Routledge
Taylor & Francis Group

The Toronto Empathy Questionnaire: Scale Development and Initial Validation of a Factor-Analytic Solution to Multiple Empathy Measures

R. Nathan Spreng,[1,*] Margaret C. McKinnon,[2,3,*] Raymond A. Mar,[4] and Brian Levine[1,5,6]

[1]*Rotman Research Institute at Baycrest Centre, Toronto, Ontario, Canada*
[2]*Mood Disorders Program, St. Joseph's Healthcare, Hamilton, Ontario, Canada*
[3]*Department of Psychiatry and Behavioural Neuroscience, McMaster University, Hamilton, Ontario, Canada*
[4]*Department of Psychology, York University, Toronto, Ontario, Canada*
[5]*Department of Psychology, University of Toronto, Toronto, Ontario, Canada*
[6]*Department of Medicine (Neurology), University of Toronto, Toronto, Ontario, Canada*

To formulate a parsimonious tool to assess empathy, we used factor analysis on a combination of self-report measures to examine consensus and developed a brief self-report measure of this common factor. The Toronto Empathy Questionnaire (TEQ) represents empathy as a primarily emotional process. In 3 studies, the TEQ demonstrated strong convergent validity, correlating positively with behavioral measures of social decoding, self-report measures of empathy, and negatively with a measure of Autism symptomatology. Moreover, it exhibited good internal consistency and high test–retest reliability. The TEQ is a brief, reliable, and valid instrument for the assessment of empathy.

Empathy is an important component of social cognition that contributes to one's ability to understand and respond adaptively to others' emotions, succeed in emotional communication, and promote prosocial behavior. The term *empathy* is derived from Titchener's (1909; Wispé, 1986) translation of the German word *Einfühlung,* meaning "feeling into" (Wispé, 1987). Generally speaking, it refers to the consequences of perceiving the feeling state of another as well as the capacity to do so accurately. Despite the prominence of the empathy construct in developmental research (Sagi & Hoffman, 1976; Ungerer, 1990; Zahn-Waxler, Friedman, & Cummings, 1983) and cross-species investigations of empathic capabilities (Masserman, Wechkin, & Terris, 1964; Rice & Gainer, 1962), a clear, consensual definition of the construct of empathy remains elusive.

Recent research into empathy has emphasized the distinction between cognitive and emotional components of the construct (Preston & de Waal, 2002). These components assume various definitions. Put simply, however, emotional empathy is commonly thought of as an emotional reaction (e.g., compassion) to another's emotional response (e.g., sadness). This reaction is not dependent on a cognitive understanding of why a person is suffering (Rankin, Kramer, & Miller, 2005), although it may facilitate understanding and action. By contrast, cognitive empathy involves an intellectual or imaginative apprehension of another's emotional state, often described as overlapping with the construct of theory of mind (understanding the thoughts and feelings of others) and used interchangeably by some authors (Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004). Numerous authors have focused on distinguishing empathy from

the related concepts of emotional contagion, sympathy, and perspective taking surveyed in some self-report measures of empathy (Omdahl, 1995; Wispé, 1986, 1987). Whereas emotional contagion (also referred to as personal distress) involves the perceiver assuming the emotional state of the target, sympathy is thought to reflect a state of "feeling sorry" for the target with or without an associated behavioral response (Preston & de Waal, 2002). Perspective taking, in contrast, involves the apprehension of another's thought and feeling states through the assessment of visual, auditory, or situational cues (Rankin et al., 2005), without any personal emotional response.

Agreement among researchers and theoreticians on the interrelated processes contributing to empathy has been elusive. Although the processes described previously (perspective taking, sympathy, personal distress, emotional contagion, theory of mind) are referred to as *empathic,* there is little agreement in the literature as to whether they are distinct from empathy as an accurate affective insight into the feelings of another or are facets of a central process required for empathic responding. Indeed, the current corpus of self-report measures of empathy reflects these differing constructs, resulting in significant heterogeneity among measures (Ickes, 1997). In the face of such heterogeneity, one useful approach may be to ask what is common among these different conceptions, allowing one to examine the consensus, or core, opinion on this important process.

It is important to note that a multifaceted measure may be preferable in some situations. We are not proposing that multi-factorial approaches be replaced with a unidimensional measure or that empathy itself be viewed as a single, homogenous construct. Rather, the field of empathy measurement lacks a sufficient tool for examining this construct at the broadest level, and it is this gap that we endeavor to remedy. A useful parallel may be drawn with early intelligence research, which has suffered a similar period of confusion populated by multiple conceptions. When a single underlying factor was extracted from the multiple

tests, this "*g* factor" proved a useful tool in intelligence research (Spearman, 1904). Moreover, the utility of *g* was not achieved at the cost of other conceptions of intelligence. With a similar aim, we sought to derive a single-factor representation of the currently heterogeneous empathy construct to create a useful tool for empathy research that can complement, rather than replace, current multifactorial approaches. Importantly, this consensus measure was derived statistically, using factor analysis, rather than through intuition.

CURRENT SELF-REPORT MEASURES OF EMPATHY

The Empathy Scale (Hogan, 1969), one of the first measures to achieve widespread use, contains four separate dimensions: social self-confidence, even-temperedness, sensitivity, and nonconformity. A psychometric analysis of the scale, however, indicates questionable test–retest reliability and low internal consistency along with poor replication of its previously hypothesized factor structure (Froman & Peloquin, 2001). Indeed, several authors have suggested that the four factors measured by this scale are better suited to the measurement of social skills, broadly speaking, than a central tendency toward empathic behavior (Baron-Cohen & Wheelwright, 2004; Davis, 1983). Hogan's (1969) Empathy Scale has been widely employed as a measure of cognitive empathy (e.g., Eslinger, 1998) but has recently been supplanted in popularity by the Interpersonal Reactivity Index (IRI; Davis, 1983), discussed following.

The Questionnaire Measure of Emotional Empathy (QMEE; Mehrabian & Epstein, 1972) reemphasizes the original definition of the empathy construct (Titchener, 1909; Wispé, 1986). The scale contains seven subscales that together show high split-half reliability, indicating the presence of a single underlying factor thought to reflect affective or emotional empathy. Mehrabian, Young, and Sato (1988) suggested more recently, however, that rather than measuring empathy per se, the scale more accurately reflects general emotional arousability. In response, an unpublished, revised version of the measure, the Balanced Emotional Empathy Scale (Mehrabian, 2000) taps respondents' reactions to others' mental states (cf. Lawrence et al., 2004).

The IRI (Davis, 1983) contains four subscales: Perspective Taking and Fantasy in addition to Empathic Concern and Personal Distress—each pair purported to tap cognitive and affective components of empathy, respectively. As pointed out by Baron-Cohen and Wheelwright (2004), however, the Fantasy and Personal Distress subscales of this measure contain items that may more properly assess imagination (e.g., "I daydream and fantasize with some regularity about things that might happen to me") and emotional self-control (e.g., "In emergency situations I feel apprehensive and ill at ease"), respectively, than theoretically derived notions of empathy. Indeed, the Personal Distress subscale appears to assess feelings of anxiety, discomfort, and a loss of control in negative environments. Factor analytic and validity studies have suggested that the Personal Distress subscale may not assess a central component of empathy (Cliffordson, 2001). Instead, Personal Distress may be more related to the personality trait of neuroticism, whereas the most robust components of empathy appear to be represented in the Empathic Concern and Perspective Taking subscales (Alterman, McDermott, Cacciola, & Rutherford, 2003).

Other self-report measures of empathy have been developed to target specific populations. These include the Scale of Ethnocultural Empathy (Wang et al., 2003), the Jefferson Scale of Physician Empathy (Hojat et al., 2001), the Nursing Empathy Scale (Reynolds, 2000), the Autism Quotient (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) and the Japanese Adolescent Empathy Scale (Hashimoto & Shiomi, 2002). Although these instruments were designed for use with specific groups, aspects of these scales may be suitable for assessing a general capacity for empathic responding. That is, all of these diverse scales touch on an aspect of empathy, broadly speaking.

The Autism Quotient (Baron-Cohen, Wheelwright, Skinner, et al., 2001) was developed to measure Autism spectrum disorder symptoms. Baron-Cohen, Wheelwright, Skinner, et al. viewed a deficit in theory of mind as the characteristic symptom of this disease (Baron-Cohen, 1995) and a number of items from this measure relate to broad deficits in social processing (e.g., "I find it difficult to work out people's intentions."). Thus, any measure of empathy should exhibit a negative correlation with this measure. The magnitude of this relation, however, will necessarily be attenuated by the other aspects of the Autism Quotient, which measure unrelated constructs (e.g., attentional focus and local processing biases).

Additional self-report measures of social interchange appearing in the neuropsychological literature contain items tapping empathic responding including the Dysexecutive Questionnaire (Burgess, Alderman, Evans, Wilson, & Emslie, 1996) and a measure of emotion comprehension developed by Hornak, Rolls, and Wade (1996). These scales focus on the respondent's ability to identify the emotional states expressed by another (e.g., "I recognize when others are feeling sad."). Current theoretical notions of empathy emphasize the requirement for understanding of another's emotions to form an empathic response (Bernieri, 2001). Only a small number of items on current measures of empathy, however, assess this ability.

In this study, we attempted to formulate a consensus among the many scales in use to gauge the empathy construct. Using exploratory factor analysis (EFA), we forced the items to load onto a single factor, thereby assembling a group of highly related items from across many measures of empathic responding, bringing about a unidimensional factor of empathy. Our aim was to identify what is common among different conceptions of empathy as operationalized by published measures of this construct. In a series of three studies, we constructed the Toronto Empathy Questionnaire (TEQ) and demonstrated the TEQ's construct validity through associations with behavioral and self-report measures of interpersonal sensitivity as well as its internal consistency and test–retest reliability.

STUDY 1

We began by submitting responses to every self-report measure of empathy we were able to identify to an EFA, determining what were common across these previously published measures. Items were forced to load on to a single factor, forming the basis of our questionnaire that was then examined for factorial integrity, internal consistency, and reliability.

Methods

Participants.   A total of 200 University of Toronto under-
graduates (100 female), mean age 18.8 years ($SD = 1.2$), partici-
pated for course credit in a psychology course, satisfying general
recommendations for sample size in factor analysis aimed at de-
termining the stability of component patterns (Guilford, 1954;
Russell, 2002). We carefully observed a balance of genders for
initial scale development.

Materials.   We conducted a review of the literature with the
aim of collecting all available measures related, even tangen-
tially, to the self-report of empathic processes or the assess-
ment of deficits in empathic ability. Questions were selected
from several published self-report empathy measures including
the IRI (28 items; Davis, 1983), Hogan's Empathy Scale (15
items; Hogan, 1969), the QMEE (nine items; Mehrabian & Ep-
stein, 1972), a reworded Balanced Emotional Empathy Scale
(12 items; Mehrabian, 2000), the Scale of Ethnocultural Empa-
thy (four items; Wang et al., 2003), Jefferson Scale of Physician
Empathy (six items; Hojat et al., 2001), Nursing Empathy Scale
(eight items; Reynolds, 2000), Japanese Adolescent Empathy
Scale (10 items; Hashimoto & Shiomi, 2002), and the Measure
of Emotional Intelligence (three items; Schutte et al., 1998), for
a total of 95 items after redundant questions were removed. An
additional 36 questions were composed based on the literature
concerning individuals with altered empathic responding due to
neurological or psychiatric disease, with the addition of mod-
ified items from the Dysexecutive Questionnaire (four items;
Burgess et al., 1996) and a measure of emotion comprehension
developed by Hornak et al. (seven items; 1996). Factor analy-
sis with 200 participants and 142 items yielded an independent
observation-to-item ratio of 1.4:1 that exceeds the minimum
1.2:1 ratio capable of recovering a population factor structure
(Barrett & Kline, 1981; see MacCallum, Widaman, Zhang, &
Hong, 1999).

To ensure consistency across sampled items, we reworded
questions to assess frequency of behavior rather than to pose
general statements or tendencies. Responses were given using
a 5-point Likert scale corresponding to various levels of fre-
quency (i.e., *never, rarely, sometimes, often, always*) as opposed
to agreement with individual statements, a method used in sev-
eral of the scales described previously.

Two additional self-report measures were administered in
their entirety to establish convergent and discriminant validity:
the IRI, comprising four subscales of seven items each (Davis,
1983) and the 50-item Autism Quotient (Baron-Cohen, Wheel-
wright, Skinner, et al., 2001). We expected the subscales of the
IRI to be positively related to the TEQ, given that these sub-
scales reflect the content of the majority of empathy measures.
Within this measure, we predicted that the Empathic Concern
subscale would show the strongest association with the TEQ
followed by the Perspective Taking subscale, as these subscales
are thought to map closely onto emotional and cognitive con-
structs of empathy. We did not expect the Fantasy and Personal
Distress subscales of this measure to show a strong associa-
tion with the TEQ given their close relation to imagination and
emotional self-control (Baron-Cohen & Wheelwright, 2004).
Finally, we predicted that the Autism Quotient would be nega-
tively related to the TEQ, as it measures a degree of deficit in
social processing. We expected this relation to be moderated,

however, by the presence of items in this scale unrelated to
empathic responding.

Statistical analysis.   We determined a consensus account
of empathy using an EFA examining the structure of intercor-
relations among items. We conducted an iterated principal-axis
factor analysis with squared multiple correlations of each item
with all other items as the initial communality estimates on
responses for each item. We forced items from this EFA to
load onto a single factor. To devise a unidimensional empa-
thy questionnaire that maximized item-remainder coefficients
and factor loadings, we eliminated items that had low item-
remainder coefficients (below .30), those that failed to improve
internal consistency, and items possessing factor loadings lower
than .40. We then conducted a second EFA with the 16 retained
items to more completely document the factor structure of the
questionnaire.

We then assessed convergent and discriminant validity of the
newly devised 16-item TEQ by calculating Pearson correlations
with the IRI and the Autism Quotient. We assessed gender dif-
ferences in the TEQ by an independent samples t test and by
calculating the effect size with Cohen's d. We also determined
correlations between the IRI subscales and the Autism Quotient.

Results and Discussion

Initial eigenvalues greater than 1 and their variance explained
are provided in Table 1. A total of 41 factors with an eigen-
value greater than 1 suggested a multiplicity of factors in the
self-report of empathy and related constructs (according to
the Kaiser criteria). Conducting an EFA with a forced single
factor yielded 55 items with loadings above .40, drawing on
items from each scale. When more than 10 items load at .40
or above, a single component can be considered a stable rep-
resentation of the population parameter with this sample size
(Guadagnoli & Velicer, 1988; Stevens, 2002). To form a brief
scale, we then culled these 55 items with loadings above .40
to maximize internal consistency and item-remainder coeffi-
cients. This process led to the formation of the 16-item TEQ
(see Appendix). The TEQ contains an equal number of posi-
tively and negatively worded/scored items from a number of
different scales as well as newly composed items (Table 2).
Unidimensional factor loadings ranged from .41 to .65 ($M =
.51$, $SD = .07$; Table 2).

Item-remainder coefficients were sound, ranging from .36 to
.59 (Table 2); internal consistency was also good: Cronbach's
$\alpha = .85$. In a second EFA of the 16-item TEQ, the first 5 eigen-
values were 5.23, 1.43, 1.13, 1.06, and 0.93. There is a dis-
continuity between the first and second factor consistent with
a unidimensional structure. Factor coefficients are reported in
Table 2 in which the items were forced to load on a single fac-
tor, ranging from .42 to .65 ($M = .53$, $SD = .08$). This analysis
yielded four items with loadings above .60, an indication that
the factor is reliable regardless of sample size (Guadagnoli &
Velicer, 1988; Stevens, 2002). We further explored the factor
structure of the newly formed TEQ in an independent sample in
Study 2.

Participants' total scores on TEQ items positively correlated
with the IRI subscale Empathic Concern, $r = .74$, $p < .001$.
Four items within the TEQ were reworded Empathic Concern
subscale items. When these items were removed from the TEQ

TABLE 1.—Exploratory factor analysis results.

Factor	Eigenvalue	% of Variance	Cumulative %
1	19.30	13.59	13.59
2	8.46	5.96	19.55
3	6.80	4.79	24.34
4	4.51	3.18	27.52
5	3.98	2.81	30.32
6	3.61	2.54	32.86
7	3.33	2.35	35.21
8	2.94	2.07	37.28
9	2.74	1.93	39.21
10	2.60	1.83	41.04
11	2.41	1.70	42.73
12	2.37	1.67	44.40
13	2.21	1.56	45.96
14	2.16	1.52	47.48
15	2.03	1.43	48.91
16	1.96	1.38	50.29
17	1.90	1.34	51.63
18	1.84	1.30	52.92
19	1.79	1.26	54.18
20	1.74	1.22	55.40
21	1.69	1.19	56.59
22	1.68	1.18	57.78
23	1.63	1.15	58.92
24	1.56	1.10	60.03
25	1.56	1.10	61.12
26	1.51	1.06	62.19
27	1.50	1.05	63.24
28	1.46	1.03	64.27
29	1.42	1.00	65.27
30	1.41	0.99	66.27
31	1.34	0.94	67.21
32	1.26	0.88	68.09
33	1.22	0.86	68.95
34	1.19	0.84	69.79
35	1.16	0.82	70.61
36	1.15	0.81	71.42
37	1.12	0.79	72.21
38	1.12	0.79	72.99
39	1.07	0.75	73.75
40	1.07	0.75	74.50
41	1.03	0.73	75.23

total score, the correlation remained high, $r = .71$, $p < .001$, suggesting that TEQ items used to measure empathy tap a construct similar to that measured by the Empathic Concern subscale of the IRI. The TEQ had a lower, but still positive, correlation with the IRI subscale of Perspective Taking despite containing no items from this scale, $r = .35$, $p < .001$. Thus, our measure of the broadest level of empathy, although clearly closer to an emotional measure of empathy, still captured variance associated with a more cognitive measure of empathy.

The TEQ scores exhibited a negative correlation with the Autism Quotient, as hypothesized, $r = -.30$, $p < .001$. Individuals scoring highly on our measure tended to report less social processing and communication difficulties as assessed by the Autism Quotient. As predicted, the magnitude of this association was not as great as that for the IRI in which the Autism Quotient measures other symptoms of this disorder not specifically related to social functioning and thus not expected to relate systematically to our measure of empathy. Relations to the Autism Quotient are intended only to demonstrate divergence with related, although conceptually quite different, measures. Means and standard deviations of all measures can be found in Table 3.

TABLE 2.—Toronto Empathy Questionnaire sources and psychometric attributes.

					Factor Loadings	
			Item Remainder			
					Study 1 142-	Study 1 16-
Item	Source	Study 1	Study 2	Study 3	Item EFA	Item EFA
1	Hashimoto & Shiomi, 2002; Mehrabian, 2000	.41	.38	.40	.46	.48
2	Davis, 1983	.54	.43	.66	.56	.59
3	Mehrabian, 2000	.54	.47	.70	.60	.64
4	Hornak, Rolls, & Wade, 1996; Mehrabian, 2000	.47	.56	.54	.50	.47
5	Hogan, 1969	.52	.54	.63	.57	.61
6	Davis, 1983	.43	.59	.64	.47	.50
7	Mehrabian & Epstein, 1972	.43	.40	.47	.47	.50
8	Hornak, Rolls, & Wade, 1996	.36	.37	.34	.41	.44
9	Hogan, 1969	.40	.39	.37	.44	.48
10	Mehrabian, 2000	.42	.47	.36	.45	.47
11	New item	.48	.51	.43	.51	.44
12	New item	.58	.59	.51	.62	.65
13	Hogan, 1969; Mehrabian, 2000	.59	.71	.71	.66	.62
14	Davis, 1983	.50	.46	.53	.55	.58
15	Mehrabian & Epstein, 1972	.40	.39	.38	.44	.42
16	Davis, 1983	.47	.43	.35	.53	.53

Note. EFA = exploratory factor analysis. All items from the Interpersonal Reactivity Index (Davis, 1983) were derived from the Empathic Concern subscale.

We observed no effect of gender in this sample of the TEQ (Table 4), suggesting that men and women provided equivalent responses on our measure.

The IRI subscales also demonstrated significant associations with the Autism Quotient. Consistent with the theory of mind deficits associated with Autism spectrum disorder, the IRI subscale Perspective Taking was negatively associated with the Autism Quotient, $r = -.23$, $p < .01$. A positive association, however, was observed between the IRI subscale Personal

TABLE 3.—Means and standard deviations of self-report and behavioral measures for all studies.

	Study 1		Study 2		Study 3	
Measure	M	SD	M	SD	M	SD
Toronto Empathy Questionnaire	44.54	7.70	47.27	7.48	46.95	7.47
Interpersonal Responsivity Index						
Empathic Concern	3.81	0.63	3.89	0.70		
Personal Distress	2.62	0.77	2.65	0.67		
Perspective Taking	3.48	0.71	3.45	0.69		
Fantasy	3.49	0.82	3.69	0.87		
Autism Quotient	18.28	5.24			17.94	4.42
Reading the Mind in the Eyes Task			27.24	3.20		
Interpersonal Perception Task			9.81	1.72		
Empathy Quotient					40.69	10.36

Note. All means and standard deviations are within the normal ranges of healthy adults previously reported (cf. Davis, 1980; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Costanzo & Archer, 1994; Baron-Cohen & Wheelwright, 2004). Interpersonal Responsivity Index subscale means are reported as the mean Likert (1–5) score.

Distress and the Autism Quotient, $r = .36$, $p < .01$. This association suggests that individuals reporting greater emotional arousability report greater difficulties with social processing and communication and may not represent a core component of empathy. Additionally, there was a slight negative or no relationship with the other subscales: Empathic Concern, $r = -.10$, $p > .10$; and Fantasy, $r = -.02$, $p > .75$. The low association between the Autism Quotient and Empathic Concern suggests that the subscale's construct of empathy is unrelated to self-reported proficiency in social processing and communication. We examined the relationship between self-reported empathy and social processing more explicitly in Study 2.

STUDY 2

From the current corpus of heterogeneous self-report measures of empathy, we identified items that together assess a common construct of empathy. This led to the creation of a unidimensional empathy questionnaire, the TEQ, which possesses high internal consistency and demonstrated convergent and discriminant validity. In a second study, we aimed to further demonstrate the TEQ's factorial integrity, internal consistency, and expand on its construct validity.

In processing interpersonal information, an empathic individual must discriminate and interpret stimuli relevant to the goals of social processing. This interpersonal information must subsequently be interpreted accurately to facilitate the task of responding in an empathic fashion (Bernieri, 2001). We assessed the relation of the TEQ to two behavioral measures that also require the processing of complex interpersonal stimuli: The Reading the Mind in the Eyes Test–Revised (MIE; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) and the Interpersonal Perception Task–15 (IPT–15; Costanzo & Archer, 1994). Together, these measures assess processes that are described commonly in the theoretical literature surrounding empathic accuracy (e.g., emotion comprehension, perspective taking; Sagi & Hoffman, 1976; Ungerer, 1990; Zahn-Waxler et al., 1983).

The utility of any self-report measure is improved greatly if associations can be found with task-based measures (which in this case are presumably less influenced by factors such as socially desirable responding). Indeed, scores on a valid scale of empathy should be systemically related to the correct identification and comprehension of social stimuli as assessed by these measures. However, most self-report measures of empathy are not systematically associated with performance on interpersonal sensitivity tasks (e.g., Ickes, 1997) except in rare instances when other factors, such as the targets' trait expressivity, is taken into account (Zaki, Bolger, & Ochsner, 2008). Here, we predicted that in assessing the broadest level of empathy, the TEQ would have more success in predicting empathic performance than did these measures.

Method

Participants. A total of 79 University of Toronto students (55 female) aged, on average, 18.9 years ($SD = 3.0$) participated for course credit in psychology.

Materials. In addition to the newly formed TEQ, this new sample of participants completed the IRI (Davis, 1983), the MIE (Baron-Cohen, Wheelwright, Hill, et al., 2001) and the IPT–15 (Costanzo & Archer, 1994).

The MIE is an adult test of mentalizing that presents respondents with 36 still pictures of actors' eye regions and asks which of four possible mental states the person currently possesses (Baron-Cohen, Wheelwright, Hill, et al., 2001). All participants were presented with a list of terms used in the task and were provided with the opportunity to read an explanation and example for each. This list of terms and definitions remains with each participant throughout testing for reference. Correct responses on the MIE indicate an ability to understand and pair mental-state terms with static nonverbal cues. High functioning individuals with Asperger's syndrome or autism perform worse on this measure compared to age- and IQ-matched controls, indicating that the test is sensitive to rather subtle individual differences in social perception (Baron-Cohen, Wheelwright, Hill, et al., 2001).

The IPT–15 is a video containing 15 unscripted interactions between two or more individuals (Costanzo & Archer, 1994). Following each vignette, a multiple-choice question is presented that has an objective and true answer (e.g., "Who is the child of the two adults?"). Respondents must closely attend to dynamic nonverbal cues (e.g., prosody, posture, gesture, etc.) to select the correct answer. The answer to this question is never explicitly conveyed. Participants reliably score significantly above chance, and scores on this measure are highly correlated with peer ratings of interpersonal sensitivity and social skills (Costanzo & Archer, 1989).

Analysis. We examined the validity of the TEQ by correlating total scores with the IRI subscales, MIE, and IPT–15. We assessed gender differences by an independent samples t test, and the effect size was determined by calculating Cohen's d. As a secondary goal, we again examined the structure of this measure by calculating item-remainder coefficients and Cronbach's alpha. We then employed two tests to reexamine the structural validity of the TEQ. Parallel analysis and Velicer's minimum average partial test (B. P. O'Connor, 2000; Steger, 2006; Velicer, 1976) are statistical methods that enable one to objectively determine the number of factors in a data set. Parallel analysis provides the eigenvalues from a factor analysis of a randomly permuted data set. Here, we performed random permutations of raw TEQ data (matching for sample size, number of items, and scoring range). We then plotted and compared the eigenvalues of the random permutations from the 95th percentile with the real data. The number of factors present in the data was observed at the point of intersection on the scree plot. Next, we performed Velicer's minimum average partial test to determine the number of factors (or components) in the TEQ. In the minimum average partial test, a complete principal components analysis is performed, after which the first principal component is partialled out of the correlations among the variables, and the average squared partial correlation is noted. This procedure was repeated using the first 2 principal components, then the first 3, an so forth. The number of components whose partialling out resulted in the minimum average partial is the number of components related to systematic, rather than unsystematic, variance in the original correlation matrix.

Results and Discussion

The TEQ correlated positively with the IRI subscales of Empathic Concern, $r = .74$, $p < .001$; Perspective Taking, $r = .29$, $p < .01$; and unlike in Study 1, Fantasy, $r = .52$, $p <$

TABLE 4.—Gender differences in TEQ scores across studies.

| | Male | | Female | | | | | |
Study	M	SD	M	SD	t	df	p	Effect Size Cohen's d
1	44.45	8.19	44.62	7.22	0.16	198	0.87	.02
2	43.46	7.79	48.93	6.77	3.16	77	<.05	.73
3	43.63	7.93	48.33	6.90	2.39	63	<.01	.63

Note. TEQ = Toronto Empathy Questionnaire.

.001. Scores on the TEQ also correlated with the behavioral measures of social comprehension: MIE, $r = .35$, $p < .01$; IPT–15, $r = .23$, $p < .05$. This was true even though these two measures themselves were uncorrelated, $r = .08$, $p > .45$. The lack of correlation between the MIE and IPT–15 illustrates the problematic heterogeneity that is commonly observed with regard to empathy measurement (Ickes, 1997) and emphasizes the need for a measure that represents core empathy or what is common among these diverse measures. Furthermore, these associations with behavioral measures of interpersonal sensitivity demonstrate validity extending beyond agreement with other self-report measures. Importantly, the magnitude of these associations is not trivial. The association with the MIE falls within the top third of all effect sizes observed in psychology for measures that do not share method variance, and the correlation with the IPT–15 lies within the middle third (Hemphill, 2003).

Unlike the TEQ, the IRI subscales demonstrated a slight negative or no relationship with the MIE: Empathic Concern, $r = -.15$, $p < .05$; Perspective Taking, $r = -.16$, $p < .01$; Personal Distress, $r = -.14$, $p < .05$; and Fantasy, $r = -.06$, $p > .30$. Additionally, the IRI exhibited statistically nonsignificant relationships with the IPT–15, which were weaker but similar in value to the TEQ: Empathic Concern, $r = .17$, $p > .10$; Perspective Taking, $r = .20$, $p > .05$; Personal Distress, $r = -.11$, $p > .30$; and Fantasy, $r = .10$, $p > .40$. Thus, although the TEQ was highly related to the Empathic Concern subscale of the IRI (Study 1), it performed better than the IRI when predicting actual social cognitive performance on measures related to empathic accuracy.

Unlike Study 1, gender differences were observed in this sample (Table 4). Consistent with previous self-report measures of empathy (e.g., Davis, 1983), a moderate effect was observed: Women scored higher than men.

Item-remainder coefficients for the TEQ were sound, ensuring that all the items assess the same construct, and ranged from .37 to .71; and internal consistency was good, Cronbach's $\alpha = .85$. An examination of the scree plot of the real and permuted data (Figure 1) indicated that the number of factors in the data set was one. Velicer's minimum average partial test found systematic variance in the TEQ related to a single component, with the smallest average squared correlation of .0231 (Table 5). The parallel analysis and the minimum average partial test provide converging evidence that the TEQ comprises a single factor.

STUDY 3

To explore further the psychometric properties of the TEQ, we once again investigated convergent and discriminant validity through associations with self-report measures of empathy and Autism spectrum disorder symptomatology as well as test–
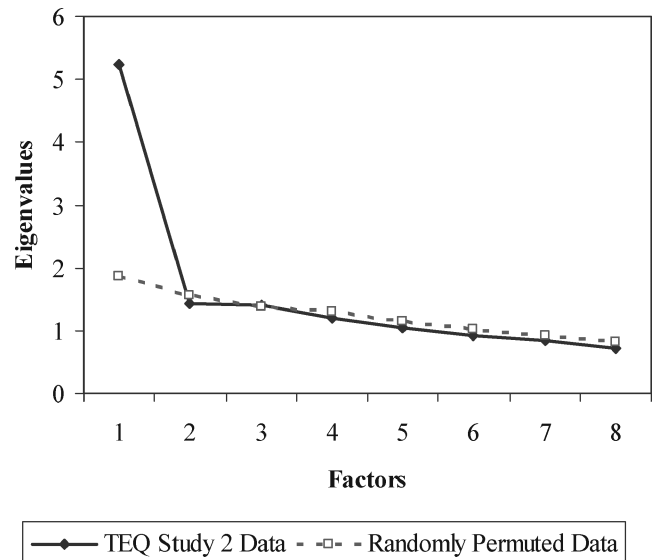

FIGURE 1.—Scree plot of the eigenvalues for Study 2 exploratory factor analysis and randomly permuted raw data. TEQ = Toronto Empathy Questionnaire.

retest reliability on a second set of responses given by returning participants from Study 2. The aim of this study was to extend the findings from Study 1 by examining the relation of the TEQ to additional measures of social cognitive processing related to empathy as well as the stability of our measure over time. We included a new self-report measure of empathy developed by Baron-Cohen and Wheelwright (2004): the Empathy Quotient. The development of this 80-item questionnaire was theoretically driven, and it was evaluated psychometrically on individuals with Asperger's syndrome and matched neurologically intact controls. Because this scale was not available when Study 1 was conducted, it was not included in the original battery given to our respondents. As predicted by Baron-Cohen and Wheelwright (2004), individuals with Asperger's syndrome scored lower on this measure of empathy than controls. We expected

TABLE 5.—Velicer's minimum average partial test results for Study 2.

Components	Average Squared Correlations
0	.0892
1	**.0231**
2	.0272
3	.0309
4	.0372
5	.0474
6	.0556
7	.0668
8	.0835
9	.1020
10	.1360
11	.1689
12	.2236
13	.2930
14	.4761
15	1

Note. The smallest minimum average partial, indicating the number of components in the data, is in bold.

TEQ scores to be positively associated with the Empathy Quotient and negatively associated with the Autism Quotient.

Methods

Participants. A total of 65 University of Toronto students (46 female) aged, on average, 18.6 years ($SD = 2.3$) returned from Study 2 a mean of 66.1 days ($SD = 6.35$, range = 57–84) following their initial participation and received course credit for participating.

Materials. In addition to the TEQ, participants completed the Empathy Quotient (Baron-Cohen & Wheelwright, 2004), and the Autism Quotient (Baron-Cohen, Wheelwright, Skinner, et al., 2001).

Analysis. We examined the validity of the TEQ by correlating its total with the Empathy Quotient and Autism Quotient. Additionally, we calculated item-remainder coefficients and Cronbach's alpha. We determined test–retest reliability by calculating the correlation between returning participants' scores attained during Study 2 and readministration of the TEQ. To assess an effect of attrition, we calculated a paired-samples t test to determine differences in TEQ score between test administrations. We assessed gender differences in the TEQ by an independent samples t test and Cohen's d.

Results and Discussion

As predicted, the TEQ correlated positively with the Empathy Quotient, $r = .80$, $p < .001$, and negatively with the Autism Quotient, $r = -.33$, $p < .01$. Item-remainder coefficients for the TEQ were sound, ranging from .34 to .71 (see Table 1). Moreover, the internal consistency of our measure remained good, $\alpha = .87$. Finally, the TEQ demonstrated high test–retest reliability, $r = .81$, $p < .001$. Differences in TEQ means (Table 3) were not significant between test administration, $t(64) = 1.51$, $p > .10$. As in Study 2, a moderate effect of gender was observed (Table 4).

GENERAL DISCUSSION

The construct of empathy has assumed various definitions as reflected by the heterogeneous nature of current self-report measures of empathy. In an EFA, we determined what was shared by the corpus of empathy questionnaires by determining a single common factor. We then used items forming this factor to construct a new unidimensional scale, the TEQ, for the assessment of empathy. This new scale captures the underlying consensus among questionnaire measures currently in use and may prove an important tool in capturing performance on this elusive construct. The items represented in this single factor suggest that among current measures of empathy, the most commonly measured construct reflects primarily an emotional process or an accurate affective insight into the feeling state of another. The results of Studies 1 through 3 demonstrate that the TEQ possesses a robust single factor structure, high internal consistency, and convergent validity with existing self-report scales as well as behavioral measures of interpersonal skills and high test–retest reliability. Overall, the TEQ is a psychometrically sound, easily administered, and brief self-report measure of empathy.

Toronto Empathy Questionnaire

Emphasis on the emotional components of empathic responding in the TEQ is consistent with the approach taken by other researchers in forming self-report measures of empathy (e.g., Mehrabian & Epstein, 1972). For example, a confirmatory factor analysis of the IRI found one general dimension of empathy at the apex, Empathic Concern; this dimension overlaps to a great extent with Perspective Taking and Fantasy (Cliffordson, 2002). Consistent with this finding, the TEQ correlated highly with the IRI subscales of Empathic Concern (Studies 1 and 2) and to a lesser degree Perspective Taking (Studies 1 and 2) and Fantasy (Study 2). Taken together, these results suggest that the four-factor (i.e., multiple subscale) solution implicit in the IRI may not be necessary to capture empathic responding in self-report measures.

Cognitive accounts of empathy, although not mutually exclusive to affective accounts, emphasize aspects of social responding involving the ability to take the perspective of another (Allport, 1961; Mead, 1934), role-taking (Mead, 1934), and the ability to infer and predict another's behavior or mental state (Baron-Cohen & Wheelwright, 2004; Dennett, 1987). The TEQ demonstrated an association with this cognitive account and correlated with the IRI subscales of Perspective Taking and Fantasy, described previously as the cognitive components of empathy (Davis, 1983). This association suggests significant overlap across the cognitive and affective components of empathy described in the literature in which intercorrelation of emotional and cognitive accounts of empathic responding may indicate shared processes (for similar accounts of theory of mind reasoning, see Leslie, Friedman, & German, 2004). Indeed, evidence from neuroimaging and monkey research suggests that the cognitive and affective empathy may be mediated in different domains but are represented by the same underlying process in viscero-motor mirror neurons, neurons that fire in response to both executing and observing a goal-directed action or emotional experience of another (Gallese, 2003; Gallese, Keysers, & Rizzolatti, 2004).

The TEQ contains 16 questions that encompass a wide range of attributes associated with the theoretical facets of empathy. The affective aspect of empathic responding is thought to be related to such phenomena as emotional contagion (Eisenberg & Miller, 1987; Lipps, 1903), emotion comprehension (Haxby, Hoffman, & Gobbini, 2000), sympathetic physiological arousal (Levenson & Ruef, 1992) and con-specific altruism (Rice, 1964), all of which are represented in TEQ items. Two items specifically target the perception of an emotional state in another that stimulates the same emotion in oneself (Items 1 and 4). One item assesses emotion comprehension in others (Item 8). Other items address the assessment of emotional states in others by indexing the frequency of behaviors demonstrating appropriate sensitivity (Items 2, 7, 10, 12, and 15). The TEQ also contains items tapping sympathetic physiological arousal (Items 3, 6, 9, and 11) and altruism (Items 5, 14, and 16). Finally, one item probes the frequency of behaviors engaging higher order empathic responding such as prosocial helping behaviors (Item 13). Eight items are negatively scored (2, 4, 7, 10, 11, 12, 14, and 15), reflecting the frequency of situational indifference toward another individual on the previously described parameters. Taken together, these items represent a wide variety of empathy-related behaviors that have been described in current literature surrounding this process.

Associations With Other Measures

We predicted that the TEQ would diverge from measures surveying autism spectrum disorder because the latter taps deficits in social processing among other symptoms of this disorder. Consistent with this prediction, the TEQ showed a negative correlation with poor interpersonal and social responding as partially assessed by the Autism Quotient, a measure of autism spectrum disorder symptomatology (Baron-Cohen, Wheelwright, Skinner, et al., 2001), demonstrating concurrent validity. As expected, the magnitude of this association was not too great in light of the fact that the Autism Quotient also measures other symptoms of autism not related to social skill. The TEQ demonstrated convergent validity in the positive correlations observed, not only with self-report measures of empathy but with two behavioral measures that require the processing of complex interpersonal stimuli. Interpersonal information must be interpreted accurately to facilitate the task of responding in an empathic fashion (Bernieri, 2001). This is in contrast to previous findings in which empathy questionnaires and behavioral tasks often do not correlate (Ickes, 1997; cf. Zaki et al., 2008). Importantly, tasks such as the MIE and IPT that directly assess interpersonal sensitivity demonstrate a higher degree of ecological validity than do self-report tasks. Behavioral measures of interpersonal sensitivity, however, carry the disadvantage of being time and effort intensive. The TEQ provides a quick and easy way of assessing interpersonal sensitivity in a way consistent with these behavioral measures whereas providing substantial time savings and ease of administration. Notably, the IRI, a commonly used self-report measure of empathy, demonstrated weaker and statistically unreliable associations with these same tasks in our data set (see also Mar, Oatley, Hirsh, dela Paz, & Peterson, 2006).

The TEQ also correlated highly with a significantly lengthier measure of empathic responding, the 80-item Empathy Quotient (Baron-Cohen & Wheelwright, 2004). Shorter questionnaires such as the TEQ are especially useful for inclusion in mass-testing packets, Internet research, or in any other instance in which time and participant fatigue is an issue.

CONCLUSION

In developing the TEQ, we created a parsimonious scale that is short, clear, and homogenous and has strong psychometric properties including a robust single factor structure, high internal consistency, construct validity, and test–retest reliability. One limitation of this study is that our data were derived from a relatively small sample composed of college-aged students. Further work is required to assess the generalization of our findings to a wider age range. The observed central tendency and variability of the IRI, Autism Quotient, Empathy Quotient, MIE, and IPT–15 across our studies are, however, consistent with previously publications, suggesting that these samples are generalizable. Inconsistent gender differences, with effect sizes ranging from trivial to moderate, will need to be addressed in larger sample sizes. The TEQ, with its brevity and ease of administration, could be useful in patient populations.

Altered empathic responding has been reported in patients with Axis I (clinical syndromes; Deardorff, Kendall, Finch, & Sitarz, 1977; L. E. O'Connor, Berry, Weiss, & Gilbert, 2002) and Axis II (developmental and personality disorders; Guttman & Laporte, 2000; Tantam, 1995) psychiatric disorders, as well as in neurological patients with "acquired sociopathy" (Blair & Cipolotti, 2000), frontal lobe lesions (Eslinger, 1998), and frontotemporal lobar degeneration (Rankin et al., 2005). These deficits pose serious challenges to the quality of life of the patient, family members, and caregivers. Work is currently underway in our laboratory to develop a second caregiver-report measure based on the TEQ. Deficits in empathic understanding may be better understood through assessment and quantification, leading to effective intervention.

REFERENCES

Allport, G. W. (1961). *Pattern and growth in personality*. New York: Holt, Rinehart & Winston.

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,* 257–265.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory-of-mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34,* 163–175.

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,* 241–251.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31,* 5–17.

Barrett, P. T., & Kline, P. (1981). The observation to variable ratio in factor analysis. *Personality Study and Group Behavior, 1,* 23–33.

Bernieri, F. J. (2001). Toward a taxonomy of interpersonal sensitivity. In J. Hall & F. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 3–20). Mahwah, NJ: Lawrence Erlbaum Associates.

Blair, R. J., & Cipolotti, L. (2000). Impaired social response reversal. A case of "acquired sociopathy." *Brain, 123,* 1122–1141.

Burgess, P. W., Alderman, N., Evans, J. J., Wilson, B. A., & Emslie, H. (1996). The Dysexecutive Questionnaire. In B. A. Wilson, N. Alderman, P. W. Burgess, H. Emslie, & J. J. Evans (Eds.), *Behavioral assessment of the dysexecutive syndrome*. Bury St. Edmunds, England: Thames Valley Test Company.

Cliffordson, C. (2001). Parents' judgments and students' self-judgments of empathy: The structure of empathy and agreement of judgments based on the Interpersonal Reactivity Index (IRI). *European Journal of Psychological Assessment, 17,* 36–47.

Cliffordson, C. (2002). The hierarchical structure of empathy: Dimensional organization and relations to social functioning. *Scandinavian Journal of Psychology, 43,* 49–59.

Costanzo, M., & Archer, B. (1994). *The Interpersonal Perception Task 15 (IPT–15): A guide for researchers and teachers*. Berkeley: University of California Center for Media and Independent Learning.

Costanzo, M., & Archer, D. (1989). Interpreting the expressive behavior of others: The Interpersonal Perception Task. *Journal of Nonverbal Behavior, 13,* 225–245.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalogue of Selected Documents in Psychology, 10,* 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multi-dimensional approach. *Journal of Personality and Social Psychology, 44,* 113–126.

Deardorff, P. A., Kendall, P. C., Finch, A. J. Jr., & Sitarz, A. M. (1977). Empathy, locus of control and anxiety in college students. *Psychological Reports, 40,* 1236–1238.

Dennett, D. (1987). *The intentional stance*. Cambridge, MA: MIT Press/Bradford Books.

Eisenberg, N., & Miller, P. A. (1987). Empathy and prosocial behavior. *Psychological Bulletin, 101,* 91–119

Eslinger, P. J. (1998). Neurological and neuropsychological bases of empathy. *European Neurology, 39,* 193–199.

Froman, R. D., & Peloquin, S. M. (2001). Rethinking the use of the Hogan Empathy Scale: A critical psychometric analysis. *American Journal of Occupational Therapy, 55,* 566–572.

Gallese, V. (2003). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society of London: Series B: Biological Sciences, 358,* 517–528.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8,* 396–403.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103,* 265–275.

Guilford, J. (1954). *Pychometric methods*. New York: McGraw-Hill.

Guttman, H. A., & Laporte, L. (2000). Empathy in families of women with borderline personality disorder, anorexia nervosa, and a control group. *Family Process, 39,* 345–358.

Hashimoto, H., & Shiomi, K. (2002). The structure of empathy in Japanese adolescents: Construction and examination of an empathy scale. *Social Behavior and Personality, 30,* 593–602.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science, 4,* 223–233.

Hemphill, J. F. (2003). Interpreting the magnitude of correlation coefficients. *American Psychologist, 58,* 78–79.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33,* 307–316.

Hojat, M., Mangione, S., Gonnella, J. S., Nasca, T., Veloski, J. J., & Kane, G. (2001). Empathy in medical education and patient care. *Academic Medicine, 76,* 669.

Hornak, J., Rolls, E. T., & Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioral changes following ventral frontal lobe damage. *Neuropsychologia, 34,* 247–261.

Ickes, W. E. (1997). *Empathic accuracy*. New York: Guilford.

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: reliability and validity of the Empathy Quotient.*Psychological Medicine, 34,* 911–919.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind." *Trends in Cognitive Sciences, 8,* 528–533.

Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology, 63,* 234–246.

Lipps, T. (1903). Einfühlung, innere Nachahmung und Organempfindung [Empathy, inner emotion and perceptual experience]. *Archiv für die gesamte Psychologie, 1,* 465–519.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4,* 84–99.

Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality*, 40, 694–712.

Masserman, J. H., Wechkin, S., & Terris, W. (1964). "Altruistic" behavior in rhesus monkeys. *American Journal of Psychiatry, 121,* 584–585.

Mead, G. H. (1934). *Mind, self and society*. Chicago: Chicago University Press.

Mehrabian, A. (2000). Manual for the Balanced Emotional Empathy Scale (BEES). Unpublished manuscript. Available from Albert Mehrabian, 1130 Alta Mesa Road, Monterey, CA 93940.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40,* 525–543.

Mehrabian, A., Young, A. L., & Sato, S. (1988). Emotional empathy and associated individual differences. *Current Psychology: Research & Reviews, 7,* 221–240.

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

O'Connor, L. E., Berry, J. W., Weiss, J., & Gilbert, P. (2002). Guilt, fear, submission, and empathy in depression. *Journal of Affective Disorders, 71,* 19–27.

Omdahl, B. L. (1995). *Cognitive appraisal, emotion, and empathy*. Mahwah, NJ: Lawrence Erlbaum Associates.

Preston, S. D., & de Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral Brain Sciences, 25,* 1–71.

Rankin, K. P., Kramer, J. H., & Miller, B. L. (2005). Patterns of cognitive and emotional empathy in frontotemporal lobar degeneration. *Cognitive Behavioral Neurology, 18,* 28–36.

Reynolds, W. (2000). *The measurement and development of empathy in nursing*. Aldershot, England: Ashgate.

Rice, G. E. (1964). Aiding behavior vs. fear in the albino rat. *Psychological Record, 14,* 165–170.

Rice, G. E., & Gainer, P. (1962). "Altruism" in the albino rat. *Journal of Comparative and Physiological Psychology, 55,* 123–125.

Rankin, K. P., Kramer, J. H., & Miller, B. L. (2005). Patterns of cognitive and emotional empathy in frontotemporal lobar degeneration. *Cognitive Behavioral Neurology, 18,* 28–36.

Sagi, A., & Hoffman, M. L. (1976). Empathic distress in the newborn. *Developmental Psychology, 12,* 175–176.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D., Cooper, J. T., Golden, C., et al. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25,* 167–77.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology, 15,* 201–193.

Steger, M. F. (2006). An illustration of issues in factor extraction and identification of dimensionality in psychological assessment data. *Journal of Personality Assessment, 86,* 263–272.

Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Lawrence Erlbaum Associates.

Tantam, D. (1995). Empathy, persistent aggression, and antisocial personality disorder. *Journal of Forensic Psychiatry, 6,* 10–18.

Titchener, E. (1909). *Experimental psychology of the thought process*. New York: Macmillan.

Ungerer, J. A. (1990). The early development of empathy. *Motivation and Emotion, 14,* 93–106.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41,* 321–327.

Wang, Y., Davidson, M., Yakushko, O. F., Savoy, H. B., Tan, J. A., & Bleier, J. K. (2003). The scale of ethnocultural empathy: Development, validation, and reliability. *Journal of Counseling Psychology, 50,* 221–234.

Wispé, L. (1986). The distinction between empathy and sympathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology, 50,* 314–321.

Wispé, L. (1987). History of the concept of empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 17–37). Cambridge, England: Cambridge University Press.

Zahn-Waxler, C., Friedman, S. L., & Cummings, E. M. (1983). Children's emotions and behaviors in response to infants' cries. *Child Development, 54,* 1522–1528.

Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science, 19,* 399–404.

APPENDIX

Toronto Empathy Questionnaire Instructions

Below is a list of statements. Please read each statement *carefully* and rate how frequently you feel or act in the manner

described. Circle your answer on the response form. There are no right or wrong answers or trick questions. Please answer each question as honestly as you can.

1. When someone else is feeling excited, I tend to get excited too
2. Other people's misfortunes do not disturb me a great deal
3. It upsets me to see someone being treated disrespectfully
4. I remain unaffected when someone close to me is happy
5. I enjoy making other people feel better
6. I have tender, concerned feelings for people less fortunate than me
7. When a friend starts to talk about his/her problems, I try to steer the conversation towards something else
8. I can tell when others are sad even when they do not say anything
9. I find that I am "in tune" with other people's moods
10. I do not feel sympathy for people who cause their own serious illnesses
11. I become irritated when someone cries
12. I am not really interested in how other people feel
13. I get a strong urge to help when I see someone who is upset
14. When I see someone being treated unfairly, I do not feel very much pity for them
15. I find it silly for people to cry out of happiness
16. When I see someone being taken advantage of, I feel kind of protective towards him/her

Scoring Item responses are scored according to the following scale for positively worded Items 1, 3, 5, 6, 8, 9, 13, 16. Never = 0; Rarely = 1; Sometimes = 2; Often = 3; Always = 4. The following negatively worded items are reverse scored: 2, 4, 7, 10, 11, 12, 14, 15. Scores are summed to derive total for the Toronto Empathy Questionnaire.