Publishers.

Batchelder L., Brosnan, M., Ashwin, C. (2017) The development and validation of the Empathy Components Questionnaire (ECQ). *PloS ONE*, *12*(1), e0169185. doi:10.1371/journal.pone.0169185

Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), Social neuroscience. *The social neuroscience of empathy* (pp. 3-15). Cambridge, MA: MIT Press.

Batson, C. D. (2011). *Altruism in humans*. New York, NY: Oxford University Press.

Beadle, J. N., Sheehan, A. H., Dahlben, B., & Gutchess, A. H. (2015). Aging, empathy, and prosociality. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *70*(2), 215-224. doi:10.1093/geronb/gbt091

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238-246. doi:10.1037/0033-2909.107.2.238

Bonino, S., Lo Coco, A., & Tani, F. (1998). *Empatia. I processi di condivisione delle emozioni* [Empathy: The processes of emotion sharing]. Florence, Italy: Giunti.

Brown, S. (2016, May 22). *Measures of shape: Skewness and kurtosis*. Retrieved from https://brownmath.com/stat/shape.htm

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5497-5502.

Cassels, T. G., Chan, S., Chung, W., & Birch, S. A. J. (2010). The role of culture in affective empathy: Cultural and bicultural differences. *Journal of Cognition and Culture*, *10*(3-4), 309-326. doi:10.1163/156853710X531203

Cuff, B. M., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review*, *8*(2), 144-153.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113-126. doi:10.1037/0022-3514.44.1.113

Decety, J. (2005). Perspective taking as the royal avenue to empathy. In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 143-157). New York, NY: Guilford Press.

Decety, J., & Cowell, J. M. (2014). The complex relation between morality and empathy. *Trends in Cognitive Sciences*, *18*(7), 337-339. doi:10.1016/j.tics.2014.04.008

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*(2), 406-412. doi:10.1177/1534582304267187

Decety, J., & Moriguchi, Y. (2007). The empathic brain and its dysfunction in psychiatric populations: Implications for intervention across different clinical conditions. *BioPsychoSocial Medicine*, *1*, 22-65. doi:10.1186/1751-0759-1-22

De Waal, F. (2009). *Primates and philosophers: How morality evolved*. Princeton, NJ: Princeton University Press.

Eisenberg, N., & Eggum, N. D. (2009). Empathic responding: Sympathy and personal distress. In J. Decety & W. Ickes (Eds.), *Social neuroscience. The social neuroscience of empathy* (pp. 71-83). Cambridge, MA: MIT Press.

Feshbach, S., & Feshbach, N. D. (1991). Aggression and altruism: A personality perspective. In C. Zahn-Waxler, E. M. Cummings, & R. Iannotti (Eds.), *Altruism and aggression: Biological Origins* (pp. 189-217). Cambridge, UK: Cambridge University Press.

Gerdes, K. E., Lietz, C. A., & Segal, E. A. (2011). Measuring empathy in the 21st century: Development of an empathy index rooted in social cognitive neuroscience and social justice. *Social Work Research*, *35*(2), 83-93. doi:10.1093/swr/35.2.83

Gerdes, K. E., Segal, E. A., & Lietz, C. A. (2010). Conceptualising and measuring empathy. *British Journal of Social Work*, *40*(7), 2326-2343. doi:10.1093/bjsw/bcq048

Gilet, A.-L., Mella, N., Studer, J., Grühn, D., & Labouvie-Vief, G. (2013). Assessing dispositional empathy in adults: A French validation of the Interpersonal Reactivity Index (IRI). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *45*(1), 42-48. doi:10.1037/a0030425

Goldberg, L. R., & Velicer, W. F. (2006). Principles of exploratory factor analysis. In S. Strack (Ed.), *Differentiating normal and abnormal personality* (pp. 209-237). New York, NY: Springer Publishing Company.

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading.* New York, NT: Oxford University Press.

Greeno, E. J., Ting, L., & Wade, K. (2018). Predicting empathy in helping professionals: Comparison of social work and nursing students. *Social Work Education*, *37*(2), 173-189. doi:10.1080/02615479.2017.1389879

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Studies in emotion and social interaction. Emotional contagion.* New York, NY: Cambridge University Press; Paris, France: Editions de la Maison des Sciences de l'Homme.

Hawk, S. T., Keijsers, L., Branje, S. J. T., Van der Graaff, J., de Wied, M., & Meeus, W. (2013). Examining the Interpersonal Reactivity Index (IRI) among early and late adolescents and their mothers. *Journal of Personality Assessment*, *95*(1), 96-106. doi:10.1080/00223891.2012.696080

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, *33*(3), 307-316. doi:10.1037/h0027580

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice.* New York, NY: Cambridge University Press.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179-185. doi:10.1007/BF02289447

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55. doi:10.1080/10705519909540118

Ickes, W. J. (1997). *Empathic accuracy*. New York, NY: Guilford Press.

Ilgunaite, G., Giromini, L., & Di Girolamo, M. (2017). Measuring empathy: A literature review of available tools. *BPA – Applied Psychology Bulletin*, *65*(280), 2-28.

Jahoda, G. (2005). Theodor Lipps and the shift from "sympathy" to "empathy". *Journal of the History of the Behavioral Sciences*, *41*(2), 151-163. doi:10.1002/jhbs.20080

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*(4), 589-611. doi:10.1016/j.adolescence.2005.08.010

Kenny, D. A. (2015, November 24). *Measuring model fit.* Retrieved from http://davidakenny.net/cm/fit.htm

Konrath, S., Meier, B. P., & Bushman, B. J. (2018). Development and validation of the single item trait empathy scale (SITES). *Journal of Research in Personality*, *73*, 111-122. doi:10.1016/j.jrp.2017.11.009

Lennon, R., & Eisenberg, N. (1987). Gender and age differences in empathy and sympathy. In N. Eisenberg & J. Strayer (Eds.), *Cambridge studies in social and emotional development. Empathy and its development* (pp. 195-217). New York, NY: Cambridge University Press.

Lietz, C. A., Gerdes, K. E., Sun, F., Geiger, J. M., Wagaman, A. M., & Segal, E. A. (2011). The empathy assessment index (EAI): A confirmatory factor analysis of a multidimensional model of empathy. *Journal of the Society for Social Work Research*, *2*, 104-124. doi:10.5243/jsswr.2011.6

Lipps, T. (1903). Einfuhlung, innere nachahmung und organempfindung [Empathy, inner imitation, and organ sensation]. *Archiv für die Gesamte Psychologie*, *1*, 465-519.

Lonigro, A., Laghi, F., Baiocco, R., & Baumgartner, E. (2014). Mind reading skills and empathy: Evidence for nice and nasty ToM behaviours in school-aged children. *Journal of Child and Family Studies*, *23*(3), 581-590. doi:10.1007/s10826-013-9722-5

Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communications*, *34*(4), 407-428. doi.10.1515/COMM.2009.025

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, *40*(4), 525-543. doi:10.1111/j.1467-6494.1972.tb00078.x

Miles, J., & Shevlin, M. (2007). A time and a place for incremental fit indices. *Personality and Individual Differences*, *42*(5), 869-874. doi:10.1016/j.paid.2006.09.022

Morelli, S. A., Lieberman, M. D., & Zaki, J. (2015). The emerging study of positive empathy. *Social and Personality Psychology Compass*, *9*(2), 57-68. doi:10.1111/spc3.12157

Mota, S., Leckelt, M., Geukes, K., Nestler, S., Humberg, S., Schröder-Abé, M., . . . Back, M. D. (2019). A comprehensive examination of narcissists' self-perceived and actual socioemotional cognition ability. *Collabra: Psychology*, *5*(1), 6. doi:10.1525/collabra.174

Mul, C.-l., Stagg, S. D., Herbelin, B., & Aspell, J. E. (2018). The feeling of me feeling for you: Interoception, alexithymia and empathy in autism. *Journal of Autism and Developmental Disorders*, *48*(9), 2953-2967. doi:10.1007/s10803-018-3564-3

Muthén, L. K., & Muthén, B. (2009). *Mplus short course topic 1. Introductory — Advanced factor analysis and structural equation modeling with continuous outcomes.* Retrieved from http://www.statmodel.com/course_materials.shtml

Nomura, K., & Akai, S. (2012). Empathy with fictional stories: Reconsideration of the fantasy scale of the interpersonal reactivity index. *Psychological Reports*, *110*(1), 304-314. doi:10.2466/02.07.09.11.PR0.110.1.304-314

Norman, G. L., & Streiner, D. L. (1994). *Biostatistics: The bare essentials*. St. Louis, MO: Mosby.

Pelligra, V. (2011). Empathy, guilt-aversion, and patterns of reciprocity. *Journal of Neuroscience, Psychology, and Economics*, *4*(3), 161-173. doi:10.1037/a0024688

Reniers, R. L. E. P., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A Questionnaire of Cognitive and Affective Empathy. *Journal of Personality Assessment*, *93*(1), 84-95. doi:10.1080/00223891.2010.528484

Reik, T. (1948). *Listening with the third ear. The inner experience of a psychoanalyst*. Oxford, UK: Farrar, Straus.

Rueckert, L., & Naybar, N. (2008). Gender differences in empathy: The role of the right hemisphere. *Brain and Cognition*, *67*(2), 162-167. doi:10.1016/j.bandc.2008.01.002

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia*, *45*(13), 3054-3067.

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain: A Journal of Neurology*, *132*(3), 617-627. doi:10.1093/brain/awn279

Shamay-Tsoory, S. G., & Lamm, C. (2018). The neuroscience of empathy-from past to present and future. *Neuropsychologia*. Advance Online Publication. doi:/10.1016/j.neuropsychologia.2018.04.034

Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology*, *24*(18), 875-878. doi:10.1016/j.cub.2014.06.054

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, *1156*, 81-96.
doi:10.1111/j.1749-6632.2009.04418.x

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*(5661), 1157-1162.
doi:10.1126/science.1093535

Vincent, T. C. (2012). From sympathy to empathy: Baudelaire, Vischer, and Early Modernism. *Mosaic: A Journal for the Interdisciplinary Study of Literature*, *45*(1), 1-15.

Williams, W. C., Morelli, S. A., Ong, D. C., & Zaki, J. (2018). Interpersonal emotion regulation: Implications for affiliation, perceived support, relationships, and well-being. *Journal of Personality and Social Psychology*, *115*(2), 224-254.
doi:10.1037/pspi0000132

Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology*, *50*(2), 314-321.
doi:10.1037/0022-3514.50.2.314

Wispé, L. (1987). History of the concept of empathy. In N. Eisenberg & J. Strayer (Eds.), *Cambridge studies in social and emotional development. Empathy and its development* (pp. 17-37). New York, NY: Cambridge University Press.

Zahn-Waxler, C., Robinson, J. L., & Emde, R. N. (1992). The development of empathy in twins. *Developmental Psychology*, *28*(6), 1038-1047.
doi:10.1037/0012-1649.28.6.1038

APPENDIX 1

CFA Standardized Factor Loadings for Affective and Cognitive Empathy

Cognitive Empathy Items	Loading	Affective Empathy Items	Loading
IRI 3 (Cognitive)	*.391*	IRI 2 (affective subscale)	.506
IRI 8 (Cognitive)	.537	*IRI 4 (affective subscale)*	*.327*
IRI 11 (Cognitive)	.518	IRI 9 (affective subscale)	.428
IRI 15 (Cognitive)	.222	IRI 14 (affective subscale)	.460
IRI 21 (Cognitive)	.441	*IRI 18 (affective subscale)*	*.322*
IRI 25 (Cognitive)	.405	IRI 20 (affective subscale)	.566
IRI 28 (Cognitive)	.462	IRI 22 (affective subscale)	.492
BES 3 (Cognitive)	.419	*BES 1 (affective subscale)*	*.574*
BES 6 (Cognitive)	.341	BES 2 (affective subscale)	.564
BES 9 (Cognitive)	.419	BES 4 (affective subscale)	.310
BES 10 (Cognitive)	.230	BES 5 (affective subscale)	.677
BES 12 (Cognitive)	.586	*BES 7 (affective subscale)*	*.588*
BES 14 (Cognitive)	.413	*BES 8 (affective subscale)*	*.612*
BES 16 (Cognitive)	.534	BES 11 (affective subscale)	.619
BES 19 (Cognitive)	*.385*	*BES 13 (affective subscale)*	*.437*
BES 20 (Cognitive)	*.440*	BES 15 (affective subscale)	.373
HIFDS 7 (Cognitive)	.457	BES 17 (affective subscale)	.614
HIFDS 8 (Cognitive)	.565	*BES 18 (affective subscale)*	*.586*
HIFDS 9 (Cognitive)	.492	HIFDS 1 (affective subscale)	.410
HIFDS 10 (cognitive subscale)	.575	HIFDS 2 (affective subscale)	.659
HIFDS 11 (cognitive subscale)	.621	HIFDS 3 (affective subscale)	.670
HIFDS 12 (cognitive subscale)	.593	HIFDS 4 (affective subscale)	.374
EAI 1 (cognitive subscale)	.597	HIFDS 5 (affective subscale)	.486
EAI 10 (cognitive subscale)	.368	HIFDS 6 (affective subscale)	.545
EAI 13 (cognitive subscale)	.487	EAI 5 (affective subscale)	.442
EAI 17 (cognitive subscale)	.587	EAI 7 (affective subscale)	.425
QCAE 15 (cognitive subscale)	.526	EAI 14 (affective subscale)	.421
QCAE 16 (cognitive subscale)	.438	QCAE 2 (affective subscale)	.309
QCAE 18 (cognitive subscale)	.489	QCAE 7 (affective subscale)	.539
QCAE 19 (cognitive subscale)	.586	QCAE 8 (affective subscale)	.286
QCAE 20 (cognitive subscale)	.507	QCAE 9 (affective subscale)	.436
QCAE 21 (cognitive subscale)	.540	QCAE 10 (affective subscale)	.625
QCAE 22 (cognitive subscale)	.502	QCAE 11 (affective subscale)	.426
QCAE 24 (cognitive subscale)	.523	QCAE 12 (affective subscale)	.664
QCAE 25 (cognitive subscale)	.476	QCAE 13 (affective subscale)	.374
QCAE 26 (cognitive subscale)	.548	QCAE 14 (affective subscale)	.395
QCAE 27 (cognitive subscale)	.529	*QCAE 17 (affective subscale)*	*.260*
QCAE 28 (cognitive subscale)	.418	QCAE 23 (affective subscale)	.406
QCAE 30 (cognitive subscale)	.411	*QCAE 29 (affective subscale)*	*.456*
QCAE 31 (cognitive subscale)	.306	-	-

Note: Reversed-scored items were italicized. BES = Basic Empathy Scale; EAI = Empathy Assessment Index; HIFDS = How I Feel in Different Situations Scale; IRI = Interpersonal Reactivity Index; QCAE = Questionnaire of Cognitive and Affective Empathy.

APPENDIX 2

Suggested Modification Indices

Origina factor	Suggested factor	Item	MI	EPC	Std. EPC
Aff	Cog	QCAE 23	60.41	0.718	0.279
Aff	Cog	QCAE 8	31.102	–0.57	–0.222
Aff	Cog	IRI 9	25.368	0.552	0.215
Aff	Cog	BES 15	23.795	–0.689	–0.268
Aff	Cog	QCAE 14	20.742	–0.419	–0.163
Aff	Cog	IRI 2	18.173	0.512	0.199
Aff	Cog	IRI 20	16.109	0.47	0.183
Cog	Aff	EAI 17	14.142	–0.422	–0.19
Aff	Cog	EAI 14	13.895	0.549	0.213
Cog	Aff	QCAE 16	12.922	–0.243	–0.109
Aff	Cog	BES 17	10.911	–0.401	–0.156
Aff	Cog	IRI 18	10.877	0.462	0.18

Note. MI = modification index; EPC = expected parameter change; Std. EPC = standardized expected parameter change. Aff = affective; Cog = cognitive; BES = Basic Empathy Scale; EAI = Empathy Assessment Index; IRI = Interpersonal Reactivity Index; QCAE = Questionnaire of Cognitive and Affective Empathy.

APPENDIX 3

EFA Oblimin-Rotated Pattern Matrix for Two-Factor Model

Item scale (subscale)	F1	F2
BES 5 (affective subscale)	**.729**	–.037
HIFDS 2 (affective subscale)	**.660**	–.008
QCAE 12 (affective subscale)	**.660**	.014
BES 17 (affective subscale)	**.650**	–.083
HIFDS 3 (affective subscale)	**.615**	.138
BES 11 (affective subscale)	**.599**	.066
BES 7 (affective subscale)	***.586***	*–.011*
QCAE 14 (affective subscale)	**.576**	–.182
BES 2 (affective subscale)	**.575**	–.015
QCAE 10 (affective subscale)	**.574**	.004
QCAE 8 (affective subscale)	**.562**	–.180
HIFDS 6 (affective subscale)	**.559**	.046
QCAE 7 (affective subscale)	**.559**	.025
BES 8 (affective subscale)	***.542***	*.102*
BES 18 (affective subscale)	***.541***	*.212*
BES 15 (affective subscale)	**.537**	–.114
QCAE 9 (affective subscale)	**.534**	–.022
HIFDS 5 (affective subscale)	**.509**	.058
QCAE 11 (affective subscale)	**.487**	.040
IRI 20 (affective subscale)	**.481**	.172
IRI 14 (affective subscale)	***.471***	*.071*
QCAE 13 (affective subscale)	**.465**	–.028
IRI 22 (affective subscale)	**.446**	.125
QCAE 29 (affective subscale)	***.428***	*.048*
BES 4 (affective subscale)	**.426**	–.04
BES 1 (affective subscale)	***.424***	*.111*
QCAE 2 (affective subscale)	***.416***	*–.080*
QCAE 26 (cognitive subscale)	–.042	**.680**
QCAE 22 (cognitive subscale)	–.109	**.675**
QCAE 24 (cognitive subscale)	–.087	**.665**
HFIDS 11 (cognitive subscale)	.059	**.645**
QCAE 20 (cognitive subscale)	–.040	**.633**
QCAE 27 (cognitive subscale)	–.062	**.624**
QCAE 19 (cognitive subscale)	.002	**.620**
QCAE 21 (cognitive subscale)	.067	**.612**
HIFDS 12 (cognitive subscale)	.088	**.609**
HIFDS 8 (cognitive subscale)	.127	**.583**
BES 16 (cognitive subscale)	.048	**.577**
QCAE 15 (cognitive subscale)	.003	**.574**
QCAE 25 (cognitive subscale)	–.062	**.561**
BES 12 (cognitive subscale)	.138	**.546**
QCAE 16 (cognitive subscale)	–.059	**.537**
QCAE 23 (affective subscale)	.178	**.505**
HIFDS 10 (cognitive subscale)	.025	**.501**
BES 20 (cognitive subscale)	*.072*	***.491***
HIFDS 9 (cognitive subscale)	.078	**.442**

Note: Reversed-scored items were italicized. Bold values represent the loading for each factor. BES = Basic Empathy Scale; HIFDS = How I Feel in Different Situations Scale; IRI = Interpersonal Reactivity Index; QCAE = Questionnaire of Cognitive and Affective Empathy; F1 = affective response; F2 = perceived social acuity.

APPENDIX 4

EFA Oblimin-Rotated Pattern Matrix for Five-Factor Model

	F1	F2	F3	F4	F5
IRI 3 (cognitive subscale)	***.536***	*.082*	*.156*	*–.017*	*–.099*
IRI 25 (cognitive subscale)	**.764**	.037	–.176	–.06	.069
IRI 28 (cognitive subscale)	**.685**	–.055	–.001	–.007	.047
EAI 1 (cognitive subscale)	**.647**	.106	.039	.114	–.029
QCAE 18 (cognitive subscale)	**.599**	.041	–.124	.12	–.085
EAI 17 (cognitive subscale)	**.546**	–.09	.186	.07	–.098
IRI 21 (cognitive subscale)	**.534**	–.027	.165	–.032	.104
IRI 11 (cognitive subscale)	**.519**	.034	.23	.066	–.021
QCAE 30 (cognitive subscale)	**.513**	.016	–.029	.066	.058
IRI 8 (cognitive subscale)	**.512**	–.024	.261	.008	.018
IRI 15 (cognitive subscale)	***.446***	-	*–.072*	*–.124*	*–.017*
QCAE 28 (cognitive subscale)	**.425**	–.013	.01	.051	.033
HIFDS 2 (affective subscale)	.011	**.651**	–.017	.04	.044
HIFDS 3 (affective subscale)	.061	**.641**	.041	.113	.003
BES 17 (affective subscale)	.026	**.629**	–.006	–.047	.014
QCAE 14 (affective subscale)	–.041	**.627**	-	–.1	–.082
BES 5 (affective subscale)	.042	**.613**	–.101	.017	.158
QCAE 10 (affective subscale)	–.027	**.604**	–.046	.082	–.001
QCAE 12 (affective subscale)	.002	**.597**	–.029	.055	.136
QCAE 9 (affective subscale)	–.098	**.591**	.105	.03	–.056
BES 2 (affective subscale)	–.011	**.571**	.077	–.001	.013
QCAE 8 (affective subscale)	–.02	**.557**	–.08	–.107	.034
QCAE 7 (affective subscale)	.093	**.549**	–.147	.098	.04
HIFDS 6 (affective subscale)	.134	**.528**	–.072	.037	.073
QCAE 13 (affective subscale)	–.093	**.511**	.239	–.034	–.044
BES 7 (affective subscale)	*.068*	***.486***	*.179*	*–.086*	*.064*
BES 15 (affective subscale)	–.005	**.444**	.032	–.101	.112
BES 14 (cognitive subscale)	.106	.003	**.576**	.062	.081
BES 3 (cognitive subscale)	.057	.037	**.477**	.129	.068
BES 20 (cognitive subscale)	*.025*	*.007*	***.417***	*.286*	*.058*
QCAE 22 (cognitive subscale)	–.066	–.047	.005	**.697**	.003
QCAE 20 (cognitive subscale)	–.093	.023	.028	**.689**	.012
QCAE 26 (cognitive subscale)	.032	–.056	–.064	**.687**	.098
QCAE 24 (cognitive subscale)	–.047	–.081	.066	**.635**	.092
QCAE 27 (cognitive subscale)	.12	–.031	–.05	**.606**	–.006
QCAE 19 (cognitive subscale)	.114	.101	.016	**.605**	–.127
QCAE 16 (cognitive subscale)	–.024	–.018	–.12	**.605**	.042
QCAE 21 (cognitive subscale)	.11	.122	.017	**.572**	–.058
QCAE 25 (cognitive subscale)	.042	.022	–.031	**.564**	–.089
QCAE 15 (cognitive subscale)	–.088	.049	.194	**.559**	–.01
HIFDS 8 (cognitive subscale)	.175	.07	–.009	**.494**	.119
HIFDS 11 (cognitive subscale)	.010	.073	.236	**.488**	.047
BES 12 (cognitive subscale)	.133	.138	.073	**.469**	.001
QCAE 11 (affective subscale)	–.055	.056	–.076	.05	**.788**
QCAE 29 (affective subscale)	*.026*	*.072*	*.077*	*–.015*	***.554***
HIFDS 5 (affective subscale)	.088	.149	.113	–.048	**.549**
EAI 13 (affective subscale)	.183	–.081	.241	.127	**.463**
QCAE 2 (affectice subscale)	*–.087*	*.175*	*.081*	*–.086*	***.413***

Note: Reversed-scored items were italicized. Bold values represent the loading for each factor. BES = Basic Empathy Scale; EAI = Empathy Assessment Index; HIFDS = How I Feel in Different Situations Scale; IRI = Interpersonal Reactivity Index; QCAE = Questionnaire of Cognitive and Affective Empathy; F1= perspective taking; F2 = emotional contagion; F3 = perceptiveness to happiness; F4 = perceived social acuity; F5 = emotional involvement in fiction.

Appendix 5

Bivariate Correlations of Factors in Five-Factor Model

	1	2	3	4
1. Perspective taking	-			
2. Emotional contagion	.23	-		
3. Perceptiveness to happiness	.20	.08	-	
4. Perceived social acuity	.35	.19	.29	-
5. Emotional involvement in fiction	.15	.47	.10	.21