



### A Brief Form of the Interpersonal Reactivity Index (B-IRI)

Journal:	<i>Journal of Personality Assessment</i>
Manuscript ID	JPA-2015-476.R1
Manuscript Type:	General Submission
Keywords:	Construct Validation < Content or Topic, Emotional Intelligence < Content or Topic, Test Development < Content or Topic, Confirmatory Factor Analysis/SEM < Methods and Statistics, Exploratory Factor Analysis < Methods and Statistics

SCHOLARONE™  
Manuscripts

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Running head: INTERPERSONAL REACTIVITY INDEX, BRIEF FORM, EMPATHY

**A Brief Form of the Interpersonal Reactivity Index (B-IRI)**

For Peer Review Only

### Abstract

The Interpersonal Reactivity Index (IRI) is a standardized self-report measure of disposition to empathic responsiveness for the general adult population (the domain for which it was developed), and for the general adolescent population. The IRI has a number of problems, however, including some uncertainty about its factor structure, low reliabilities, and poor readability of some items for people with limited literacy skills. To address these issues, we constructed an abbreviated form of the index, the Brief IRI (B-IRI). Three studies demonstrated that this 16-item B-IRI has a clear and coherent factor structure, adequate internal consistency, measurement invariance across gender and age, and theoretically meaningful associations with a range of external criteria that support its construct validity. The B-IRI substantially preserves the psychometric properties of the long form, and we recommend its use in all research settings.

*Keywords:* IRI, brief form development, measurement invariance, confirmatory factor analysis, validity construct

### A Brief Form of the Interpersonal Reactivity Index (B-IRI)

Empathy is a core dimension of social functioning, relating to affective and cognitive responses to others' emotional circumstances. The term describes the tendency to consider others' viewpoints and vicariously experience their emotional states (Davis, 1980, 1983, 1996; Hoffman, 2008). As Decety and Lamm (2006) put it, empathy is the ability to experience and understand what others feel without confusing oneself with others" (p. 1146). It is a key component of social interactions, providing a basis for prosocial behavior (Eisenberg, Fabes, & Spinrad, 2006), inhibiting aggressive behavior towards others (Jolliffe & Farrington, 2004), and bullying (Gini, Albiero, Benelli, & Altoè, 2007), maintaining social relations (Ingoglia, Lo Coco, Liga, & Lo Cricchio, 2011), and enhancing psychological well-being (Musick & Wilson, 2003). Considering its relevance in several areas of intervention to promote individuals' psychosocial adjustment (e.g., Feshbach & Feshbach, 2009), it is essential for psychologists and educators to have valid and reliable procedures for measuring empathic responsiveness.

Various self-report measures of empathy have been developed over the years. The *Interpersonal Reactivity Index* (IRI; Davis, 1980, 1983) is a scale widely used to measure individual differences in disposition to empathic responsiveness (for a review, see Pedersen, 2009; Yu & Kirk, 2009). The IRI's popularity is due to several desirable qualities; it is based on a multidimensional conceptualization of empathy, and regarded as one of the most comprehensive measures of self-reported empathic disposition. The 28-item IRI consists of four 7-item subscales measuring four different dimensions of interpersonal reactivity (Davis, 1980, 1983). In Davis's model, empathy is seen as a set of distinct, but related constructs, two of which are cognitive dimensions representing two different types of antecedent of the experience of emotions in response to emotions felt by others: *Fantasy* refers to the tendency to imagine

1  
2  
3 feeling and acting like fictitious characters portrayed in movies, books, etc.; and *Perspective*  
4  
5 *Taking* refers to the spontaneous tendency to adopt another person's psychological viewpoint.  
6  
7  
8 The other two IRI dimensions are purely and typically emotional and represent two different  
9  
10 ways of vicariously participating in other people's emotions: *Empathic Concern* refers to "other-  
11  
12 oriented" feelings of sympathy and concern for unfortunate others; *Personal Distress* refers to  
13  
14 "self-oriented" feelings of personal anxiety and unease in tense interpersonal settings.  
15  
16

17  
18 The IRI is a good self-report measure of disposition to empathic responsiveness for the  
19  
20 general adult population (the domain for which it was developed), and for the general adolescent  
21  
22 population. It has been adapted for many different countries and cultures, and numerous studies  
23  
24 have supported the adequacy of its four-factor structure among US, Swiss, Dutch-speaking  
25  
26 Belgian, Dutch, Swedish, and Chinese adults (Cliffordson, 2002; Davis, 1980; De Corte, Buysse,  
27  
28 Verhofstadt, Roeyers, Ponnet, & Davis, 2007; Gilet, Mella, Studer, Grünh, & Labouuvie-Vief,  
29  
30 2012; Hawk, Keijsers, Branjie, van der Graff, de Wied, & Meeus, 2013; Huang, Li, Sun, Chen,  
31  
32 & Davis, 2012), and among Dutch, Italian, Spanish, and Swedish preadolescents and adolescents  
33  
34 (Albiero, Ingoglia, & Lo Coco, 2006; Cliffordson, 2001; Hawk et al., 2013; Tello, Egido, Ortiz,  
35  
36 & Gandara, 2013). Using simplified items also produced similar results in US 7- to 10-year-olds  
37  
38 (Litvack-Miller, McDougall, & Romney, 1997). A recent study also examined IRI measurement  
39  
40 invariance across age and gender (Hawk et al., 2013), showing the invariance of the factor  
41  
42 loadings and factor variances and covariances. Notwithstanding, Hawk and colleague did not  
43  
44 evaluate the invariance at the level of item intercepts and residual variances.  
45  
46  
47  
48  
49

50  
51 The scale has nonetheless exhibited a number of problems too, which underscore the need  
52  
53 to further investigate certain validity issues. For example, although several studies noted earlier  
54  
55 have replicated the four factor model, as noted by De Corte et al. (2007) various other studies  
56  
57  
58  
59  
60

1  
2  
3 have failed to replicate it. Several studies have identified different factor solutions (Alterman,  
4  
5 McDermott, Cacciola, & Rutherford, 2003; Cliffordson, 2002; Poulos, Elison, & Lennon, 2004;  
6  
7  
8 Siu & Shek, 2005; Tello et al., 2013): for instance, some authors found a higher-order model  
9  
10 with two global factors (Cliffordson, 2002), while others judged that a one-dimensional model  
11  
12 represented the IRI data best (Alterman et al., 2003). Moreover, although they globally  
13  
14 confirmed the four-factor structure hypothesized by Davis, some researchers also reported the  
15  
16 need to delete some items, or to add some residual covariances into the factor structure analyses,  
17  
18 in order to reach a good fit to the data. In a study on Italian preadolescents and adolescents, for  
19  
20 example, Albiero et al. (2006) removed three items from the model and added four residual  
21  
22 covariances (intended as the covariance between the measurement errors associated with two  
23  
24 items) involving negatively-worded items. In a study on Dutch-speaking Belgian adults, De  
25  
26 Corte et al. (2007) added four residual covariances involving some of the items on the Fantasy  
27  
28 subscale; and the authors suggested that this could be due to some semantic overlap giving rise to  
29  
30 a covariance between the Fantasy items. Finally, a number of problems emerge when the IRI is  
31  
32 used with special populations (such as offenders), including a different factor structure, low  
33  
34 reliabilities and a poor readability of some items (Beven, O'Brien-Malone, & Hall, 2004;  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 difficulties with the negatively-worded items, so they proposed using a shortened version of the  
4  
5 IRI, omitting the negatively-worded items. Using this shorter version, the component structure of  
6  
7 the IRI was replicated, the reliabilities were adequate, and differential validity was demonstrated.  
8  
9

10  
11 Given the above-mentioned concerns, we believe that the field of study on empathy  
12  
13 would benefit from an appropriately-developed brief version of the IRI that contains no  
14  
15 negatively-worded or poorly-understood items, and that reduces the semantic overlap identified  
16  
17 by De Corte et al. (2007). Guidelines for constructing and refining scales (Floyd & Widaman,  
18  
19 1995; Smith, McCarthy, & Anderson, 2000), and for developing short forms (Smith et al., 2000),  
20  
21 indicate a number of rational, quantitative steps that can be taken to shorten and improve the IRI.  
22  
23 These guidelines suggest that short form developers conduct content analyses of each factor in  
24  
25 the scale, then administer an abbreviated form to a different sample from the one considered  
26  
27 when the scale was developed. They will then need to: show that each factor meets reasonable  
28  
29 reliability standards; demonstrate the factor structure of the short form, using both exploratory  
30  
31 factor analyses (EFA) and confirmatory factor analyses (CFA); and validate their short form. The  
32  
33 shortened version of the IRI proposed by Lauterbach and Hosser (2007) fails to comply with  
34  
35 these guidelines: the authors did not clearly explain why they omitted some items; they  
36  
37 investigated their scale's dimensionality by performing a series of principal component analyses  
38  
39 instead of using EFA and CFA; and they did not replicate their results in an independent sample.  
40  
41 By means of a series of studies, we followed the guidelines step by step to construct a brief form  
42  
43 of the IRI (B-IRI) that would contain no poor items, and that would have a simple factor  
44  
45 structure, improving the separability of the individual scales while maintaining their internal  
46  
47 consistency and core content coverage. We tested the B-IRI factor structure in different samples  
48  
49 and assessed its reliability. We also investigated its measurement invariance across gender and  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

age (comparing adolescents and emerging adults). We provided evidence of its construct validity by examining whether scores obtained on the B-IRI subscales showed any gender- or age-related differences, and patterns of associations consistent with theoretical accounts of empathy and prior empirical studies on the topic.

### Study 1

The purpose was to construct a brief form of the IRI, shortening the scale while retaining its psychometric properties and aiming for subscales containing four items each. A scale of this length would be more practical to administer and consequently more appealing to scholars working in various areas who might benefit from using this instrument, but it would have to strike a good balance between the goal of brevity and theoretical and psychometric concerns (assuring adequate content coverage and reliability, for instance).

### Method

#### *Participants*

The first sample was composed of 1,104 Italian adolescents and emerging adults (38% male), ranging in age from 14 to 30 years ( $M = 17.64$  years,  $SD = 3.03$ ), and coming from several northern and southern regions of Italy. All participants were Caucasian. The adolescents were 818 high-school students (41% male) recruited from randomly-selected schools in several urban areas. They ranged in age from 14 to 19 years ( $M = 16.22$  years,  $SD = 1.64$ ). They were living in one household with their parent(s); 94% of them came from intact, two-parent families, 5% had parents who had separated or divorced, and one of the parents had died in 1% of cases. The adolescents' parents were socioeconomically diverse (14% professional or managerial, 36% tradespeople, 15% skilled workers, 33% unskilled workers, 2% retired); and 55% of the adolescents' parents had completed their secondary schooling. The emerging adults were 286



BRIEF IRI

8

1  
2  
3 university students (31% male) attending several universities. They ranged in age from 20 to 30  
4  
5 years ( $M = 21.64$  years,  $SD = 2.23$ ). They were all unmarried; 35% were residential students  
6  
7 (living on campus), 17% lived off campus without their family, 48% lived off campus with their  
8  
9 family; 90% came from two-parent families, 9% had parents who had separated or divorced, and  
10  
11 one of the parents had died in 1% of cases. The emerging adults' parents were socioeconomically  
12  
13 diverse (22% professional or managerial, 21% tradespeople, 36% skilled workers, 10% unskilled  
14  
15 workers, 11% retired); and 68% had parents who had completed their secondary schooling.  
16  
17  
18

### 19 *Measure*

20  
21 Participants were administered the 28-item IRI (Davis, 1980, 1983). They were asked to  
22  
23 indicate the extent to which each item described them on a 5-point Likert-type scale ranging  
24  
25 from 1 (*Does not describe me at all*) to 5 (*Describes me very well*). In the present study, the  
26  
27 subscales had an adequate internal consistency: Cronbach's alpha was .78, .73, .71, and .76 for  
28  
29 Fantasy, Empathic Concern, Perspective Taking and Personal Distress, respectively.  
30  
31  
32

### 33 *Procedure*

34  
35 The Institutional Review Boards (IRB) of the University of Palermo and the University  
36  
37 of Padua approved this study, which was conducted in conformity with the guidelines for the  
38  
39 ethical treatment of human participants of the Italian Association of Psychology. Participation in  
40  
41 the study was voluntary and anonymous, and participants received no compensation; 98% of the  
42  
43 students initially contacted agreed to take part in the research. Prior permission was obtained  
44  
45 from each participant, from the university dean or school director, and from the course professors  
46  
47 involved. Participants received written information about the study. Parents' written informed  
48  
49 consent was obtained for minors under the age of 18. The IRI was administered collectively  
50  
51 during class sessions under the supervision of four psychology undergraduates; it took no longer  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 than 15 minutes to complete. Research assistants provided verbal instructions in addition to the  
4  
5 written instructions accompanying the questionnaire. The same procedure was employed in  
6  
7 subsequent studies.  
8  
9

### 10 *Procedure for developing the brief form*

11  
12 Our general aim was to shorten the IRI scale with no loss of its psychometric properties  
13 (in terms of reliability and validity), aiming for subscales containing four items each that could  
14 still ensure an adequate breadth of content domain coverage. To achieve this result, we  
15 considered the content of each of the previously-described factors, and tried to retain the items  
16 most closely describing: the tendency to imagine feeling and acting like fictitious characters  
17 portrayed in novels, movies, etc., for the *Fantasy* subscale; the tendency to adopt another  
18 person's psychological point of view for *Perspective Taking*; "other-oriented" feelings of  
19 sympathy and concern for unfortunate others for *Empathic Concern*; and "self-oriented" feelings  
20 of personal anxiety in tense interpersonal contexts for *Personal Distress*. Three independent  
21 judges expert in dealing with empathic issues examined the items' content for this purpose.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35

36 **Rational identification of problematic items.** In a first step, we assessed the quality of  
37 the IRI items in terms of the clarity of their content. The judges found all IRI items clearly  
38 defined except for item 20 ("I am often quite touched by things that I see happen"), which  
39 seemed more vague than the other items on the Empathic Concern subscale. Then the negatively-  
40 worded items were assessed. The inclusion of such items is controversial: it is usually thought to  
41 improve respondents' level of attention and avoid response styles such as acquiescence bias  
42 (Anastasi & Urbina, 1997), but psychometric analyses have found that they can pose problems. It  
43 has been shown, for example, that respondents may calibrate negatively-worded items differently  
44 (Bode, Smith, & Smith, 2004; Muris, Meesters, & Gobel, 2001), and such sentences take longer  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 to process (Clark & Chase, 1972), meaning that they probably demand greater literacy skills.  
4  
5 Based on the previously-mentioned studies on offenders (Beven et al., 2004; Lauterbach &  
6  
7  
8 Hossler, 2007), we generally aimed to omit such items. Notwithstanding, with regard to the  
9  
10 Empathic Concern subscale, we decided to retain one negatively-worded item since this subscale  
11  
12 contains four positively-worded items but the content of item 20 was considered vague by the  
13  
14 judges. To decide which of these negatively-worded items to retain (item 4, “Sometimes I don’t  
15  
16 feel very sorry for other people when they are having problems”, item 14, “Other people’s  
17  
18 misfortunes do not usually disturb me a great deal”, and item 18, “When I see someone being  
19  
20 treated unfairly, I sometimes don’t feel very much pity for them”), the judges examined their  
21  
22 content and identified item 18 as the one most adequately describing the content of this subscale,  
23  
24 highlighting the feeling of sympathy for others. Globally, in this first step, we removed eight  
25  
26 negatively-worded items from the IRI: 7 (“I am usually objective when I watch a movie or play,  
27  
28 and I don’t often get completely caught up in it”) and 12 (“Becoming extremely involved in a  
29  
30 good book or movie is somewhat rare for me”) on the *Fantasy* subscale; 4 and 14 on the  
31  
32 *Empathic Concern* subscale; 3 (“I sometimes find it difficult to see things from the “other guy’s”  
33  
34 point of view”) and 15 (“If I’m sure I’m right about something, I don’t waste much time listening  
35  
36 to other people’s arguments”) on the *Perspective Taking* subscale; and 13 (“When I see someone  
37  
38 get hurt, I tend to remain calm”) and 19 (“I am usually pretty effective in dealing with  
39  
40 emergencies”) on the *Personal Distress* subscale. We also removed item 20 on the *Empathic*  
41  
42 *Concern* subscale since its content was considered vague by the judges.  
43  
44  
45  
46  
47  
48  
49

50  
51 **Empirical identification of the most appropriate items based on internal consistency**  
52  
53 **analyses.** The items to retain were empirically identified using metrics of scale consistency, such  
54  
55 as item-total correlations to identify items that demonstrated clear quantitative relationships with  
56  
57  
58  
59  
60

one another, supporting the cohesiveness of the constructs and domains established during the initial definition of the domains (Floyd & Widaman, 1995; Smith et al., 2000). After removing the negatively-worded and poor items, we then omitted any items with a corrected item-total correlation  $< .30$ , since when a corrected item-total correlation coefficient value is  $< .30$ , it indicates that the item contributes very little to the homogeneity of the scale (Streiner, Norman, & Cairney (2014). This applied to three items: 1 (“I daydream and fantasize, with some regularity, about things that might happen to me”, *Fantasy*), 21 (“I believe that there are two sides to every question and try to look at them both”, *Perspective Taking*), and 10 (“I sometimes feel helpless when I am in the middle of a very emotional situation”, *Personal Distress*).

## Results

### *Descriptive statistics for the B-IRI items*

The means, standard deviations, skewness, and kurtosis of the B-IRI, and Pearson’s correlation coefficients are given in Table 1. The data had a normal univariate distribution, the skewness and kurtosis values being approximately in the range  $-1.0$  and  $+1.0$  (Muthén & Kaplan, 1985).

(insert Table 1 about here)

### *Structural validity of the B-IRI*

We conducted an EFA to examine the factor structure of the B-IRI. We performed the principal axis factoring, and the extracted factors were rotated obliquely. A 4-factor solution was chosen, based on theoretical considerations and the empirical criterion of the random data parallel analyses (Horn, 1965). The eigenvalues derived from the actual data were compared to the eigenvalues derived from the random data. Factors were retained as long as the  $i$ th eigenvalue from the actual data was greater than the  $i$ th eigenvalue from the random data (O’Connor, 2000).

Both Kaiser's (1961) criterion (items with eigenvalues  $\geq 1$ ) and the scree test (Cattell, 1966) were secondarily checked for agreement. Table 2 shows the factor pattern matrix. The 4-factor B-IRI model explained 40.66% of the total variance in the data. There was no cross-loading among the items (cross loadings ranged from .01 to .18), and all factor loadings were higher than .40. All items saturated in the expected factors. Correlations between factors ranged from .07 to .41<sup>1</sup>.

(insert Table 2 about here)

### ***Internal consistency of the B-IRI***

Corrected item-total scale correlation coefficients and Cronbach's alpha were computed to assess the internal consistency of the B-IRI subscales. The results showed an adequate internal consistency. The item-total scale correlation ranged from .55 to .64 for *Fantasy*, from .40 to .54 for *Empathic Concern*, from .34 to .58 for *Perspective Taking*, and from .48 to .54 for *Personal Distress*. Cronbach's alpha was .79, .69, .65, and .71 for *Fantasy*, *Empathic Concern*, *Perspective Taking*, and *Personal Distress*, respectively.

### **Study 2**

We examined the psychometric properties of the B-IRI in more depth in a new, independent sample. First, we tested the hypothesized 4-factor structure of the scale. Then we tested the B-IRI's measurement invariance across gender and age by comparing males with females, and adolescents with emerging adults. Extensive research has addressed the issues of gender differences in empathy (e.g., Eisenberg & Lennon, 1983; De Corte et al., 2007), and empathy development from adolescence to adulthood (e. g., Davis, 1996; Hoffman, 2008), but few studies have examined whether scales such as the IRI have a similar psychometric structure

---

<sup>1</sup> Similar results were found when EFA was run for males and females separately, and for adolescents and emerging adults separately.

1  
2  
3 across genders and age groups. Hawk et al. (2013) recently found the IRI a psychometrically  
4  
5 invariant measure for early and late adolescents, boys and girls, adolescents and their mothers.  
6  
7 Tello et al. (2013) likewise tested the IRI's measurement invariance across gender in a sample of  
8  
9 adolescents, finding the scale psychometrically invariant for boys and girls. The latter authors  
10  
11 did not confirm the 4-factor model, however, but a hierarchical model characterized by five first-  
12  
13 order factors and one second-order factor. More research is therefore needed on the measurement  
14  
15 invariance of this scale. We also assessed the B-IRI's reliability in terms of the internal  
16  
17 consistency of its subscales, and we examined whether the scores on the B-IRI subscales showed  
18  
19 any gender- or age-related differences.  
20  
21  
22  
23

## 24 Method

### 25 *Participants*

26  
27  
28  
29 The second sample consisted of 836 Italian adolescents and emerging adults (48% male),  
30  
31 ranging in age from 15 to 27 years ( $M = 20.54$ ,  $SD = 3.26$ ), and coming from several northern  
32  
33 and southern regions of Italy. All participants were Caucasian. The adolescents were 354 high-  
34  
35 school students (49% male) recruited from randomly-selected schools located in several urban  
36  
37 areas. They ranged in age from 15 to 19 years ( $M = 17.31$  years,  $SD = 1.19$ ); they were living in  
38  
39 one household with their parent(s); 93% of them came from intact, two-parent families, 6% had  
40  
41 divorced or separated parents, and 1% came from a family in which one of the parents had died.  
42  
43 The adolescents' parents were socioeconomically diverse (16% professional or managerial, 34%  
44  
45 tradespeople, 17% skilled workers, 31% unskilled workers, 2% retired); and 57% of the  
46  
47 adolescents' parents had completed their secondary schooling. The emerging adults were 482  
48  
49 university students (47% male) attending two universities. They ranged in age from 20 to 27  
50  
51 years ( $M = 22.91$  years,  $SD = 2.02$ ). They were all unmarried; 36% were residential students  
52  
53  
54  
55  
56  
57  
58  
59  
60

(living on campus), 17% lived off campus without their family, 47% lived off campus with their family; 90% came from two-parent families, 9% had parents who had separated or divorced, and one of the parents had died in 1% of cases. Emerging adults' parents were socioeconomically diverse (28% professional or managerial, 24% tradespeople, 26% skilled workers, 9% unskilled workers, 13% retired); and 69% had parents who had completed their secondary schooling.

### ***Measure***

Participants were administered the 16-item B-IRI (see Appendix). It consists of only positively-worded items, including the originally negatively-worded IRI item 18 of the *Empathic Concern* subscale. In order to make it positively worded like the remaining B-IRI items, it was changed as follows: "When I see someone being treated unfairly, I feel very much pity for them" (see Study 1 for a more complete description of the item selection procedure). Each subscale consists of 4 items. Participants were asked to indicate the extent to which each item described them on a 5-point Likert-type scale ranging from 1 (*Does not describe me at all*) to 5 (*Describes me very well*).

### ***Procedure***

The same procedure was used as for Study 1. The B-IRI was administered collectively during class sessions.

## **Results**

### ***Confirmatory factor analysis of the B-IRI***

A CFA was performed to test the B-IRI's factor structure. The CFA was based on examining the covariance matrix using the EQS 6.1 structural equation modeling software (Bentler, 2006). The 4-factor model was evaluated, allowing each item to load on the hypothesized factor and setting all other factor loadings at zero. Factor covariances were free



parameters to be estimated. To establish the measurement scale of each factor, their variance was fixed at 1.0. Since items exhibited a multivariate non-normal distribution (the normalized Mardia's coefficient was 39.81,  $p < .001$ ), the robust maximum likelihood estimation method was used, which adjusts standard errors of parameter estimates and chi-square statistics (SB  $\chi^2$ ) to account for non-normality (Satorra & Bentler, 1994). Goodness of fit indexes showed that the model generally fitted the data well:  $\chi^2(98) = 344.89$ , SB  $\chi^2(98) = 271.76$ ,  $p < .001$ , robust CFI = .932, RMSEA = .046, although the CFI was lower than the recommended value of .95 (Kenny & McCoach, 2003). The standardized solution is shown in Figure 1. All parameters were significant, except for the covariance between *Perspective Taking* and *Personal Distress*.

(insert Figure 1 about here)

### ***B-IRI measurement invariance across gender***

After judging the fit of the 4-factor model in the sample as a whole, we examined the gender invariance for the model by running a multisample CFA. We tested five levels of group invariance, i.e. configural, metric, factor variance, factor covariance, and residual variance invariance (Cheung & Rensvold, 2002). *Configural invariance* was supported by a good fit for an unconstrained baseline model (M0) in which factor loadings, factor covariances, and residual variances differed between the groups. *Metric invariance* was tested by comparing M0 with a model in which all loadings were simultaneously constrained across groups (M1). We then tested invariance in the **factor variances and covariances (M2), intercepts (M3), and residual variances (M4)**, respectively, comparing M0 with models in which their respective parameters were constrained. For all tests, a significant deterioration in the model's fit would indicate non-invariance and would also lead to a significant Satorra-Bentler adjusted chi-square difference (Satorra & Bentler, 2001). Because these chi-square difference tests detect small discrepancies



with no practical or theoretical implications in sample sizes  $> 200$ , we considered CFI decreases  $\geq .010$ , supplemented by RMSEA increases  $\geq .015$ , as more meaningful indicators (Chen, 2007; Cheung & Rensvold, 2002). The goodness of fit indexes are given in Table 3. Across the difference tests, the baseline model fit was acceptable. Some chi-square-tests were significant in the additional tests of invariance, but no worsening of CFI or RMSEA exceeded thresholds of  $\geq .010$  or  $\geq .015$ , respectively, suggesting no meaningful group differences. The only exception was represented by the comparison M4 vs. M0 which showed a significant worsening of CFI. The model was therefore modified releasing the equality constraint imposed on the residual variance of B-IRI item 14. Globally, our results revealed a full metric invariance, a factor variance/covariance invariance, an intercept invariance and a partial residual variance invariance, specifically the residual variance of B-IRI item 14 was higher for males (.69) than for females (.37).

(insert Table 3 about here)

### ***B-IRI measurement invariance across age***

A multisample CFA was also run to examine measurement invariance across age groups, comparing adolescents ( $n = 354$ , high school students ranging in age from 15 to 19 years) and emerging adults ( $n = 482$ , university students ranging in age from 20 to 27 years). The goodness of fit indexes are given in Table 3. Across the difference tests, the baseline model fit was acceptable. Some chi-square-tests were significant in the additional tests of invariance, but no worsening of CFI or RMSEA exceeded thresholds of  $\geq .010$  or  $\geq .015$ , respectively, suggesting no meaningful group differences. Globally, our results revealed full metric invariance, factor variance/covariance invariance, intercept invariance and residual variance invariance.

### ***Internal consistency of the B-IRI***

Corrected item-total scale correlation coefficients and Cronbach's alpha were computed to test the internal consistency of the B-IRI subscales. The results showed adequate internal consistency values. The item-total scale correlation ranged from .58 to .70 for *Fantasy*, from .40 to .54 for *Empathic Concern*, from .40 to .51 for *Perspective Taking*, and from .43 to .55 for *Personal Distress*. Cronbach's alpha was .82, .68, .69, and .71 for *Fantasy*, *Empathic Concern*, *Perspective Taking*, and *Personal Distress*, respectively.

### ***Gender- and age-related differences in the B-IRI subscales***

A 2 (gender) x 2 (age) multivariate analysis of variance (MANOVA) was performed to examine gender- and age-related differences in the mean levels of the B-IRI subscales. The results only revealed a significant multivariate effect for gender [Wilks's lambda = .85,  $F(4, 829) = 36.78$ ,  $p < .001$ ,  $\eta^2 = .15$ ]. Univariate ANOVAs revealed significant effects for all subscales (see Table 4), with females reporting higher scores in all B-IRI subscales than males.

(insert Table 4 about here)

### **Study 3**

When studying the psychometric properties of a measure, it is important to consider construct validity. We thus examined whether scores on the B-IRI subscales showed patterns of associations consistent with theoretical accounts of empathy and prior empirical studies. For this purpose, we examined each subscale in terms of its interrelations with measures of psychological functioning, including self-other differentiation, emotional fragility, and masculinity and femininity, which can provide further evidence of validity. Extensive prior research had shown that the IRI had specific links with a range of psychosocial processes.

*Fantasy* represents the tendency to become deeply involved in the fictitious world of books and movies. According to Davis (1983), *Fantasy* indicates a significant tendency to report

1  
2  
3 a lack of emotional invulnerability and, by implication, a certain degree of emotional  
4  
5 vulnerability; high fantasizers are more susceptible to emotional responses. Davis found this  
6  
7 subscale correlated with measures of sensitivity to others and measures of chronic emotionality  
8  
9 and masculinity. We thus expected *Fantasy* to be linked positively with emotional fragility and  
10  
11 femininity, and negatively with masculinity and self-other differentiation.  
12  
13

14  
15 *Perspective Taking* concerns the tendency to spontaneously adopt another person's  
16  
17 psychological viewpoint. Davis (1983) found *Perspective Taking* negatively related to  
18  
19 fearfulness, positively related to other-oriented sensitivity measures, and unrelated to masculinity  
20  
21 (which primarily reflects an emotional invulnerability and lack of responsiveness to emotional  
22  
23 situations). We thus expected *Perspective Taking* to be linked positively with femininity,  
24  
25 negatively with self-other differentiation, and unrelated to masculinity.  
26  
27

28  
29 *Empathic Concern* measures sympathy for others in need. Davis (1983) found similar  
30  
31 results for *Fantasy* and *Empathic Concern* with regard to emotionality; subjects scoring high on  
32  
33 the *Empathic Concern* subscale displayed mild tendencies towards emotional vulnerability  
34  
35 (femininity) and chronic fearfulness and insecurity. The scale was strongly associated with a  
36  
37 measure of selflessness and concern for others. We thus expected *Empathic Concern* to be linked  
38  
39 positively with emotional fragility and femininity, and negatively with masculinity and self-other  
40  
41 differentiation.  
42  
43

44  
45 *Personal Distress* represents an inability to regulate negative empathic arousal (Davis,  
46  
47 1983; Decety & Jackson, 2004; Eisenberg & Eggum, 2009). Davis (1983) found high *Personal*  
48  
49 *Distress* scores in individuals exhibiting a marked tendency towards emotional vulnerability  
50  
51 (masculinity) and a moderate tendency towards chronic fearfulness. Although the general pattern  
52  
53 of these relationships was similar to those seen for *Fantasy* and *Empathic Concern*, their  
54  
55  
56  
57  
58  
59  
60

magnitude was considerably greater with the *Personal Distress* subscale towards a type of emotionality characterized by fearfulness, uncertainty and vulnerability (Davis, 1983).

Consistent and significant positive correlations were also found between *Personal Distress* and self-oriented measures of sensitivity to others; this indicates a consistent tendency for high *Personal Distress* scores to be associated with a concern for the self as the object of others' evaluations. We consequently expected *Personal Distress* to be linked positively with emotional fragility and femininity, and negatively with masculinity and self-other differentiation.

We also replicated some of the analyses already performed in Study 2. Specifically, we examined the factorial structure of the B-IRI, the internal consistency of the B-IRI subscales, and their gender- and age-related differences, in a new, independent sample.

## Method

### *Participants*

The third sample included 649 Italian adolescents and emerging adults (40% male), ranging in age from 18 to 30 years ( $M = 22.34$ ,  $SD = 2.35$ ), and coming from several northern and southern regions of Italy. All participants were Caucasian. The adolescents were 273 high-school students (41% male) recruited from randomly-selected schools located in several urban areas. They ranged in age from 18 to 19 years ( $M = 18.22$  years,  $SD = 0.64$ ). They were living in one household with their parent(s); 95% of them came from intact, two-parent families, 4% had parents who had separated or divorced, and one of the parents had died in 1% of cases. The adolescents' parents were socioeconomically diverse (16% professional or managerial, 34% tradespeople, 17% skilled workers, 31% unskilled workers, 2% retired); and 61% of the adolescents' parents had completed their secondary schooling. The emerging adults were 377 university students (39% male) attending several universities. They ranged in age from 20 to 30

1  
2  
3 years ( $M = 24.28$  years,  $SD = 2.23$ ). They were all unmarried; 65% of them were residential  
4  
5 students (living on campus), 23% lived off campus without their family, 12% lived off campus  
6  
7 with their family; 92% came from two-parent families, 7% had parents who had separated or  
8  
9 divorced, and one of the parents had died in 1% of cases. The emerging adults' parents were  
10  
11 socioeconomically diverse (28% professional or managerial, 28% tradespeople, 25% skilled  
12  
13 workers, 11% unskilled workers, 8% retired); and 68% had parents who had completed their  
14  
15 secondary schooling.  
16  
17  
18

### 19 20 *Measures*

21  
22 **Empathy.** Participants were administered the 16-item B-IRI and asked to indicate the  
23  
24 extent to which each item described them on a 5-point Likert-type scale ranging from 1 (*Does*  
25  
26 *not describe me at all*) to 5 (*Describes me very well*).  
27  
28

29  
30 **Self-other differentiation.** Participants were administered the Self-Other Differentiation  
31  
32 Scale (SODS; Olver, Aries, & Batgos, 1989) to assess the degree to which they experienced a  
33  
34 separate sense of self in their relationships with others. The SODS scale comprises 11 items (e.g.  
35  
36 “If someone close to me finds fault with what I do, I find my self-evaluation lowered”), assessing  
37  
38 such relational aspects as deferring to the wishes of others, taking on the interests and  
39  
40 orientations of others, and over-reliance on others for criteria of self-worth. Each item is rated as  
41  
42 *true* (0) or *false* (1); the sum of the individual item ratings produces a total score, where higher  
43  
44 scores indicate a greater self-other differentiation. In the present study, the scale had an adequate  
45  
46 internal consistency: Cronbach's alpha was .70.  
47  
48  
49

50  
51 **Emotional fragility.** The Scale for the Measurement of Emotional Fragility (SMEF;  
52  
53 Caprara, Perugini, Barbaranelli, & Pastorelli, 1991) was administered to assess participants'  
54  
55 emotional fragility. It consists of 30 items (e.g., “It's difficult for me to choke back tears when  
56  
57  
58  
59  
60

1  
2  
3 I'm excited") that measure an individual's tendency to experience a sense of inadequacy and  
4  
5 distress in circumstances perceived as dangerous. The items were presented as declarative  
6  
7 statements; participants were asked to rate the degree to which each statement was true of  
8  
9 themselves on a 5-point scale (1 = *Very untrue*; 5 = *Very true*). In the present study, the subscales  
10  
11 had an adequate internal consistency: Cronbach's alpha was .80.  
12  
13

14  
15 **Masculinity and femininity.** To measure masculine and feminine personality styles  
16  
17 among men and women, participants were administered the Bem Sex Role Inventory (BSRI;  
18  
19 Bem, 1974), comprising two subscales: (a) masculinity (20 items), which includes characteristics  
20  
21 perceived as typical of men (e.g., assertive, strong personality, and dominant); and (b) femininity  
22  
23 (20 items) concerning characteristics perceived to be typical of women (e.g., emotional,  
24  
25 sympathetic, and understanding). Participants assessed how well each personality trait described  
26  
27 themselves on a 5-point scale (1 = *Almost never true*, 5 = *Almost always true*). In the present  
28  
29 study, the subscales had an adequate internal consistency: Cronbach's alpha was .82 for  
30  
31 masculinity and .84 for femininity.  
32  
33  
34  
35

### 36 *Procedure*

37  
38 The same procedure was used as for Study 1. The B-IRI was administered collectively  
39  
40 during class sessions.  
41  
42

## 43 **Results**

### 44 *Confirmatory factor analysis of the B-IRI*

45  
46 A CFA was performed to test the B-IRI's factor structure on a new sample. The goodness  
47  
48 of fit indexes showed that the model generally fitted the data well:  $\chi^2(98) = 281.60$ ,  $SB\chi^2(98) =$   
49  
50  $239.50$ ,  $p < .001$ , robust CFI = .925, RMSEA = .046, although the CFI was lower than the  
51  
52 recommended value of .95 (Kenny & McCoach, 2003). All factor loadings were significant; they  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 ranged from .35 (B-IRI item 9) to .77 (B-IRI item 14). All factor covariances were significant,  
4  
5 with the sole exception of the covariance between *Perspective Taking* and *Personal Distress*;  
6  
7 significant factor correlations ranged from .19 (*Empathic Concern-Personal Distress*) to .46  
8  
9 (*Empathic Concern-Perspective Taking*).  
10  
11

### 12 ***Internal consistency of the B-IRI***

13  
14  
15 Corrected item-total scale correlation coefficients and Cronbach's alpha were computed  
16  
17 to assess the internal consistency of the B-IRI subscales. The results showed adequate internal  
18  
19 consistency values. The item-total scale correlation ranged from .51 to .65 for *Fantasy*, from .38  
20  
21 to .46 for *Empathic Concern*, from .37 to .51 for *Perspective Taking*, and from .46 to .55 for  
22  
23 *Personal Distress*. Cronbach's alpha was .79, .68, .68, and .72 for *Fantasy*, *Empathic Concern*,  
24  
25 *Perspective Taking*, and *Personal Distress*, respectively.  
26  
27  
28

### 29 ***Gender- and age-related differences in the B-IRI subscales***

30  
31  
32 We assessed gender- and age-related differences in the B-IRI subscales by performing a 2  
33  
34 x 2 MANOVA. The results only revealed a significant multivariate effect for gender [Wilks's  
35  
36 Lambda = .86,  $F(4, 642) = 26.82, p < .001, \eta^2 = .14$ ]. Like for Study 2, univariate ANOVAs  
37  
38 revealed significant effects for all variables. Females reported higher levels of *Fantasy*,  
39  
40 *Perspective Taking*, *Empathic Concern* and *Personal Distress* than males.  
41  
42  
43

### 44 ***Intercorrelations between the B-IRI and other study variables***

45  
46 Structural equation modeling (SEM) was used to examine the associations between the B-  
47  
48 IRI subscales and other study variables. All constructs were represented as latent variables using  
49  
50 two parceled indicators. Each parcel was the sum of a subset of scale items, and each item from  
51  
52 the scale was assigned to one of the parcels (Bandalos, 2002). Although use of parcels is a  
53  
54 subject of some controversy (e.g., Little, Cunningham, Shahar, & Widaman, 2002), parcels  
55  
56  
57  
58  
59  
60



1  
2  
3 provide a more stable set of manifest variables on which to base structural models than do the  
4  
5 individual scale scores.  
6

7  
8 Goodness of fit indexes showed that the model fitted the data well:  $\chi^2 (76) = 233.93$ , SB  
9  
10  $\chi^2 (76) = 213.96$ ,  $p < .001$ , robust CFI = .952, RMSEA = .046. All factor loadings were  
11  
12 significant; they ranged from .59 to .90. Table 5 shows the interrelations among the latent  
13  
14 variables. Both *Fantasy* and *Empathic Concern* correlated positively with femininity and  
15  
16 emotional fragility, and negatively with self-other differentiation and masculinity. *Perspective*  
17  
18 *Taking* correlated positively with femininity and emotional fragility, and negatively with self-  
19  
20 other differentiation, and did not correlate with masculinity. Finally, *Personal Distress* correlated  
21  
22 positively with emotional fragility, and negatively with self-other differentiation, and did not  
23  
24 correlate with femininity.  
25  
26  
27

28  
29 (insert Table 5 about here)  
30

### 31 32 **General discussion** 33

34 Over the course of a series of studies, we developed a brief form of the IRI, then tested its  
35  
36 factor structure in three independent samples, assessed its measurement invariance across gender  
37  
38 and age groups, examined its reliability, and produced evidence of its construct validity by  
39  
40 examining whether scores on the B-IRI subscales showed any gender- and age-related  
41  
42 differences, and patterns of associations consistent with theoretical accounts of empathy and  
43  
44 prior empirical studies. We succeeded in reducing the length of the scale compared to the IRI,  
45  
46 while preserving a reasonable scale reliability and validity.  
47  
48  
49

50 The B-IRI is a 16-item instrument consisting of four subscales, each of which includes  
51  
52 four positively-worded items, since they seem to require fewer literacy skills than negatively-  
53  
54  
55  
56  
57  
58  
59  
60



worded sentences (Beven et al. 2004; Lautherbach & Hosser, 2007). This makes the B-IRI easier to use, and more appropriate than the IRI, with special populations who have low literacy skills.

Our results confirmed a four-factor structure for the B-IRI, replicating the structure hypothesized by Davis (1980, 1983) for the IRI. The model showed a good fit to the data without adding any residual variance to improve the fit (Albiero et al., 2006; De Corte et al., 2007). Our results also substantially confirmed the B-IRI's measurement invariance across gender and age groups. The B-IRI seems to be psychometrically invariant between adolescents and emerging adults. Although the B-IRI was substantially invariant between males and females, we would recommend that investigators consider the role of gender when applying the B-IRI because our results revealed a partial residual variance invariance. This finding suggests that B-IRI item 14 seems to be more reliable for females than for males, as suggested by the lower residual variances revealed in our female subsample. Despite these gender-related differences, the subscales showed an overall **reasonable** reliability in terms of internal consistency.

We also provided evidence of the B-IRI's construct validity, analyzing the gender- and age-related differences in its factors and patterns of association with some of the constructs conceptually associated with empathy. Our results found evidence of significant gender-related differences in all B-IRI subscales, with females reporting higher scores on all dimensions than males, confirming past studies on the IRI (e.g., Albiero et al., 2006; Davis & Franzoi, 1991; Gilet et al., 2012; Hawk et al., 2013). No age-related differences were indicated, though other studies on the IRI have found conflicting results on this issue, with some identifying age-related differences in some IRI subscales (Gilet et al., 2012; Hawk et al., 2013), and others finding none (Quince, Parker, Wood, & Benson, 2011). Our result might be explained, however, by the fact

1  
2  
3 that only a limited range of participants' ages was considered in our studies, and this may have  
4  
5 prevented any consistent differences from becoming apparent.  
6  
7

8 As expected, *Fantasy* was associated positively with emotional fragility and femininity,  
9  
10 and negatively with self-other differentiation, while it was unexpectedly unrelated with  
11  
12 masculinity. As hypothesized, *Perspective Taking* was associated positively with femininity,  
13  
14 negatively with self-other differentiation, and unrelated with masculinity, but it also surprisingly  
15  
16 correlated positively with emotional fragility. *Empathic Concern* was associated positively with  
17  
18 femininity and emotional fragility, and negatively with masculinity and self-other differentiation,  
19  
20 as expected. *Personal Distress* was associated positively with emotional fragility and femininity,  
21  
22 and negatively with masculinity and self-other differentiation. Although the construct validity of  
23  
24 a scale cannot be judged definitively from the outcome of a single study, our results provide  
25  
26 evidence to support the B-IRI's construct validity.  
27  
28  
29  
30  
31

32 This study has some limitations. First, more constructs than those provided are needed to  
33  
34 ascertain the construct validity of the B-IRI. Previous studies have examined this issue much  
35  
36 more extensively for the IRI. Second, some goodness of fit indexes (CFI in Studies 2 and 3, and  
37  
38 chi-square difference tests in Study 2) were lower than the minimum value recommended in the  
39  
40 literature. Particularly the chi-square difference tests used in the evaluation of measurement  
41  
42 invariance across gender need to be cautiously interpreted, even if the other fit criteria used in  
43  
44 the study were satisfying. Nonetheless, more research is needed about the measurement  
45  
46 invariance of both the IRI and the B-IRI. Third, all studies reported here were conducted on  
47  
48 Italian samples, though additional research is needed to examine the applicability of the B-IRI to  
49  
50 other cultural settings. Fourth, the age range of our participants was very narrow. Future research  
51  
52 will need to investigate age-related differences in the B-IRI subscales on a wider range of age  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 groups. Fifth, this study did not investigate the stability over time of the findings generated by  
4  
5 the B-IRI. Sixth, two subscales, Empathic Concern and Perspective Taking, did not show a  
6  
7 completely adequate reliability. All these issues underline the need of further research on the B-  
8  
9 IRI psychometric properties.  
10  
11

12  
13 Short-form instruments typically contain a simple subset of items drawn from the  
14  
15 corresponding longer form of the instruments. They often have significantly weakened  
16  
17 psychometric properties in relation to their longer counterparts. The B-IRI is unusual in this  
18  
19 sense, as it is a brief instrument that retains the psychometric properties of the full instrument  
20  
21 from which it was derived. Thus, we recommend using the B-IRI instead of the IRI, especially  
22  
23 for special populations with limited literacy skills. The B-IRI is more practical to administer and  
24  
25 likely to be more appealing to a broad variety of researchers who might benefit from using this  
26  
27 tool, which balances the goal of brevity with theoretical and psychometric requirements. We  
28  
29 believe that this scale will be useful in future investigations on empathic responsiveness.  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**References**

- 1  
2  
3  
4  
5  
6 Albiero, P., Ingoglia, S., & Lo Coco, A. (2006). Contributo all'adattamento italiano  
7  
8 dell'Interpersonal Reactivity Index di Davis. *Testing Psicometria Metodologia*, *13*(2),  
9  
10 107-125.
- 11  
12  
13 Alterman, A. I., McDermott, P. A., Cacciola, J. S., & Rutherford, M. J. (2003). Latent structure of  
14  
15 the Davis Interpersonal Reactivity Index in methadone maintenance patients. *Journal of*  
16  
17 *Psychopathology and Behavioral Assessment*, *25*, 257-265.
- 18  
19  
20 Anastasi, A., & Urbina, S. (1997). *Psychological Testing*. Upper Saddle River, NJ: Prentice-  
21  
22 Hall. .
- 23  
24  
25 Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and*  
26  
27 *Clinical Psychology*, *42*, 155-162.
- 28  
29  
30 Bentler, P. M. (2006). *EQS 6 Structural Equations Program Manual*. Encino: CA. Multivariate  
31  
32 Software, Inc.
- 33  
34  
35 Beven, J. P., O'Brien-Malone, A., & Hall, G. (2004). Using the interpersonal reactivity index to  
36  
37 assess empathy in violent offenders. *International Journal of Forensic Psychology*, *1*(2),  
38  
39 33-41.
- 40  
41  
42 Bode, R. K., Smith, E. V., & Smith, R. M. (2004). Partial credit model and pivot anchoring.  
43  
44 *Introduction to Rasch Measurement*, 279-295.
- 45  
46  
47 Caprara, G. V., Perugini, M., Barbaranelli, C., & Pastorelli, C. (1991). *Scala per la misura della*  
48  
49 *fragilità emotiva*. Firenze, Italy: Organizzazioni Speciali.
- 50  
51  
52 Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*,  
53  
54 *1*(2), 245-276.
- 55  
56  
57  
58  
59  
60

BRIEF IRI

28

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance.

*Structural Equation Modeling, 14*, 464–504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing

measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures.

*Cognitive Psychology, 3*, 472-517.

Cliffordson, C. (2001). The hierarchical structure of empathy: Dimensional organization and

relations to social functioning. *Scandinavian Journal of Psychology, 43*, 49-59.

Cliffordson, C. (2002). Parents' judgments and students' self-judgments of empathy. *European*

*Journal of Psychological Assessment, 17*, 36-47.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS*

*Catalog of Selected Documents in Psychology, 10*, 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a

multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126.

Davis, M. H. (1996). Interpersonal reactivity index. *Empathy: A Social Psychological Approach*,

55-6.

Davis, M. H., & Franzoi, S. L. (1991). Stability and change in adolescent self-consciousness and

empathy. *Journal of Research in Personality, 25*(1), 70-87.

De Corte, K., Buysse, A., Verhofstadt, L. L., Roeyers, H., Ponnet, K., & Davis, M. H. (2007).

Measuring empathic tendencies: Reliability and validity of the Dutch version of the

Interpersonal Reactivity Index. *Psychologica Belgica, 47*(4), 235-260.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral*

*and Cognitive Neuroscience Reviews, 3*(2), 71-100.

- 1  
2  
3 Decety, J., & Lamm, C. (2006). Human empathy through the lens of social neuroscience. *The*  
4  
5  
6 *Scientific World Journal*, 6, 1146-1163.  
7
- 8 Eisenberg, N., & Eggum, N. D. (2009). Empathic responding: Sympathy and personal  
9  
10 distress. *The Social Neuroscience of Empathy*, 71-83.  
11
- 12 Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial Development. In N. Eisenberg,  
13  
14 W. Damon, & R.M Lerner (Eds.). *Handbook of Child Psychology. Vol. 3, Social,*  
15  
16 *Emotional, and Personality Development (6th ed.)* (pp. 646-718). Hoboken, NJ: Wiley.  
17
- 18 Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related  
19  
20 capacities. *Psychological Bulletin*, 94(1), 100.  
21  
22
- 23 Feshbach, N. D., & Feshbach, S. (2009). Empathy and education. In J. Decety & W. J. Ickes  
24  
25 (Eds.), *The Social Neuroscience of Empathy* (pp 85–98). Cambridge: The MIT Press.  
26  
27
- 28 Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of  
29  
30 clinical assessment instruments. *Psychological Assessment*, 7(3), 286.  
31  
32
- 33 Gilet, A. L., Mella, N., Studer, J., Grünh, D., & Labouvie-Vief, G. (2013). Assessing  
34  
35 dispositional empathy in adults: A French validation of the Interpersonal Reactivity Index  
36  
37 (IRI). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du*  
38  
39 *comportement*, 45(1), 42.  
40  
41
- 42 Gini, G., Albiero, P., Benelli, B., & Altoè, G. (2007). Does empathy predict adolescents' bullying  
43  
44 and defending behavior? *Aggressive Behavior*, 33, 1-10.  
45  
46
- 47 Hawk, S. T., Keijsers, L., Branje, S. J., Graaff, J. V. D., Wied, M. D., & Meeus, W. (2013).  
48  
49 Examining the Interpersonal Reactivity Index (IRI) among early and late adolescents and  
50  
51 their mothers. *Journal of Personality Assessment*, 95(1), 96-106.  
52  
53
- 54 Hoffman, M. L. (2008). Empathy and prosocial behavior. *Handbook of Emotions*, 3, 440-455.  
55  
56  
57  
58  
59  
60

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179-185.
- Huang, X., Li, W., Sun, B., Chen, H., & Davis, M. H. (2012). The validation of the Interpersonal Reactivity Index for Chinese teachers from primary and middle schools. *Journal of Psychoeducational Assessment*, *30*(2), 194-204.
- Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior*, *9*, 441-476.
- Kaiser, H. F. (1961). A note on Guttman's lower bound for the number of common factors1. *British Journal of Statistical Psychology*, *14*(1), 1-2.
- Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural equation modeling*, *10*(3), 333-351.
- Ingoglia, S., Coco, A. L., Liga, F., & Cricchio, M. G. L. (2011). Emotional separation and detachment as two distinct dimensions of parent-adolescent relationships. *International Journal of Behavioral Development*, *35*(3), 271-281.
- Lauterbach, O., & Hosser, D. (2007). Assessing empathy in prisoners: A shortened version of the Interpersonal Reactivity Index. *Swiss Journal of Psychology*, *66*(2), 91-101.
- Litvack-Miller, W., McDougall, D., & Romney, D. M. (1997). The structure of empathy during middle childhood and its relationship to prosocial behavior. *Genetic, Social, and General Psychology Monographs*, *123*(3), 303-324.
- Muris, P., Meesters, C., & Gobel, M. (2001). Reliability, validity, and normative data of the Penn State Worry Questionnaire in 8–12-yr-old children. *Journal of Behavior Therapy and Experimental Psychiatry*, *32*(2), 63-72.
- Musick, M. A., & Wilson, J. (2003). Volunteering and depression: The role of psychological and

- 1  
2  
3 social resources in different age groups. *Social Science & Medicine*, 56(2), 259-269.
- 4  
5  
6 Muthén, B., & Kaplan, D. (1985). A comparison of methodologies for the factor analysis of non-  
7  
8 normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, 38,  
9  
10 171-189.
- 11  
12  
13 O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components  
14  
15 using parallel analysis and Velicer's MAP test. *Behavior research methods, instruments,*  
16  
17 & computers, 32(3), 396-402.
- 18  
19  
20 Olver, R. R., Aries, E., & Batgos, J. (1989). Self-other differentiation and the mother-child  
21  
22 relationship: The effects of sex and birth order. *Journal of Genetic Psychology*, 150(3),  
23  
24 311-321.
- 25  
26  
27 Pedersen, R. (2009). Empirical research on empathy in medicine: A critical review. *Patient*  
28  
29 *Education and Counseling*, 76(3), 307-322.
- 30  
31  
32 Poulos, S., Elison, J., & Lennon, R. (2004). The hierarchical structure of the Interpersonal  
33  
34 Reactivity Index. *Social Behavior and Personality*, 32, 355-360.
- 35  
36  
37 Quince, T. A., Parker, R. A., Wood, D. F., & Benson, J. A. (2011). Stability of empathy among  
38  
39 undergraduate medical students: A longitudinal study at one UK medical school. *BMC*  
40  
41 *Medical Education*, 11(1), 90.
- 42  
43  
44 Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in  
45  
46 covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent Variables*  
47  
48 *Analysis: Applications for Developmental Research* (pp. 399-419). Thousand Oaks, CA:  
49  
50 Sage.
- 51  
52  
53 Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment  
54  
55 structure analysis. *Psychometrika*, 66(4), 507-514.
- 56  
57  
58  
59  
60



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Siu, A. M., & Shek, D. T. (2005). Validation of the Interpersonal Reactivity Index in a Chinese context. *Research on Social Work Practice, 15*(2), 118-126.

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12*(1), 102.

Streiner, D. L., Norman, G. R., & Cairney, J. (2014). *Health measurement scales: A practical guide to their development and use*. Oxford: Oxford University Press.

Tello, F. P. H., Egido, B. D., Ortiz, M. A. C., & Gandara, M. D. B. (2013). Interpersonal Reactivity Index: Analysis of invariance and gender differences in Spanish youths. *Child Psychiatry & Human Development, 44*(2), 320-333.

Yu, J., & Kirk, M. (2009). Evaluation of empathy measurement tools in nursing: Systematic review. *Journal of Advanced Nursing, 65*(9), 1790-1806.

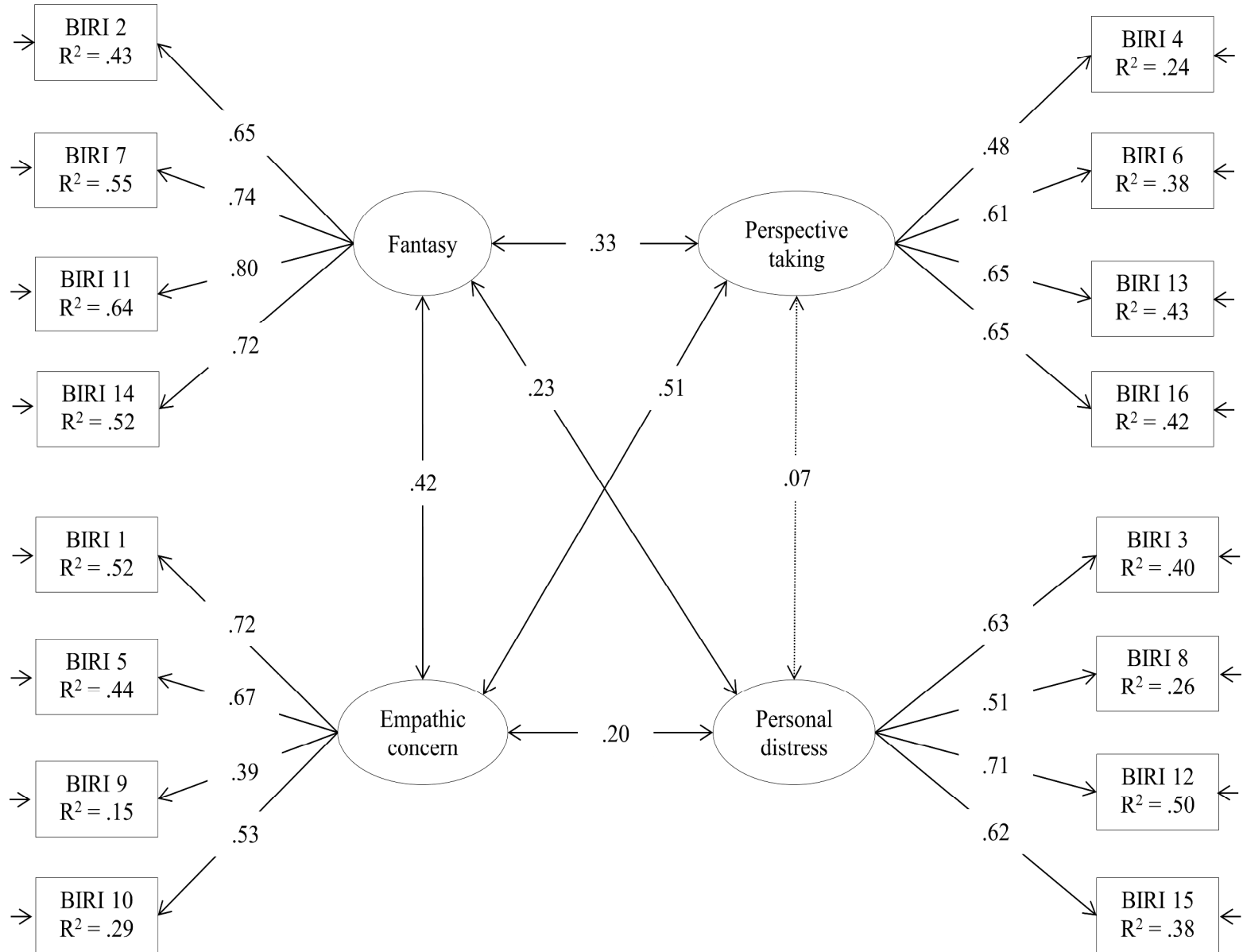
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Figure Captions

*Figure 1.* CFA on the B-IRI items. Standardized solution ( $n = 836$ ).

*Note.* All parameter estimates were significant at  $p < .05$ , except that represented with a dotted line.

For Peer Review Only



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

BRIEF IRI

Table 1

Mean, standard deviation, skewness and kurtosis of the B-IRI items, and Pearson product-moment correlation coefficients (n = 1,104).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1															
2	.24	1														
3	.22	.24	1													
4	.07	-.00	-.03	1												
5	.47	.17	.15	.14	1											
6	.27	.18	.10	.27	.21	1										
7	.24	.45	.22	-.02	.20	.20	1									
8	.16	.17	.46	-.06	.13	.10	.25	1								
9	.28	.15	.04	.04	.38	.12	.14	.05	1							
10	.34	.20	.21	.08	.33	.21	.15	.18	.25	1						
11	.16	.46	.17	-.00	.16	.16	.59	.21	.12	.23	1					
12	.11	.15	.41	-.07	.11	.07	.15	.35	.04	.16	.14	1				
13	.22	.15	.07	.30	.18	.38	.14	.05	.13	.20	.10	.06	1			
14	.25	.45	.19	.03	.20	.25	.50	.17	.17	.22	.47	.09	.23	1		

## BRIEF IRI

31

15	.10	.18	.36	-.04	.03	.03	.12	.32	-.08	.07	.12	.43	.03	.11	1	1
16	.25	.15	.07	.22	.29	.30	.13	.09	.19	.25	.10	.08	.44	.14	.06	1
M	3.60	3.47	3.22	3.64	2.99	3.06	3.28	3.39	3.78	3.08	3.54	3.66	3.11	3.50	2.91	2.86
SD	1.00	0.88	0.90	0.97	1.09	1.08	1.00	1.09	0.91	1.06	0.91	1.11	1.13	0.99	1.08	1.06
Skeness	-0.27	-0.14	-0.07	-0.54	-0.61	-0.43	-0.06	0.09	-1.10	-0.60	-0.27	0.39	0.04	-0.35	0.79	0.00
Kurtosis	-0.05	-0.70	-0.69	-0.365	0.15	0.01	-0.61	-0.64	0.86	0.11	-0.42	-0.31	-0.52	-0.52	0.31	-0.58

*Note.* Correlation coefficients higher than .10 were significant at  $p < .05$ .

## BRIEF IRI

Table 2

*Results of the Exploratory Factor Analysis on the full B-IRI items (n = 1,104).*

	Fantasy	Empathic Concern	Perspective Taking	Personal Distress
Eigenvalues	2.04	3.92	1.28	1.57
Variance before rotation	8.98%	21.01%	4.20%	6.46%
7	<b>.77</b>	.01	.03	.01
11	<b>.77</b>	.02	.06	.01
14	<b>.66</b>	.02	.09	.03
2	<b>.58</b>	.03	.02	.07
5	.05	<b>.74</b>	.01	.02
1	.04	<b>.55</b>	.07	.10
9	.04	<b>.54</b>	.04	.11
10	.05	<b>.41</b>	.10	.13
13	.06	.05	<b>.73</b>	.03
6	.14	.04	<b>.50</b>	.03
16	.02	.18	<b>.49</b>	.06
4	.07	.02	<b>.47</b>	.08
12	.05	.02	.01	<b>.66</b>
3	.03	.09	.03	<b>.64</b>
15	.01	.13	.03	<b>.63</b>
8	.07	.06	.04	<b>.55</b>

## BRIEF IRI

33

Table 3

*Multigroup tests of the B-IRI model invariance across gender and age (males = 401 vs. females = 435, adolescents = 354 vs. emerging adults = 482).*

<i>Gender invariance</i>	$\chi^2$	df	SB $\chi^2$	CFI	RMSEA	$\Delta\chi^2$	$\Delta$ df	$\Delta$ CFI	$\Delta$ RMSEA
M0 Configural invariance	456.75***	196	366.86***	.929	.046	-	-	-	-
M1 Full metric invariance	490.66***	212	396.17***	.923	.046	30.36*	16	-.006	-
M2 Factor variance/covariance invariance	502.68***	218	405.58***	.922	.045	38.62	22	-.007	-.001
M3 Intercept invariance	696.43***	234	581.07***	.928	.049	250.00***	38	-.001	.003
M4 Residual variance invariance	779.90***	250	636.10***	.915	.051	279.24***	54	.014	.005
M4a Partial residual variance invariance	752.74***	249	618.31***	.921	.050	265.39***	53	-.008	.004
<i>Age invariance</i>	$\chi^2$	df	SB $\chi^2$	CFI	RMSEA	$\Delta\chi^2$	$\Delta$ df	$\Delta$ CFI	$\Delta$ RMSEA
M0 Configural invariance	456.31***	196	356.15***	.937	.044	-	-	-	-
M1 Full metric invariance	493.40***	212	387.51***	.931	.045	31.36*	16	-.006	.001
M2 Factor variance/covariance invariance	497.98***	218	391.57***	.932	.044	34.95*	22	-.005	-
M3 Intercept invariance	533.80***	234	428.56***	.930	.045	72.79**	38	-.007	.001
M4 Residual variance invariance	549.54***	250	439.72***	.932	.043	81.95**	54	-.005	-.001

\*  $p < .05$ ; \*\*\*  $p < .001$ .

BRIEF IRI

34

Table 4

*Means, standard deviations of the B-IRI subscales by gender,  $F$  and  $\eta^2$  values ( $n = 836$ ).*

	Males		Females		$F(1,832)$	Cohen's $d$
	M	SD	M	SD		
Fantasy	2.86	0.83	3.28	0.80	57.61***	.53
Empathic concern	3.71	0.72	4.06	0.56	65.98***	.56
Perspective taking	3.31	0.73	3.47	0.64	12.17***	.24
Personal distress	2.26	0.68	2.67	0.69	71.19***	.59

\*\*\*  $p < .001$ .



## BRIEF IRI

35

Table 5

*Correlations between latent variables in the model tested for the construct validity study (n = 649).*

		1	2	3	4	5	6	7	8
<b>B-IRI</b>	1 Fantasy	1							
	2 Empathic concern	.46***	1						
	3 Perspective taking	.31***	.51***	1					
	4 Personal distress	.29***	.29***	.11	1				
<b>SODS</b>	5 Self-other differentiation	-.24***	-.15**	-.13**	-.44***	1			
<b>SMEF</b>	6 Emotional fragility	.37***	.36***	.15**	.62***	-.41***	1		
<b>BSRI</b>	7 Masculinity	-.16**	-.19**	-.11	-.46***	.43***	-.44***	1	
	8 Femininity	.30***	.83***	.38***	.07	-.10	.17***	.10	1

\*\*  $p < .01$ ; \*\*\*  $p < .001$ .

BRIEF IRI

**Appendix.****The Brief Interpersonal Reactivity Index (subscale and IRI item number)**

- 
1. I often have tender, concerned feelings for people less fortunate than me. (EC - 2)
  2. I really get involved with the feelings of the characters in a novel. (F - 5)
  3. In emergency situations, I feel apprehensive and ill-at-ease. (PD - 6)
  4. I try to look at everybody's side of a disagreement before I make a decision. (PT - 8)
  5. When I see someone being taken advantage of, I feel kind of protective towards them. (EC - 9)
  6. I sometimes try to understand my friends better by imagining how things look from their perspective. (PT - 11)
  7. After seeing a play or movie, I have felt as though I were one of the characters. (F - 16)
  8. Being in a tense emotional situation scares me. (PD - 17)
  9. When I see someone being treated unfairly, I feel very much pity for them. (EC - 18 originally expressed as a reversed item)
  10. I would describe myself as a pretty soft-hearted person. (EC - 22)
  11. When I watch a good movie, I can very easily put myself in the place of a leading character. (F-23)
  12. I tend to lose control during emergencies. (PD - 24)
  13. When I'm upset at someone, I usually try to "put myself in his shoes" for a while. (PT - 25)
  14. When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were happening to me. (F - 26)
  15. When I see someone who badly needs help in an emergency, I go to pieces. (PD - 27)
  16. Before criticizing somebody, I try to imagine how I would feel if I were in their place. (PT - 28)
-