International Journal of Psychological Studies; Vol. 9, No. 1; 2017
ISSN 1918-7211 E-ISSN 1918-722X
Published by Canadian Center of Science and Education

Mentalizing: Seeking the Underlying Dimensions

Marjaana Lindeman[1] & Jari Lipsanen[1]

[1] Division of Cognitive Psychology and Neuropsychology, Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland

Correspondence: Marjaana Lindeman, Division of Cognitive Psychology and Neuropsychology, Institute of Behavioural Sciences, University of Helsinki, P.O. Box 9, 00014, Helsinki, Finland. Tel: 358-50-448-4162. E-mail: marjaana.lindeman@helsinki.fi

Received: November 3, 2016 Accepted: November 26, 2016 Online Published: December 5, 2016

doi:10.5539/ijps.v9n1p10 URL: http://dx.doi.org/10.5539/ijps.v9n1p10

Abstract

Mentalizing (i.e., theory of mind) is a much studied construct, but the way different forms of mentalizing are related to each other is poorly understood. In this study (N = 369), we examined the dimensionality of mentalizing by addressing several forms of cognitive and affective empathy, practical mentalizing skills (i.e., understanding figurative language and social etiquette), and metacognition. The results of principal component analysis showed that sixteen mentalizing variables could be reduced to four principal components, namely affective empathy, social skills, self-insight, and views about the nature of beliefs. The components were unrelated, suggesting that they are independent aspects of mentalizing. No general mentalizing factor or overall empathy was found indicating that mentalizing is a non-hierarchical profile construct.

Keywords: empathy, mentalizing, theory of mind

1. Introduction

Mentalizing is a vital part of our lives: we think and talk much about our own and other people's thoughts, emotions, and behaviour. Children understand mental states in a different way than adults, and mentally disturbed individuals understand them in a different way than healthy adults. For example, small children may think that a thought is "white and round" (Piaget, 1929/1951, p. 51), and individuals with schizophrenia have difficulties in inferring what other people think (Brüne & Brüne-Cohrs, 2006). During the last decades, research on mentalizing has increased strongly, most probably because of the theoretical advancements in the cognitive capacity known as theory of mind. Besides developmental and clinical psychology, mentalizing among healthy people is a hot topic, addressed in several fields of psychology, including cognitive, personality, organizational, and educational psychology as well as brain science. However, the terminology is vague and a long list of constructs has been argued to characterize mentalizing. Similarly, a host of methods are in use for assessing individual differences in mentalizing, but the way the underlying constructs aggregate is not known. In the present study, we will therefore examine how diverse forms of mentalizing link to each other.

Undoubtedly, we are dealing with a multidimensional construct when we speak about mentalizing (also known as theory of mind). It can refer to cognitive and affective empathy, understanding communicative signs, the ability to interpret eye-gaze cues, understanding of the self, detecting intentionality, and understanding social outcomes, to mention but a few. To our knowledge, studies that have broadly examined how different dimensions of mentalizing are related to each other do not exist. We do not know, for example, do people who rapidly discern other people's emotions also behave tactfully in social situations, or does understanding other people's thoughts necessarily imply understanding one's own thoughts? Yet, understanding whether various aspects of mentalizing are conceptually distinct has important implications for the development of theory and assessment in the field.

The number of dimensions of mentalizing that can be explored in one study is necessarily limited. Of all the dimensions that have been brought up in the literature, we ended up with four higher order constructs, namely cognitive and affective empathy, practical skills and metacognition, and their 14 lower-order facets, all discussed below (see also Table S1 at the Appendix). The division is largely based on Apperly's (2012) theoretical model that describes different, and too often isolated approaches in recent research on theory of mind.

1.1 Empathy

One of the most examined dimensions of adults' mentalizing is empathy. Although several researchers agree that empathy is a multidimensional construct, there is no consensus about what those dimensions are. It is now quite generally accepted, however, that empathy comprises at least cognitive empathy and affective (emotional) empathy (e.g., Baron-Cohen & Wheelwright, 2004; Davis, 2006; Decety & Jackson, 2004; Shamay-Tsoory, Tomer, Goldsher, Berger, & Aharon-Peretz, 2004). Affective empathy involves an emotional response to the emotional responses of others, especially the sharing of other people's emotions and the tendency to experience personal unease when witnessing the distress of other people. Cognitive empathy involves identifying another's mental states. Based on findings from cognitive neuroscience, Decety (2011) has proposed a widely-cited model about two basic modes of empathic processes, closely resembling the ideas of affective and cognitive empathy. According to the model, bottom-up processes are shared with other primates and they include sharing the emotional experience of the other person and the capacity for two people to resonate with each other emotionally. The evolutionarily newer, higher level top-down processes include motivational and self-regulatory processes which enable prosocial behaviour and understanding of the other person's experience.

Several sub-dimensions of affective and cognitive empathy have also been put forward. In Davis's (1983) multidimensional empathy theory, cognitive empathy is divided into perspective taking (the tendency to spontaneously adopt the psychological point of view of others) and to fantasy (the tendency to transpose oneself imaginatively into the feelings and actions of fictitious characters in books, movies, and plays). Affective empathy is divided into empathic concern (other-oriented feelings of sympathy and concern for unfortunate others) and personal distress (anxiety and unease in tense interpersonal settings). Both fantasy and perspective taking can be assessed with the widely-used empathy questionnaire, Interpersonal Reactivity Index.

In turn, based on neuroscientific and developmental studies of empathy, Carré et al. (2013) have broken empathy into cognitive empathy and two types of affective empathy, namely emotional contagion and emotional disconnection. Emotional contagion refers to automatic replication of another person's emotions (e.g., facial mimicry) and to a rapid evaluation whether the emotion is pleasant or aversive. Emotional disconnection relates to withdrawal from emotion, that is, to regulatory functions which protect an individual from excessive emotions (To keep all the empathy concepts parallel, we use here the concept of emotional connection).

Finally, Muncer and Ling (2006) have suggested on empirical grounds that the short form of the popular empathy questionnaire, Empathy Quotient (EQ), measures three aspects of empathy: cognitive empathy, emotional reactivity (close to affective empathy), and social skills (close to practical mentalizing skills).

1.2 Practical Mentalizing Skills

Mentalizing can also manifest itself as social skills like understanding metaphors, irony and faux pas pas (Brüne & Brüne-Cohrs, 2006; Happé, 1993; Hooker, Verosky, Germine, Knight, & D'Esposito, 2008). Apperly (2012) has characterized comprehension of ironical and metaphorical speech as practical mentalizing skills because it reflects the ability to put mentalizing concepts into use, that is, to use theory of mind abilities in a flexible, pragmatically appropriate and context-sensitive manner. Supporting this argument, *individuals who have deficits in mentalizing*, *also have difficulties in understanding figurative language* (Langdon & Coltheart, 2004). *However*, *it is possible that* irony is processed differently than metaphors because a meta-analysis of fMRI studies showed that only irony requires the ability to make inferences regarding other peoples' thoughts, whereas metaphor comprehension requires mainly semantic processes and bears on the same neural network as literal processing (Bohrn, Altmann, & Jacobs, 2012).

In addition, several scholars have proposed that understanding tactless acts, slips and blunders in etiquette, and their emotional consequences in social situations (i.e., faux pas) reflect mentalizing skills put into practical use (e.g., Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999; Brüne & Brüne-Cohrs, 2006). Again, those *who have deficits in mentalizing*, *for example* children with Asperger syndrome (Baron-Cohen et al., 1999), *have difficulties in faux pas*. As for understanding metaphors and irony, the way understanding social etiquette is related to the other types of adults' mentalizing, for example empathy, is not known.

1.3 Metacognition

Although metacognition (thinking about thinking) has seldom been studied together with the other aspects of healthy adults' mentalizing, many researchers emphasize its importance. Metacognitive awareness is argued to evince conceptual change in the mentalizing domain (Apperly, 2012), to regulate and reflect on first-order cognition (Kuhn, 2000), and to be an important issue in discussions about theory of mind and the evolution of social intelligence (Smith, Shields, & Washburn, 2003).

We will here examine two types of metacognition, namely the way individuals understand their own thinking (i.e., insight) and the views about beliefs and knowledge in general. Based on Beck, Baruch, Balter, Steer, and Warman's (2004) work, we conceptualize insight as overconfidence in the validity of one's beliefs ("self-certainty") and the capacity and willingness to observe one's own judgments and to consider alternative explanations ("self-reflectiveness"). Views about the nature of beliefs and knowledge can also take two forms. According to Kuhn, Cheney, and Weinstock (2000), most adults are "multiplists" in that they understand that beliefs and knowledge are uncertain and that divergent claims are legitimate. Reaching the developmental endpoint of epistemological understanding, the evaluativist level, is more unusual: no more than half of adults accept that one belief can be more right than the other if that belief is better supported by argument and evidence (Kuhn et al., 2000).

1.4 The Present Study

The above discussion clearly indicates that mentalizing is a construct that involves more than one dimension. The present study is an initial step toward empirical exploration of the nature and number of the relevant dimensions of mentalizing.

Mentalizing can be multidimensional in four different ways (cf., Law, Wong, & Mobley, 1998). First, it is possible that the dimensions are independent and that they can be summed up to form a composite mentalistic cognition score. An analogical example is the concept of overall job satisfaction, which is composed of independent dimensions such as contentment with pay, co-workers and supervisor. The second alternative is that there is a latent commonality underlying the different aspects of mentalistic cognition. In this case, the dimensions are manifestations of a higher-order construct, in the same way as some theories of intelligence maintain that specific factors are manifestations of general intelligence. Third, mentalizing can be a profile construct, composed of separate forms of mentalizing which cannot be combined together into a higher-order construct. A corresponding example is personality, which may be characterized with several dimensions such as openness and extraversion, but there is no overall personality on which a person can be high or low.

In this study, the dimensionality of mentalizing is analysed with principal component analysis. Determining potential higher-order constructs is important because it allows for more theoretical parsimony and reduces complexity (Law et al., 1998). However, identifying independent constructs is also important. Narrow constructs can be psychologically more meaningful than broad constructs in explaining behaviour, and assessment of narrow constructs can increase the accuracy of prediction of individual differences (Paunonen, Rothstein, & Jackson, 1999).

2. Method

2.1 Participants and Procedure

The participants were 369 individuals (mean age = 32 years, age range 16-67; 62.1% females). Of the respondents, 37.1% were working, 43% were students, and 16.8% were employed in other activities. Fifty-four per cent had a polytechnic or a university degree, 42.8% had finished vocational school or an upper secondary school, and the others had finished grammar school.

The online data were collected in two stages. In the first stage, the participants were recruited mainly via several open internet discussion forums. In this stage, the participants filled in the Empathy Quotient Short, the Pictorial Empathy Test, and the Reading the Mind in the Eyes Test (see below). For the purpose of this study, additional data were collected 1.5 years later from the pool of individuals who participated in the first stage. The recruitment message was mailed to all individuals who participated in the first stage and who had given their e-mail address for participating in further studies (N = 1537). Of them, 237 could not be contacted because of an outdated e-mail address and 887 did not take part in the present study.

Of the 413 people who originally took part in the study, 44 were not included, because they responded to at least one of the six attention check items. We included two sets of 3-item attention checks at two points in the survey (e.g., "I like to watch cloud shapes change in the sky") with the following instruction: "The following three statements have the same response scale as before. These are statements which you should not respond to at all. They are only involved because we want to check if you read the instructions and answer the questions carefully". The respondents were given 3 weeks time to participate in the study.

2.2 Pilot Study

An online pilot study was conducted to adjust and modify those methods which have not been much used, either overall or in the context of mentalizing (see Epistemic Understanding Questionnaire, and understanding metaphors, irony and social etiquette). Only the final form of these scales are described below, for a more

detailed description of the development of the scales, see at the Appendix. The participants in the pilot study were recruited via electronic student mailing list and an Open University Psychology course. In contrast to the university courses, Open University courses in Finland are fee-charging courses available for all, regardless or age, citizenship or educational background. Thirty-three individuals participated (mean age = 29.7 yrs., age range 20-64; 29 females).

2.3 Material

2.3.1 Empathy

Self-reported empathizing was measured with the 15-item version of the Empathy Quotient Short, EQ-Short (Muncer & Ling, 2006). The EQ-Short includes three subscales, *Cognitive Empathy* (e.g., "I can tune into how someone else feels rapidly and intuitively"), *Social Skills* (e.g., "I do not tend to find social situations confusing"), and *Emotional Reactivity* (e.g., "I tend to get emotionally involved with a friend's problems"). The scale was scored using the normative weights (Baron-Cohen & Wheelwright, 2004): the response format (1 = strongly disagree, 2 = slightly disagree, 3 = slightly agree, 4 = strongly agree) was converted into scores of 0, 0, 1, and 2. The rationale for the scoring is that participants score 2 points only if they display an empathizing response strongly and 1 point if they display an empathizing response slightly.

We also used the Basic Empathy Scale in Adults (BES-A, Carré et al., 2013). BES-A includes 20 items that the participants rate on a five-point scale (1 = *strongly disagree*, 5 = *strongly agree*). The three subscales of the BES-A were used in the analyses: *Emotional Contagion* (six items, e.g., "I get caught up in other people's feelings easily"), *Emotional Connection* (six items, e.g., "My friends' emotions don't affect me much", reversed), and *Cognitive Empathy* (eight items, e.g., "When someone is feeling 'down' I can usually understand how they feel").

The third empathy scale we used was the Interpersonal Reactivity Index (IRI, Davis, 1983). IRI includes 28 five-point items (0 = *does not describe me well*, 4 = *describes me very well*) and it includes four 7-item subscales. Two of them assess cognitive empathy, namely *Fantasy* (e.g., "I really get involved with the feelings of the characters in a novel") and *Perspective Taking* (e.g., "I try to look at everybody's side of a disagreement before I make a decision"). Two subscales assess affective empathy: *Empathic Concern* (e.g., "I am often quite touched by things that I see happen") and *Personal Distress* (e.g., "In emergency situations, I feel apprehensive and ill-at-ease").

Affective empathy was also measured by the Pictorial Empathy Test (PET, Lindeman, Koirikivi, & Lipsanen, 2016). PET includes seven photographs, depicting men, women, and children feeling sad, fearful or in pain, or variations of these emotions. The photographs were presented consecutively on the screen together with the question: "How touching do you find the photograph?" (1 = *not at all*, 5 = *very much*).

Cognitive empathy was also measured by thirteen pictures from the revised version of the Reading the Mind in the Eyes Test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). The original test includes 36 photographs of the eye-region of the face of different actors and actresses, and it assesses how well individuals understand what the person in the picture is thinking or feeling. The thirteen pictures were evenly selected to represent easy, average and difficult items, based on the normative data provided by Baron-Cohen et al. (2001). For each picture, the participants were asked to choose which of the four listed emotion words, one correct and three foil terms, best describes what the person in the picture is thinking or feeling.

2.3.2 Metacognition

Metacognition was assessed with two scales, the Beck Cognitive Insight Scale (BCIS, Beck et al., 2004) and the Epistemic Understanding Questionnaire (EUQ, Kuhn et al., 2000). BCIS includes 15 four-point items (0 = *do not agree at all*, 3 = *agree completely*) and two subscales: *Self-Certainty* (e.g., "I can trust my own judgment at all times") and *Self-Reflectiveness* (e.g., "I have jumped to conclusions too fast").

EUQ included nine pairs of contrasting claims, attributed to two individuals, Robin and Chris (e.g., "Robin thinks people should take responsibility for themselves. Chris thinks people should work together to take care of each other"). After presenting each pair of claims, the participants were told that both may have some rightness and they were then asked: "Could one view be better or more right than the other?" The response options were "One could not be more right than the other" (reflecting the multiplist level, scored as 1) and "One could be more right" (reflecting the evaluativist level, scored as 2).

2.3.3 Practical Skills

Besides the above-mentioned Social Skills subscale of the EQ-Short, skills to put mentalizing abilities into practical use were operationalized as understanding metaphors and irony, and as faux pas. *Understanding metaphors* and *irony* were assessed with a task, modified after the lexical decision task (Langdon & Coltheart, 2004). The participants were presented 15 stories which ended either with a metaphorical (six stories), an ironic (six stories), or a literal statement (three filler stories) made by one of the story characters. After each story, we presented two stimulus words. The participants were required to rapidly assess how well the words corresponded to what the story character wanted to say (1 = not related, 2 = partly related, 3 = related). One of the two stimulus words corresponded either to the figurative interpretation or to the literal interpretation of the statement. The other stimulus word was a filler word which was either related or unrelated with the statement. Because we had a 3-point scale and six stories for both metaphors and ironies, the variables for *understanding metaphors and understanding irony* could both range from 1 to 18. For examples of the stories and the stimulus words, see Table S2 at the Appendix.

To assess *understanding social etiquette*, we surfed the Internet and gathered the most commonly listed etiquette mistakes to develop the Social Etiquette Scale (SES). The final scale included 11 common etiquette mistakes (e.g., "I cut in line" and "I am late for appointments"). The participants were told that that the statements describe bad habits which we all have at least occasionally, and they were asked: "If this were to happen would you feel you were doing wrong?" (1 = *no*, 5 = *surely*).

3. Results

All analyses were done using R-software version 3.3.0 (R Core Team, 2016). First, structural validity was examined separately for each scale using confirmatory factor analysis for Likert scale questionnaires and two parameter item response models (IRT, e.g., Embretson & Reise, 2000) for dichotomous items, i.e., for RMET and EUQ (Table 1). Due to a moderate violation of multivariate normal distribution assumption, all analyses were performed using robust maximum likelihood estimation with Huber-white standard errors. Confirmatory factor analysis was performed by using the Lavaan-package (Rosseel, 2012), and Mirt-package (Chalmers, 2012) was used in the two parameter IRT-analysis.

To achieve satisfactory fit, cross-loadings based on modification indexes were added in all multifactorial scales. All cross-loadings were small (< .30). Reliability of the scales were estimated using Cronbach alpha for Likert scales and marginal reliability coefficient (Dimitrov, 2003) for the dichotomous scales. As Table 1 shows, the fit indexes for the RMET and for understanding metaphors were not satisfactory. Also the reliabilities of these scales were low. These scales were therefore not included in the subsequent analyses.

Table 1. Confirmatory factor analysis, two parametric item response models, and reliabilities for the 18 facets of mentalizing included in the main study

		χ^2	df	χ^2/df	CFI	TLI	RMSEA	Reliability
EQ-Short		152.16	84	1.81	0.95	0.93	0.050	
	Social Skills							.74
	Cognitive Empathy							.82
	Emotional Reactivity							.63
BES-A		237.50	146	1.63	0.96	0.95	0.043	
	Cognitive Empathy							.82
	Emotional Contagion							.80
	Emotional Connection							.83
IRI		579.10	336	1.72	0.89	0.88	0.050	
	Fantasy							.78
	Perspective Taking							.76
	Empathic Concern							.80
	Personal Distress							.75

PET	19.03	11	1.73	0.99	0.98	0.047	.89
RMET[a]	59.31	53	1.12	0.79	0.74	0.019	.55
Understanding irony	31.39	17	1.85	0.95	0.91	0.055	.71
Understanding metaphors	40.21	19	2.12	0.92	0.88	0.063	.42
Understanding social etiquette	158.64	73	2.17	0.92	0.90	0.058	.86
EUQ[a]	50.17	20	2.51	0.99	0.98	0.066	.90
BCIS	180.34	85	2.12	0.92	0.90	0.059	
Self-Reflectiveness							.76
Self-Certainty							.66

Note. EQ-Short = Empathy Quotient Short; BES-A = Basic Empathy Scale in Adults; IRI = Interpersonal Reactivity Index; PET = Pictorial Empathy Test; RMET = Reading the Mind in the Eyes Test; EUQ = Epistemic Understanding Questionnaire; BCIS = Beck Cognitive Insight Scale.

[a] The results are based on two parametric (multidimensional) item response model and the reliabilities are marginal reliability coefficients.

Next, composite scores for the valid scales were generated and the relationships between the scales were examined using principal components analysis. The components were further rotated by using oblimin rotation. Because some observed mentalizing scores were highly skewed, robust correlation matrices between the composite scores were estimated using Robust-package (Wang et al., 2013) and robust Stahel-Donoho estimator (Maronna & Yohai, 1995). The correlation matrix was then analysed by principal components analysis (Table 2).

The number of components was determined by parallel analysis, which is one of the most accurate methods for determining the number of factors (a review: Hayton, Allen, & Scarpello, 2004). A four-component solution was extracted because it adequately explained the common variance. The first component loaded primarily on scales of affective empathy and was also named as such. Component two, named Social Skills, encompassed scores from scales assessing social functioning and cognitive empathy. The third component, Insight, represented Beck Cognitive Insight Scale. The fourth component was named Views about Beliefs because it encompassed only the EUQ. Correlations between the components were low (≤ .17), indicating that the components were independent. The four components explained 59% of the variance.

Table 2. Principal component analysis of 16 facets of mentalizing

	PC1	PC2	PC3	PC4	Communality	Uniqueness
BES-A Emotional Contagion	**.83**	-.19	-.14	.07	.68	.32
EQ-Short Emotional Reactivity	**.83**	-.04	-.08	.03	.69	.31
BES-A Emotional Connection	**.80**	-.07	.09	.10	.68	.32
IRI Empathic Concern	**.75**	.17	.13	-.03	.66	.34
PET	**.61**	.15	-.08	.14	.42	.58
IRI Fantasy	**.55**	-.01	.22	.37	.50	.50
BES Cognitive Empathy	**.54**	**.51**	.06	.22	.70	.30
Understanding social etiquette	**.51**	.17	-.27	.14	.32	.68
IRI Perspective Taking	.43	.30	.33	-.16	.47	.53
IRI Personal Distress	.26	**-.85**	-.01	.12	.79	.21
EQ-Short Social Skills	.05	**.84**	-.01	.12	.73	.27
EQ-Short Cognitive Empathy	.44	**.53**	.02	-.07	.56	.44
BCIS Self-Certainty	-.04	.06	**-.81**	.00	.66	.34
BCIS Self-Reflectiveness	.01	-.02	**.71**	.19	.53	.47
Understanding irony	.12	.36	.42	-.30	.24	.76

EUQ	-.11	.07	.06	**.88**	.78	.22
Eigenvalue	4.43	2.33	1.50	1.17		
Proportion variance	0.28	0.15	0.09	0.07		
Cumulative variance	0.28	0.42	0.52	0.59		

Note. Loadings of .50 or more are shown in bold. BES-A = Basic Empathy Scale in Adults; EQ-Short = Empathy Quotient Short; IRI = Interpersonal Reactivity Index; PET = Pictorial Empathy Test; BCIS = Beck Cognitive Insight Scale; EUQ = Epistemic Understanding Questionnaire.

4. Discussion

The results revealed that sixteen variables of mentalizing could be reduced to four principal components, reflecting affective empathy, social skills, insight, and views about the nature of beliefs. The components were unrelated indicating that they are independent aspects of mentalizing which do not affect each other. The results raise important issues about the components and their assessment, and about mentalizing as a multidimensional, profile construct.

4.1 Components of Mentalizing

The first component, named Affective Empathy, increased with several facets of mentalizing which scholars have defined as affective empathy (Carré et al., 2013; Davis, 1983; Lindeman et al., 2016; Muncer & Ling, 2006). These were the following (assessment method in parentheses): emotional contagion (BES-A), emotional reactivity (EQ-Short), emotional connection (BES-A), empathic concern (IRI), and finding images of distressed individuals as emotionally touching (PET). Being the first component, the results imply that individual differences are greater in affective empathy than in any other aspect of mentalizing that were addressed here.

Moreover, three other mentalizing facets loaded moderately on the first component, namely fantasy (IRI), cognitive empathy (BES-A), and understanding social etiquette (SES). The results for the Fantasy and Cognitive Empathy scales replicate earlier findings. Although both of these scales have originally been considered as measures of cognitive empathy, people with strong affective empathy have been shown to score high both on Fantasy (Carré et al., 2013) and on BES-A Cognitive Empathy (Albiero, Matricardi, Speltri, & Toso, 2009; Jolliffe & Farrington, 2006), suggesting that neither of these scales are indicators of cognitive empathy only. As suggested by Jolliffe and Farrington (2006), the overlap between BES-A Cognitive and Affective Empathy is plausible: for example, a high degree of emotional recognition can be expected to facilitate both greater cognitive and affective empathy.

The result about social etiquette is also plausible. Although we initially anticipated that SES would reflect practical mentalizing skills, guilt for breaking social rules, and thus possibly causing other people harm, is no doubt an affective reaction. Whereas previous studies have typically focused only on the cognitive component, i.e., detection of faux pas among children and in clinical samples (e.g., Baron-Cohen et al., 1999; Dolan & Fullam, 2004), the present results highlight the possibility that guilt for breaking social rules is an important aspect of affective empathy among adults. To increase our more general understanding of faux pas, a variety of new methods are needed alongside the SES because the other available tasks are too trivial to be used in a general, adult population (e.g., acknowledging that disparaging accidentally a gift to the gift-giver is a goof).

The second principal component correlated strongly and positively with scores on the EQ-Short Social Skills, and strongly and negatively with scores on the IRI Personal Distress. The component was named Social Skills because the items focus on behaviour in emotionally charged situations (IRI) or in social situations in general (EQ-Short). The component was moderately associated also with high scores on two cognitive empathy scales, EQ-Short and BES-A. This is in line with the findings that the EQ Cognitive Empathy cannot be differentiated from the EQ Social Skills (Kempe & Heffernan, 2011). Because principal components are new variables that account for most of the variance in the observed variables, the results imply that people's view about their social behaviour is a more important aspect of self-reported mentalizing than views about their cognitive empathy. Although it has been acknowledged that social behaviour is based on cognitive empathy (Decety, 2011), scholars in turn have emphasized the division between affective and cognitive empathy in mentalizing, not between affective empathy and social skills.

The third principal component increased with increasing willingness to acknowledge fallibility and with decreasing over-confidence in one's own beliefs. The fourth component was associated only with

epistemological understanding of the nature of beliefs and knowledge. These results indicate that metacognition about one's own thinking is independent of understanding the nature of beliefs in general, and that these two types of metacognition are independent of the two other components of mentalization. Some scholars have suggested that thinking about one's own mental states (Dimaggio & Lysaker, 2015), understanding the nature of beliefs and knowledge (Apperly, 2012; Kuhn, 2000), and metacognitive abilities overall (Smith et al., 2003) may be important in mentalizing. In this study, both metacognition components were pure and independent but they accounted for only a small part of the variance in the data. Accordingly, their role in mentalizing, in the light of results obtained, may not be very important.

4.2 Mentalizing and Understanding Figurative Language

Contrary to our expectations, understanding metaphors and irony were not related to any of the mentalizing components. The method we used for understanding metaphors had low validity and reliability. Better assessment methods are hence needed to examine whether metaphor comprehension reflects practical mentalizing skills as suggested (Apperly, 2012; Brüne & Brüne-Cohrs, 2006) or whether it is mainly a question of language processing as some other authors have presented (a meta-analysis: Bohrn et al., 2012).

In turn, assessment of understanding irony was valid and reliable. The reason why irony did not load on any of the mentalizing components can only be speculated because it is generally agreed that understanding irony requires mentalizing skills (e.g., Bohrn et al., 2012). One potential explanation for the lack of findings is that the stimulus stories were not emotional enough because earlier studies have shown that understanding irony is largely based on emotional processes (Shamay-Tsoory, Tomer, & Aharon-Peretz, 2005). It remains hence for future studies to determine whether and how irony comprehension relates to other facets of mentalizing.

4.3 Mentalizing as a Profile Construct

In the introduction, we raised the question about the nature of the multidimensionality in mentalizing. Within the limits of the methods used in this study, the data support the view that there is no g factor nor any other higher-order factors of mentalizing. Rather, the results strongly suggest that mentalizing is a profile construct, comparable to personality traits. High affective empathy, for example, does not predict anything about the person's cognitive empathy or social skills. Accordingly, a person's overall interests or abilities in mentalizing should be specified as a profile of discrete combinations of different facets of mentalizing.

The finding that there seems to be a set of independent facets of mentalizing that can make different combinations in individuals has important implications. For one thing, assessment of overall mentalizing skills requires attention to several aspects of mentalizing, and conclusions about one facet cannot be made on the basis of research results about another. Similarly, findings concerning one or some mentalizing facets do not justify conclusions about overall mentalizing.

4.4 Limitations

One limitation of this study is the small number of participants. Second, most methods were self-reports. Although self-reports can reliably measure traits, interests, and other typical performance, they are not the best methods to examine abilities, i.e., maximum performance. Third, some of the methods did not work. Besides the aforementioned test for understanding metaphors, Reading the Mind in the Eyes Test could not be included in the analyses because the model did not fit the data. One reason for the poor fit indices may be that the test measures not only mentalizing skills but also visual perception skills, as Valla et al. (2010) have suggested. Moreover, although the other fit indices of IRI were good, CFI ja TLI were only satisfactory, indicating possibly the problems in the four component theoretical model underlying IRI, as noted by Albiero et al. (2009).

Finally, it is important to note that a host of other methods than those examined here have been used to assess mentalizing. However, most of them are appropriate only for clinical purposes or for children. Because valid methods for assessing variations in healthy adults' mentalizing are not easily available, we need new methods to increase our understanding of mentalizing.

4.5 Conclusion

Researchers have appropriately pointed out that the meaning of mentalizing, or theory of mind, is vague and inconsistent, that the methods used are highly heterogeneous, and that studies using one method often inherit assumptions from other research, with little regard for whether these assumptions are appropriate (Apperly, 2012; Schaafsma, Pfaff, Spunt, & Adolphs, 2015). The present study shed light on the construct by identifying four dimensions of mentalizing which should be conceptualized as independent both in theory formation and in empirical work. The results concerning the two first components are in line with the findings from cognitive neuroscience. These results have shown that two types of empathic processes are mediated by two neural

networks: bottom-up processing of affective sharing and emotional matching, and a more advanced system for cognitive understanding of mental states which influences empathic experience and the likelihood of prosocial behavior (Decety, 2011). The way individuals understand their own thinking (i.e., insight) and the views about beliefs and knowledge in general formed two distinct components but understanding figurative language was not related to the other aspects of mentalizing. While preliminary in nature, the results are among the first to point out that mentalizing is a non-hierarchical profile construct and that conclusions about individuals' standing on one mentalizing dimension cannot be made about their standing on another.

Acknowledgements

This study was supported by the Research Funds of the Academy of Finland (265518).

References

Albiero, P., Matricardi, G., Speltri, D., & Toso, D. (2009). The assessment of empathy in adolescence: A contribution to the Italian validation of the "Basic Empathy Scale". *Journal of Adolescence*, *32*, 393-408. https://dx.doi.org/10.1016/j.adolescence.2008.01.001

Apperly, I. A. (2012). What is "theory of mind"? Concepts, cognitive processes and individual differences. *The Quarterly Journal of Experimental Psychology*, *65*, 825-839. https://dx.doi.org/10.1080/17470218.2012.676055

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. https://dx.doi.org/10.1023/B:JADD.0000022607.19833.00

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*, 407-418. https://dx.doi.org/10.1023/A:1023035012436

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. https://dx.doi.org/10.1111/1469-7610.00715

Beck, A. T., Baruch, E., Balter, J. M., Steer, R. A., & Warman, D. M. (2004). A new instrument for measuring insight: The Beck Cognitive Insight Scale. *Schizophrenia Research*, *68*, 319-329. https://dx.doi.org/10.1016/S0920-9964(03)00189-0

Bohrn, I. C., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language—A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia*, *50*, 2669-2683. https://dx.doi.org/10.1016/j.neuropsychologia.2012.07.021

Brüne, M., & Brüne-Cohrs, U. (2006). Theory of mind—Evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience & Biobehavioral Reviews*, *30*, 437-455. https://dx.doi.org/10.1016/j.neubiorev.2005.08.001

Carré, A., Stefaniak, N., D'Ambrosio, F., Bensalah, L., & Besche-Richard, C. (2013). The Basic Empathy Scale in Adults (BES-A): Factor structure of a revised form. *Psychological Assessment*, *25*, 679-691. https://dx.doi.org/10.1037/a0032297

Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*, 1-29. https://dx.doi.org/10.18637/jss.v048.i06

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113-125. https://dx.doi.org/10.1037/0022-3514.44.1.113

Davis, M. H. (2006). Empathy. In J. Stets, & J. H. Turner (Eds.), *Handbook of the Sociology of Emotions* (pp. 443-466). New York: Springer. https://dx.doi.org/10.1007/978-0-387-30715-2_20

Decety, J. (2011). Dissecting the neural mechanisms mediating empathy. *Emotion Review*, *3*, 92-108. https://dx.doi.org/10.1177/1754073910374662

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 71-100. https://dx.doi.org/10.1177/1534582304267187

Dimaggio, G., & Lysaker, P. H. (2015). Metacognition and mentalizing in the psychotherapy of patients with psychosis and personality disorders. *Journal of Clinical Psychology*, *71*, 117-124. https://dx.doi.org/10.1002/jclp.22147

Dimitrov, D. M. (2003). Marginal true-score measures and reliability for binary items as a function of their IRT parameters. *Applied Psychological Measurement*, *27*, 440-458. https://dx.doi.org/10.1177/0146621603258786

Dolan, M., & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological Medicine*, *34*, 1093-1102. https://dx.doi.org/10.1017/S0033291704002028

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Happé, F. G. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, *48*, 101-119. https://dx.doi.org/10.1016/0010-0277(93)90026-R

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*, 191-205. https://dx.doi.org/10.1177/1094428104263675

Hooker, C. I., Verosky, S. C., Germine, L. T., Knight, R. T., & D'Esposito, M. (2008). Mentalizing about emotion and its relationship to empathy. *Social Cognitive and Affective Neuroscience*, *3*, 204-217. https://dx.doi.org/10.1093/scan/nsn019

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

Kempe, V., & Heffernan, E. (2011). Digit ratio is linked to affective empathy in women. *Personality and Individual Differences*, *50*, 430-433. https://dx.doi.org/10.1016/j.paid.2010.10.024

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*, 178-181. https://dx.doi.org/10.1111/1467-8721.00088

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, *15*, 309-328. https://dx.doi.org/10.1016/S0885-2014(00)00030-7

Langdon, R., & Coltheart, M. (2004). Recognition of metaphor and irony in young adults: The impact of schizotypal personality traits. *Psychiatry Research*, *125*, 9-20. https://dx.doi.org/10.1016/j.psychres.2003.10.005

Law, K. S., Wong, C. S., & Mobley, W. H. (1998). Toward a taxonomy of multidimensional constructs. *Academy of Management Review*, *23*, 741-755.

Lindeman, M., Koirikivi, I., & Lipsanen, J. (2016). Pictorial Empathy Test. An easy-to use method of affective empathic abilities. *European Journal of Psychological Assessment*. https://dx.doi.org/10.1027/1015-5759/a000353

Maronna, R. A., & Yohai, V. J. (1995). The behavior of the Stahel-Donoho robust multivariate estimator. *Journal of the American Statistical Association*, *90*, 330-341. https://dx.doi.org/10.1080/01621459.1995.10476517

Muncer, S. J., & Ling, J. (2006). Psychometric analysis of the Empathy Quotient (EQ) scale. *Personality and Individual Differences*, *40*, 1111-1119. https://dx.doi.org/10.1016/j.paid.2005.09.020

Paunonen, S. V., Rothstein, M. G., & Jackson, D. N. (1999). Narrow reasoning about the use of broad personality measures for personnel selection. *Journal of Organizational Behavior*, *20*, 389-405. https://dx.doi.org/10.1002/(SICI)1099-1379(199905)20:3%3C389::AID-JOB917%3E3.0.CO;2-G

Piaget, J. (1929/1951). *The child's conception of the world*. London: Routledge & Kegan.

R Core Team. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1-36. https://dx.doi.org/10.18637/jss.v048.i02

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, *19*, 65-72. https://dx.doi.org/10.1016/j.tics.2014.11.007

Shamay-Tsoory, S., Tomer, R., & Aharon-Peretz, J. (2005). The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology*, *19*, 288-300. https://dx.doi.org/10.1037/0894-4105.19.3.288

Shamay-Tsoory, S., Tomer, R., Goldsher, D., Berger, B., & Aharon-Peretz, J. (2004). Impairment in cognitive and affective empathy in patients with brain lesions: Anatomical and cognitive correlates. *Journal of Clinical and Experimental Neuropsychology*, *26*, 1113-1127. https://dx.doi.org/10.1080/13803390490515531

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences*, *26*, 317-339. https://dx.doi.org/10.1017/S0140525X03000086

Valla, J. M., Ganzel, B. L., Yoder, K. J., Chen, G. M., Lyman, L. T., Sidari, A. P., ... Wong, S. K. (2010). More than maths and mindreading: Sex differences in empathizing/systemizing covariance. *Autism Research*, *3*, 174-184. https://dx.doi.org/10.1002/aur.143

Wang, J., Zamar, R., Marazzi, A., Yohai, V., Salibian-Barrera, M., Maronna, R., ... Maechler, M. (2013). *Package "robust" (pdf)*. CRAN.

Appendix

Table S1. Possible dimensions of mentalizing and their assessment in the study

Dimension	Assessment method
EMPATHY	
Cognitive empathy	
Overall	Basic Empathy Scale in Adults, Empathy Quotient Short, Reading Mind in the Eyes Test
Fantasy	Interpersonal Reactivity Index
Perspective taking	Interpersonal Reactivity Index
Affective empathy	
Overall	Pictorial Empathy Test
Emotional reactivity	Empathy Quotient Short
Contagion	Basic Empathy Scale in Adults
Disconnection	Basic Empathy Scale in Adults
Empathic concern	Interpersonal Reactivity Index
Personal distress	Interpersonal Reactivity Index
PRACTICAL SKILLS	
Social skills	Empathy Quotient Short
Faux pas	Social Etiquette Scale
Understanding metaphors	Modified Story Comprehension Task
Understanding irony	Modified Story Comprehension Task
METACOGNITION	
Views about beliefs	Epistemic Understanding Questionnaire
Self-reflectiveness	Beck Cognitive Insight Scale
Self-certainty	Beck Cognitive Insight Scale

Complete description of the Epistemic Understanding Questionnaire (EUQ)

Participants' views about beliefs in general was assessed with the EUQ (Kuhn, Cheney, & Weinstock, 2000). The scale is based on the rationale that individuals at the absolutist level see knowledge as located in the external world and knowable with certainty. In turn, the multiplist relocates the source of knowledge from the known object to the knowing subject, being thus aware of the subjective nature of believing and knowing. Individuals at the highest, i.e., evaluativist level acknowledge that two people can both have legitimate positions but one position can have more merit than the other to the extent that that position is better supported by argument and evidence.

In the questionnaire, two contrasting claims, attributed to two individuals, Robin and Chris, are presented. The claims include five judgment domains: personal taste domain, aesthetic judgments, value judgments, social truth judgments, and physical truth judgments. An example from the value judgments domain is: "Robin thinks people should take responsibility for themselves. Chris thinks people should work together to take care of each other." After each pair of claims, the participants are asked: "Can only one of their views be right, or could both have some rightness?" The response options are "Only one can be right" (reflecting the absolutist level) or "Both could have some rightness". Participants who choose the "both could have some rightness" option are asked the second question: "Could one view be better or more right than the other?" The response options are "One could not be more right than the other" (reflecting the multiplist level) and "One could be more right" (reflecting the evaluativist level).

The viability of EUQ for the present study was assessed with the pilot study. Four of the five EUQ domains were included. The personal taste domain was omitted because (Kuhn et al., 2000) found that the highest expected level in this domain was attained even by children. Twelve pairs of contrasting claims, three for each of the four domains, were thus presented to the participants.

The results showed that a vast majority of the participants (72.7%-93.9%) agreed that one could not be more right than the other in value judgments, social truth judgments, and physical truth judgments. We therefore excluded the first question and replaced it with the statement that "both may have some rightness" and proceeded directly to the second question in the main study. Because one cannot legitimately regard one art work as objectively superior to another (cf. Kuhn et al., 2000), we also excluded the aesthetic judgments domain but included one claim pair from this domain as a filler item. Besides the filler claims, we thus used nine pairs of contrasting claims, three for the above three domains, in the main study. Responses reflecting the multiplist level were scored as 1, and responses reflecting the evaluativist level were scored as 2.

Complete Description of the Tasks Assessing Understanding Metaphors and Irony

Understanding metaphors and irony were assessed with a modified lexical decision task. The participants were presented 15 stories, adjusted and modified to Finnish language from the Story Comprehension Task (Langdon & Coltheart, 2004). The stories ended either with a metaphorical (six stories), an ironic (six stories), or a literal statement (three filler stories) made by one of the story characters. Table S2 shows examples of the stories with different endings. After each story, we presented two stimulus words. The participants were required to rapidly assess how well the words corresponded to what the story character wanted to say (1 = not related, 2 = partly related, 3 = related). One of the two stimulus words corresponded either to the figurative interpretation or to the literal interpretation of the statement. The other stimulus word was a filler word which was either related or unrelated with the statement (see Table S2). Because we had a 3-point scale and six stories for both metaphors and ironies, the variables for understanding metaphors and understanding irony could both range from 1 to 18.

The feasibility of this task for the present study was assessed with the pilot study. In the pilot study we told the participants which of the stories ended with a metaphorical statement and which stories ended with an ironic statement. No time limit was set for the task. The instruction turned out to be slightly ambiguous but the responses to the stimulus words were meaningful. We therefore clarified the instruction but made no changes on the stories or the stimulus words in the main study.

The procedure in the main study was as follows. The participants were given the instruction: "In this task, you will see short stories ending in one of the story characters making a statement. Your task is to assess how the two words which we present to you correspond with what the character wants to say. You have only 10 seconds time to respond before the next story is presented". The words "what the character wants to say" were written in bold. The participants were then shown an example story ending with a bold written non-figurative statement and two stimulus words with the 3-point rating scales on the screen.

Next, two practice stories with no time limit were presented, one with a non-figurative statement and one with a metaphorical statement. The participants could read the story as long as they wanted and they were asked to move to the questions by pressing the "Continue" button when they were ready. Two common Finnish words, each with 4-8 letters, then appeared on the screen together with the question "How do these words correspond with what the character wanted to say?" The story with the highlighted statement remained on the screen. The 3-point rating scale was provided for both words. After filling in the practice tasks, the participants were told that in the next questions, they had only 10 seconds time to respond, and they were asked to move on when they were ready. The 15 stories were ordered randomly and presented in a different order for each participant. The order of the stimulus words was counterbalanced.

Table S2. Five examples of stories ending with three types of statements and four types of stimulus words rated on a 3-point scale (1 = not related, 2 = partly related, 3 = related)

	Intended meaning	Literal meaning	Filler words (related to the story)	Filler words (unrelated to the story)
Metaphorical statements:				
1. Anna wants to know more about her new roommate Pauli. Curious, Anna asks Pauli questions of his life. Anna wants to know if Pauli has a girlfriend and what he wants to do in the future. Pauli answers to Anna's questions and remarks: **"I am an open book."**	honest[a]			landscape
2. The mathematics teacher is explaining to his students how to solve a difficult math problem. He tells the students to listen carefully, because they have not studied this problem before. After writing out examples on the board, he notices the confused looks on students' faces. He says: **"You seem to be in a fog."**		mist[b]	teaching	
Ironical statements:				
3. Tomi has been working hard for the promotion he wants. He is very disappointed, since his boss gave the position Tomi wanted to another employee. Tomi goes to tell his friend about the boss' decision and says: **"I've just had some great news."**	resentful[a]			sun
4. Juho comes home from school. He has a lot of homework to do. Juho goes to his room and immediately starts doing his homework. Two hours later, Juho's mother peeks into the room and finds that Juho is still working on his homework. His mother says: **"You only like to loaf around."**	diligent[a]	inefficient[b]		
Control statements:				
5. Jani goes over to Kaisa's house. He wants to see Kaisa's new dog. When Jani arrives at Kaisa's place, the dog greets him by jumping up and licking Jani's face. Jani laughs loudly and says: **"Your dog is very enthusiastic."**			happy	shop

Note. [a] = correct (related), [b] = incorrect (not related).

Complete Description of the Development of the Social Etiquette Scale (SES)

Because the available faux pas tests are appropriate mainly for children or for clinical samples, we developed a scale to assess etiquette skills in the general population. We surfed the Internet and gathered the most commonly listed etiquette mistakes to develop the Social Etiquette Scale (SES). We tested a scale with 14 statements in the

pilot study. Three statements lowered the reliability and the SES scale with 11 items was used in the main study. The participants were told that we will ask them "about bad habits which we all have at least occasionally". The statements were presented twice, first with a question "How often do you behave like this?" (1 = never, or almost never, 5 = very often), and then with a question "If this were to happen would you feel you were doing wrong?" (1 = no, 5 = surely). Only the latter, understanding social etiquette was included in the analyses because it had higher reliability (α = .86) than the variable addressing behaviour (α = .65) and because understanding wrongdoing may reflect mentalizing skills better than behaviour (resulting from poor impulse control, for example).

Final items:

1) I am late for appointments

2) I am not listening to other people

3) If a written invitation reads "please respond" (RSVP.), I do not reply

4) I use my mobile phone a lot even when I am with other people

5) I cut in line

6) I borrow things without permission, for example food out of a shared refrigerator

7) I do not thank although I probably should

8) I don't introduce two unknown people to each other, although I know them both

9) I do not give up my seat for elderly or disabled on public transport

10) I treat waiters (or other customer servants) impolitely

11) I forget the norm of reciprocity, for example in terms of a gift or service

Sources

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, *15*, 309-328. https://dx.doi.org/10.1016/S0885-2014(00)00030-7

Langdon, R., & Coltheart, M. (2004). Recognition of metaphor and irony in young adults: The impact of schizotypal personality traits. *Psychiatry Research*, *125*, 9-20. https://dx.doi.org/10.1016/j.psychres.2003.10.005

Copyrights