Università degli Studi di Torino

Dipartimento di Psicologia

Dottorato di ricerca in
Scienze Psicologiche, Antropologiche e dell'Educazione
XXX ciclo

Development and Initial Validation of an R-PAS Empathy Index

Candidata   Tutor
Marzia Di Girolamo   Luciano Giromini
SSD M-PSI/07

A.A. 2014-2017

INDEX

I. EMPATHY

Without some empathetic skills, human beings would be seriously limited in building interpersonal relationships, because they would be blind to others' needs and desires. Nonetheless, is empathy a real construct or is it just a term used in the common language to identify behaviors and phenomena that not necessarily match with a psychological function? Danzinger (1997; Jahoda, 2005) argued that the categories used in psychology are theoretical definitions, not actual things themselves. Often, the efforts to define psychological constructs lead to a deep analytical process with the aim of identifying the main aspects of that concept, and finally synthetize them in a basic, endorsable definition. In the case of empathy, as noticed by Davis (1994), it is a multifaceted phenomenon of interest to many fields like psychology, anthropology, philosophy, and ethology; thus, «one reason it is difficult to get a good handle on empathy is that it has so many handles (p. IX) ».

1.1. WHAT IS (NOT) EMPATHY

Empathy is such a complex construct that the more we explore it, in its various facets and components, the more we get the feeling of losing its central core. More in general, the more it is analyzed, the less it can be synthesized. There is no agreement within the scientific community regarding the definition of empathy (Leiberg & Anders, 2006).

Although[1] scholars have attributed different meanings to the word "empathy" over time, all agree that this psychological construct plays a key role in human interactions. Empathy was first defined as "to feel into" (Lipps, 1903) and referred to how people come to know others' emotional states. Since then, many refinements of this complex and multifaceted construct have been

---

[1] The most part of this paragraph is present in Di Girolamo, M., Giromini, L., Winters, C. L., Serie, C. M., & de Ruiter, C. (2017). The Questionnaire of Cognitive and Affective Empathy: A Comparison Between Paper-and-Pencil Versus Online Formats in Italian Samples. Journal of personality assessment, 1-12.

proposed (Leiberg & Anders, 2006). The importance of empathy is self-evident; every time we try to understand others' behaviors or intentions, and we consequently adapt our social behavior to improve our interactions and relationships, we are empathizing.

Different definitions of empathy have been proposed in various contexts such as psychotherapy, social psychology, neuropsychology, and even ethology.

Rogers (1959) proposed that empathy would be the capability to perceive the internal frame of reference of someone else with the same emotional components and meanings. This definition is probably incomplete, when compared to more recent conceptualizations, as it only focuses on the voluntary act of "put[ting yourself] in someone else's shoes," without taking into account the emotion-contagion process, which is mainly an automatic and spontaneous one. Hoffman (1984) – on the contrary – focused more on the emotional aspects of empathy, and defined it as an "affective response more appropriate to someone else's situation than on one's own" (Hoffman, 1984, p.114).

In the last twenty years (Davis, 1994) the scientific community not only realized that empathy is characterized by an affective component (which determines the quality of the experience) and a cognitive one (which, places it in a frame of meaning), but also that these two components do not always coexist. Actually, Davis argued that in defining empathy «there is one central recurrent, and seemingly intractable problem: the term empathy is routinely used to refer to two distinctly separate phenomena, cognitive role taking and affective reactivity to others. […] a fact which contributes in on small way to the continuing semantic confusion in empathy research (Davis, 1994, p.9)».

1.1.1.   COGNITIVE EMPATHY

Cognitive empathy is the ability to understand how other people might feel, using visual, auditory, and/or situational cues. Within the research literature (e.g., Blair, 2005), cognitive empathy is sometimes considered to be a construct very close to that of Theory of Mind (ToM; Lawrence, 2004; see also Kanske,

Böcckler, Trautwein, & Singer, 2015; Brown, Thibodeau, Pierucci, & Gilpin, 2017), which is the capability to understand that we and others have mental states, that mental states of others can differ from one's own, and that others' behavior can be explained by their mental state (Frith & Frith, 2003). Some authors (Premack and Woodruff, 1978; Frith and Frith, 2003) argued that cognitive empathy is based on a high cognitive level because it is the capability to infer others' beliefs, intentions, desires, or emotions using the mentalizing capability. Some other authors (Gordon, 1986; Goldman, 2005, 2006; Gallese, et al., 2004) consider the Cognitive empathy as a process more automatic – although with different degrees – and spontaneous, based on the idea that we understand others' mental states because we reproducing or simulating them in our own minds.

Until 3, a child does not understand that other people cannot know all the things he/she know and that others may hold false beliefs. Between 3 to 4, children learn that other people could have false beliefs rather the actual state of the world. Between 6-7 years, children understand that also other people can understand the mental states of others. The highest level of this process is reached between the ages of 9-11, when children become able to recognize the so-called *faux pas*: «to understand that a faux pas has occurred, one has to represent two mental states: that the person saying it does not know that they should not say it and that the person hearing it would feel insulted or hurt (Stone, Baron-Cohen, Knight, 1998; p. 641) ». The last aspect, the one related to hurt others, involved also the affective component of empathy.

An empathic behavior can be implemented without a particular emotion being involved. Typical examples of this are those behaviors called "friendly" or "kind": if a person with crutches gets on the bus while we are seated, and there are no free seats, we could get up to let him/her sit. At that moment, our mood may not have changed, but the other-directed behavior is however put in place.

Moreover, in some circumstances it is hard to actually put us in someone else's shoes – e.g., all those times that we are exposed to war images or natural disasters – however we can still feel sad. Conversely, circumstances could make us ignore information that comes from cognitive empathy. Let us imagine that while we were facing a very important test, we saw someone worried and in

difficulty. We could perfectly appreciate the mental state of that person, but at that moment our benefit is more important than his one, and we do not put in place an other-oriented behavior. Furthermore, in some cases – suffice it to say about manipulative or Machiavellian people – cognitive empathy could be used to exploit others.

1.1.2. AFFECTIVE EMPATHY

Everyone has had the experience of being sad, or upset, in the presence of a child, and having been comforted by him with the object that he usually uses to console himself (e.g., toys, plush toys, Teddy bears). The child had no idea about the reason why you were sad, or about what makes you feel better, but he wanted you to be well. That is how an empathetic behavior can be put in place without understanding the origin or the solution of the issue.

Feeling and or understanding does not mean "reacting," and reacting does not mean "reacting appropriately". Baron-Cohen (2004) classified three types of affective arousal due to the exposure to others' experiences. The most intuitive case is when the observer's type of emotion matches exactly the one of the observed person (Eisenberg & Miller, 1987). The second case is when the two emotions are not properly overlapping, and this case could be differentiated within two further categories: the observer's feeling is appropriate to the others' experience, or it is not. Suppose we were walking down the street and we saw a group of bullies beating a little boy. The boy is probably feeling pain and fear; if we felt the same identical emotion, namely fear, we would not be able to help him, because we would be scared too. In order to help him, it is necessary to not feel fear, maybe anger, but surely not fear. In this case, the emotion is not the same one but it is appropriate to manage the situation in some way (Stotland, 1969), and the behavior. By contrast, imagine that sadness of someone else makes us happy, or at least satisfied. In this case, the emotion is different and also unappropriated. This last case was named "contrast empathy" (Stotland,

Sherman, & Shaver, 1971) or, more recently, Schadenfreude (i.e. "gloating"; Feather & Nairn, 2005; Cikara, & Fiske, 2013). However, it remains to define how much two emotions could be really considered "matching," and then what can be classified as an appropriate reaction. If a friend of ours is sad because s/he broke up with his/her obnoxious and irritating partner, we could reasonably feel contrasting emotions.

However, in a different level, those examples can furthermore split into two additional, main types of empathic affective responses. One is the personal arousal, which could be in form of distress or excitement, and is self-oriented; another is the reaction, which is other-oriented (Gibbons, 2011). However, as far as could the arousal be spontaneous, the reaction should receive modulation from the cognitive functions, at least in non-pathological conditions, because attention can be preemptively allocated when an automatic response is undesirable (Preston and de Waal, 2002).

1.1.3.   EMPATHIC BEHAVIOR

From this short overview, empathy appears as anything and everything at the same time: «Cognitive versus affective, attitude versus behavior, a momentary experience versus a life situation, shallow versus deep, and expressed versus unexpressed (Bohart & Greenberg, 1997, p. 444) ». In an attempt to integrate these seemingly contradictory definitions, Bateson (2009) more recently hypothesized that a good way to conceptualize empathy would be to consider it as 'the answer' to the following two questions. First, how can a human being know what someone else is thinking or feeling? Second, what leads this person to react with sensitivity in front of his or her suffering?

As regard the first question, Preston and de Waal (2002) introduced the so-called perception-action model of empathy. The core of this theory is that the empathetic process is automatically triggered by the view of the emotional state of another person. This process starts with a corresponding representation of

that emotional state, based on a somatic and a motor reaction in the observer, and then it could conduct to an empathetic behavior. This model is supported by the somatic marker hypothesis (Damasio, 1994) and the discovery of mirror neuron system (MNS; Rizzolatti, 1996), given that it is based on the idea that the experience of the observer overlaps with the experience of the observed person. However, everyday life shows that we can vicariously experience emotional states of others, or deduce them, even if we have never experienced similar situations before (e.g., war-related emotions, earthquakes, etc.). In addition, sometimes it is sufficient to hear, to read or even to imagine about others' experiences to provoke an empathetic feeling. Thus, a perception-action model is necessary but not sufficient to explain the entire empathy process.

To overcome these limitations, Baron-Cohen et al. (2005) have proposed a model of empathy that better accounts the fact that we can be empathetic without an emotional contagion-like process. More in detail, they identified two levels of the empathy processes: the lower one, which develops early, is the affective part of empathy and includes the contagion-like process; the higher one, which develops later, is the cognitive part of empathy and includes complex cognitive processes like the Theory of Mind (Leiberg & Anders, 2006).

With regard to the second question posed by Bateson (i.e., what leads a person to react with sensitivity in front of suffering), Preston and de Waal (2002) suggested that empathy is a prosocial behavior based on the cost/benefit in peer and kin groups. In this model, support, assistance, and help become advantageous because they are likely to be reciprocated by other members of the group (see, for example, Trivers, 1971). Another possible account for the humans' predisposition toward being sensitive or empathetic to others' suffering could be found in the emotional contagion process (Weisbuch, Ambady, Slepian, & Jimerson, 2011). In this view, the exposure to others' pain, that would automatically and intrinsically elicit some distress in the observer, and therefore the observer could choose to act empathetically simply because s/he wants to discontinue his or her own distress.

Given a perception-action view of empathy, these processes extend to the prediction and response to allies as well as competitors. The PAM can produce appropriate helping behaviors, as well as effective punishments. In both cases, the subject accesses the object's state and generates an appropriate response. Associated representations of the object and situation will determine whether the desired out- come is to produce or alleviate distress (Preston & de Waal, 2002).

The effort to stop other's pain is a behavior influenced by the balance between the stress that the other's pain provokes in the subject and the potential stress that the subject would prove it he/r will step in (bullying, hurry, relational reasons, deal of effort, etc.).

Summarizing, Hatfield, Rapson, and Le (2011) claimed that true empathy requires 3 skills: ability to share the other person's feelings, cognitive ability to intuit what another person is feeling, "socially beneficial" intention to respond compassionately to that person's distress.

Most intuitively, it could be said that empathy is a behavior implemented when others' experience is shared, both cognitively – I go into the others – and affectively – the other comes into me.

1.2. EMPATHY IN CLINICAL PSYCHOLOGY

One of the main changes in the DSM 5 is the Alternative Dimensional Model for the assessment of personality disorders. The rationale under this improvement is that, quite apart the specific pathological traits, the cross-sectional difficulties in personality disorders is impairment in ideas and feelings about self and interpersonal relationships. Personality functions are assessed by focusing on two main areas, self and interpersonal, in turn split into identity/self-direction and intimacy/empathy. Empathy is defined as the comprehension and appreciation of others' experiences and motivation; tolerance of differing perspectives; understanding the effect of own behavior on others (p. 821).

Nonetheless, this modification in personality assessment procedure is notable and helpful, even before the Fifth edition of the DSM empathy was a crucial criterion in some pathologies, namely the Narcissistic Personality Disorder (NPD) and the Antisocial Personality Disorder (AsPD).

The NPD is overall defined as «a pervasive pattern of grandiosity (in fantasy or behavior), need for admiration, and lack of empathy, beginning by early adulthood and present in a variety of contexts [...]». More specifically, the 6th and the 7th diagnostic criteria say "Is interpersonally exploitative (i.e. takes advantage of others to achieve his or her own ends)" and "Lacks empathy: is unwilling to recognize or identify with the feelings and needs of others" (American Psychiatric Association, 2013, p. 670-671)

In the AsPD, the lack of empathy is not mentioned as a basic characteristic, but the 2nd and the 5th criteria identify behaviors strongly in contrast with an adequately developed empathy capabilities: "Deceitfulness, as indicated by repeated lying, use of aliases, or conning others for personal profit or pleasure" and "Reckless disregard for safety of self and others" (American Psychiatric Association, 2013, p. 659). However, even if it is not mentioned in the DSM-5 description of these diseases, Narcissistic and Antisocial patients are also quite able to manipulate and influence people who have to do with them (Dolan & Fullam, 2004; Ritter, Lorig, Laurent, & Matthews, 2004; Stellwagen & Kerig, 2013), so their empathy deficits are more related to the emotionally evocative stimuli than with the capability to put their selves in other's shoes (Marcoux, Michon, Lemelin, Voisin, Vachon-Presseau et al., 2014).

Constructs close to the empathy, as the Theory of Mind and the capability to recognize facial expressions, in NPD and AsPD were actual investigated. In some of these studies, no significant differences were found when comparing control and clinical samples (Richell, Mitchell, Newman, Leonard, Baron-Cohen, et al., 2003; Dolan & Fullam, 2004; Glass & Newman, 2006), whereas some others research has found that patients with ASPD can have difficulties in ToM tasks and in facial expressions recognition tasks, especially with fear, sadness or anger (Blair, 2005; Munro, Dywan, Harris, McKee, Unsal, & Segalowitz, 2007; Hastings,

Tangney, & Stuewig, 2008; Wolfkühler, Majorek, Tas, Küper, Saimed, Juckel, & Brüne, 2012).

In addition, lack of empathy is not just important for personality assessment, but also in psychotic disease and in Autistic Spectrum Disorder. Given the deficits in the interpersonal skills in psychotic diseases (Ofir-Eyal, Hasson-Ohayon & Kravetz, 2014), studies were conducted in schizophrenic patients about their metacognitive abilities. Results are just partially coherent: some studies show that psychotic subjects (Weiss, Kohler, Nolan, Czobor, Volavka et al., 2006), or violent psychotic subjects (Abu-Akel and Abushua'leh, 2004; Majorek et al., 2009), have difficulties in recognizing other people's emotions or in ToM tasks; in some other researches these results are not statistically significant (Mitchell., Gumley, Reilly, Macbeth, Lysaker et al., 2012) or even overturned, with better performances being observed within the violent psychotic samples compared to the control ones (Silver, Goodman, Knoll, Isakov, & Modai, 2005).

Autistic Spectrum Disorders (ASD; APA, 2013) is a big field of study regard to social competences. The ASD, indeed, are characterized by difficulties with the social environment, in terms of interaction, comprehension and therefore prediction (Baron-Cohen, Leslie & Frith, 1985).

However, there is a big difference in the cases of NPD and the ASPD in comparison with behavior where the subject "was not empathic": one thing is to not step in, something else is to purposely provoke other's suffering, and quite something else is to feel pleasure in doing this. These studies are not without limitations (Harris & Picchioni, 2013) but underline the importance to shed light on the empathy and on its components, also in order to deepen the knowledge of the psychopathologies.

It emerges from the literature the importance to properly assess the empathy, at least in its two main aspects, namely the cognitive and the affective component.

II. RORSCHACH INKBLOT TEST

The Rorschach test was published in 1922 (Rorschach, 1922) by its author Hermann Rorschach. Ten inkblot cards compose the tool: 5 in black and white, 2 in black, white and red, and 3 colorful.

From its first publication, the test had to travel through a curious and windy path made of conflicting approaches and theories about how (and if) it works, different administration methods, criticism and skepticism (Meyer & Eblin, 2012; Mihura, Meyer, Dumitrascu, & Bombel, 2013).

2.1. HERMAN RORSCHACH

The birth and development of the Rorschach inkblot test are curiously related to a series of lucky circumstances, that have contributed significantly to determine what nowadays is one of the most important and well-known diagnostic instruments.

The psychiatrist Herman Rorschach was born in Zurich 1884, and in 1903 decided to enroll at the Faculty of Medicine in Zurich, where obtained his Doctolar degree in 1912 under the supervision of Eugene Bleuler.

The main reason why he decided to use the inkblots with patients is not very clear and no more available, but there are some clues about his life that could help formulating some hypotheses.

First, thanks to his father who was a painter, Rorschach grew up in an intellectual, artistic and cultural environment. Second, in this atmosphere, *Klecksographie* played a central role: it is an artist technique that consists in to let the ink pour slowly on a sheet, fold the sheet in half along the central axis, so scatter the inkblot surrounding on the surface in a way that it forms shapes and silhouettes. When Rorschach was young, this technique was very common in Switzerland and it was used like a game or an entertainment in which people had

to imagine what the inkblot looks like. Third, in 1912, Rorschach tested with the inkblot the pupils of his friend Konrad Gehring and purposed that the more gifted pupils showed more phantasy in the task. Fourth, he was interested in the relationship between hallucination and symbolism, which indeed became the subject of his doctoral dissertation. Fifth, his thesis supervisor was Eugene Bleuler, who in 1911 published his "*Dementia Praecox oder Gruppe der Schizophrenien*", a work where functional psychoses were differentiated from the organic ones. Sixth, Rorschach was interested in all artistic productions by psychotics, so that when he was in Waldau Hospital he made a lot of efforts to let the patients paint. Seventh, in 1917 the Bleuler's disciple Szmon Hense published an experiment with the inkblots he had conducted in 1912 (Ellenberger, 1989).

In 1919, Rorschach decided to conduct his last, but most important experiment. The elaboration of the cards, the experimentations with patients and control sample, the writing of the book and all the issues related to its publication, which took more than three years of time. Finally, in 1921, *Psychodiagnostic* was published. Six months later, April 2 1922, Herman Rorschach entered in the hospital where he worked, but this time as an emergency patient, and died because of a peritonitis.

The Rorschach method aroused interest all around the world and in many different contexts. In this work, however, I will focus on the development of the Rorschach-Performance Assessment System™ – given that it is the method used in this research, and the "last stop" of the journey of the Rorschach's test.

2.2. THE "FIVE DIFFERENT RORSCHACH TESTS".

In the '20s of the past century, the American psychiatrist David Levy went in Switzerland to meet Emil Oberholzer – disciple of Bleuler and collaborator of Herman Rorschach – and thanks to him discovered the Herman Rorschach's research. When he came back to USA, in 1926, he wrote some papers on the

Rorschach method that raised the attention of his colleague of the Guidance Institute of New York: Samuel J. Beck. Thanks to a research in developmental psychology with the Rorschach plates, the two colleagues determined the first standardization of the test. Another colleague of theirs, Marguerite Hertz, became attracted to the utility of the test and used it for researches in the same field. The similar results, made Levy, Hertz, and Beck realize the tool's potential.

In 1933, Bruno Klopfer – German psychiatrist fugitive from the Nazi Germany – moved to Zurich where, thanks to Carl G. Jung obtained a job in the C. G. Jung Institute. This experience allowed him to make contact with the plates, with the experiment of Rorschach, and to learn how to administer and code the test. When the next year he came back to the USA, his competence with the tool found the interest of American psychologists. Until then, the knowledge of the Rorschach test was hampered by the absence of a translation in English language (the first one was achieved in 1942). Therefore, Klopfer organized some seminaries for postgraduate students at the Columbia University, with the aim to teach administration and coding of the test. From those seminaries, a lot of coding issues emerged, especially with regards to the localization and shade coding. This inducted Klopfer to create several new codes. In 1936, Klopfer founded *The Rorschach Research Exchange[2],* with the aim to divulge the test and its developments: he tried to involve in this project also Beck, Levy, Oberholzen and Hertz but none of them accepted. On the contrary, this work was the cause of the chasm between Klopfer and Beck: the latter was more close to the experimental psychology of the beginning of the century, the former – clinician and therapist – tried to improve the method by focusing on clinical observations. Hertz, who published in the *Exchange* a paper about strengths and weaknesses of the two systems, initially assumed a median position but her studies brought her to assume a third point of view. The result was a tripartite approach to the Rorschach method.

---

[2] Which will become the *Journal of Projective Techniques* and then the *Journal of Personality Assessment.*

Among the disciples that attended the Klopfer classes, there was the Polish psychologist Zygmunt Piotrowsky. He was interested in studying how patients with neurological diseases could solve ambiguous stimuli, as the Rorschach inkblots were. In 1957, he published *Perceptanalysis*, a book where he integrated his knowledge about the perceptive interpretation with the Rorschach method[3]. His approach, however, did not achieve the same success as the previous ones. On the one hand, it was affected by the reputation of the other two methods and, on the other, it was a particularly eclectic approach with references taken from both neurology and literature.

Last, but not least, the Second World War brought in America also the European ego-psychologist David Rapaport, from Hungary. He worked for the *Menninger Foundation,* which commissioned him the investigation of the test in the clinical field. Thus, he approached the study of the test from a psychoanalytic perspective, and gave a notably boost to the usage of the Rorschach as a personality assessment test. Rapaport, together with Morton Gill and Roy Shafer, conducted a study on the personality assessment based on a battery of tests among which he included the Rorschach. In 1945, the results of this research were published in a work named *Diagnostic Psychological Testing.* Shafer subsequently expanded the psychoanalytic approach used by Rapaport, and exposed it in 1955 in the *Psychoanalitic Interpretation in Rorschach Testing*, which became the reference point for the comprehension of the defense mechanisms, the Ego processes, the conflicts, the reality test, and the psychopathological contents in the Rorschach test.

 As a result, these methods were so different by theories, procedures, administrations, codes and interpretations that Exner, in his work of the 1969 *The Rorschach Systems* (Exner, 1969) will define them "five different Rorschach tests".

---

[3] Among the contributions of this author there is the discrimination between the interpretation of the three movements (human, animal and inanimate) and the *interdependence principle*, namely that the meaning of Rorschach variables has sense just if interpreted in relation with the other variables.

2.3. THE COMPREHENSIVE SYSTEM BY J. E. EXNER

The *Comprehensive System* (CS; Exner, 1974) was an approach that initially was born as an effort to unify the five methods, but then became itself the reference system for the Rorschach test.

Despite the Rorschach method has had an initial steep rise, between the '50s and '60s, its development underwent a critical stop, which probably was the consequence of the previous enthusiasm. The history of the Rorschach has had its turning point thanks to a young student, who had as teachers both Klopfer and Beck: John E. Exner.

The fact of being disciple of both Klopfer and Beck, had two important effects: on the one hand, Exner held in high esteem both of them, and he was not able to prefer one over the other; on the other hand, he knew both methods. This condition led him, in 1960, to make a purpose about a comparison between Klopfer and Beck – rivals since 15 years – with Exner as moderator. The purpose failed.

Neither Beck, nor Klopfer, accepted the invitation. Nevertheless, Beck advised Exner to write a paper comparing the two methods (more or less what Hertz had done before). In 1961, Exner began this work, but suddenly he realized that he could not disregard the other author's contributions. Therefore, the next year he decided to integrate also the remaining three methods. It took him six years to complete the work, which was published in 1969 and named *The Rorschach Systems*. Klopfer, Beck, Hertz, Piotrowsky and Schafer (Rapaport has died at the beginning of the project) received a copy of the manuscript – Exner sent it to them to check whether there were any errors or issues. Klopfer and Beck were disappointed and complained about its lack of a definitive conclusion, while Hertz and Piotrowsky provided Exner with a positive feedback.

Meanwhile, in 1968, the *Rorschach Research Foundation* was born, thanks to the funding received from the National Institute of Mental Health. The latter designated Exner to define which one among the approaches had the best clinical validity and reliability. Since then, the scope of Exner has changed and he has looked at the study of the Rorschach from a different point of view.

In 1971, a computerized dataset was created to easier the direct comparison between the five systems. All the differences among protocols, assessment procedures, coding criteria and interpretation emerged more clearly, together with the absence of an empirical support. At this point, the aim had changed once again: the purpose was to create a new reference system, composed by all those elements that would pass the evidence-based proof. The CS was thus an approach based on psychometric criteria[4], and this procedure removed some codes but at the same time improved the reliability of the ones that left.

In 1997, Exner founded the Rorschach Research Council (RRC), which had the mission to increase the research data, and to improve the information provided by the CS. For two decades, this mission had been preserved. However, while it was true that Exner's works had provided a renewed scientific trust in the Rorschach method, it did not silenced all the concerns about the validity and reliability of the instrument.

In 1993, when CS was published, Archer and Krishnamurthy reviewed the literature derived from 37 studies of interrelation between MMPI and Rorschach, underlining how low were the correlations – and thereby the convergent validity – between MMPI scales and Rorschach indices. In 1997, Gregory Meyer – disciple of Exner and member of the RRC – replied to the paper. The response to Archer and Krishnamurthy (Meyer, 1997a) clarified how MMPI and Rorschach measure personality in two different, complementary, but fundamental ways. The MMPI is a self-reported questionnaire: the required task is structured, it needs reading competences, the task demanded to the subjects is about themselves and the task

---

[4] In order to warrant the reliability, there were excluded all the variables that did not reach an agreement of .85 in groups of 10-15 raters and in at least 20 protocols (Exner, 1969).

does not need particular examiner's skills to provide correct data. The Rorschach test, instead, is a performance-based test: it presents ambiguous stimuli, no reading or cultural competences are required, the task demand is about the inkblots, and the responsibility to provide correct data is in the hand of the examiner. Considered this, it is not surprising that the two tools investigate the personality in two different ways: with the former, what can be investigated are symptoms and experiences of life that the subject is aware of; what instead emerges from the latter are underlying personality characteristics, propensities, mental representations that the respondent is not necessarily mindful of.

However, the two tests have in common that the patients might want to give a certain impression to the examiner, thus a form of censure about what they are going to share might be present in both circumstances. In the case of the MMPI, it is easier to understand how the information provided by the items could be manipulated or what information is wise to not share, while in the Rorschach it is harder to figure out how an interpretation of an inkblot could say something about the personality. Despite these discrepancies, patients may put in place different response styles in front of both tasks: one more compliant, honest and disclosure (Expressive Style), and another one more resistant, deceitful and covert (Constricted Style). Meyer thus decided to (1993) re-test the convergent validity between the two tests, also after taking into account the response styles. Results showed that when subjects adopted the Expressive Style in both tests, the correlations between the MMPI scales and the Rorschach indices that are supposed to assess the same constructs (Affective Distress, Psychosis, and Wariness) were statistically significant and positive, from $r = .22$ to $r = .67$. Conversely, when subjects adopted different response styles in the two tests, the same scales were negatively related. Therefore, any efforts to manipulate one's own symptom presentation, such as "faking good" or "faking bad", would have a different impact on the two tests. Obviously, it is never possible to determine how much of the validity coefficient is due to the response style – and so to the method adopted –, and how it is determined by a genuine trait variance. However, it is a resource and not a limitation the fact that different kinds of measures give

different kinds of information about the personality, especially in the assessment of not ego-syntonic conditions.

A year before the foundation of the RRC, Wood, Nezworski, and Stejskal (1996) published a paper where they brought into question the validity and the reliability of the Rorschach, together with the nature of the whole research base. More in detail, criticisms were about the lack of incremental validity, the influence that R (number of responses) has on the other scores and indices (given that R is influenced by educational level, social class, and intelligence), and about the incoherence between some CS indices presented as "core component" of psychopathologies vs. the "multiple signs" diagnoses required by the DSM-IV. Issues were also related to the reliability. In particular, they pointed out the scores vulnerability to the contamination by situational factors, the influence of the examiner training in the obtained data, the impossibility to do the test-retest procedure because of the memory-effect, and finally they criticized the lack of availability of the Workshop Studies, of which 63% were unpublished. Meyer (1997b) replied to this paper too.

Meyer's answer was focused on clarifying the rationale under statistical methodologies used by CS proponents. In particular, Meyer contested how Wood and colleagues used the classical Cohen's Kappa, without adapting it to the complexity of the C.S. variables. Thus, Meyer replicated the meta-analysis studies using a Cohen's Kappa, after defining chance for the base agreement rates, and results shown from k = .72 to k = .96, with an average of k = .86. So, the interrater reliability of the CS was excellent (Cicchetti, 1994). Finally, the author explained that – in those studies where it was examined the test-retest reliability – the time interval between the "test" and the "retest" was 3 years. Given that a 3-month time is widely accepted as a time interval sufficient to avoid the memory-effect, 3 years should be more than enough. Nevertheless, one of the studies conducted by Exner has tested this possibility: in a test-retest study, the sample was split in two, and at one group – the experimental one – was asked to give responses that were different from the ones they has given the previous time. Event in that case, the median retest coefficient was k = .87.

The Wood's responses was rapid (1997), and then the Meyer's one was too (1997b). In brief time, the controversy became a sophisticated "methodological-fight", made of clarifications, reasons for methodological choices, replications of meta-analyses, criteria for exclude/include studies in meta-analyses, comparisons with other instruments, accusations and defenses about the applicability of the normative data, and from time to time tones became even more sharp. Each publication, replication, and comment seemed to be the last one, which would have put an end to the debate: actually, it was just the beginning.

In 1999, indeed, Garb called for a moratorium aimed to prevent the use of the Rorschach in all psychological uses but the research: «I am calling for a moratorium on the use of Rorschach Inkblot Test in clinical and forensic (but not research) settings. This moratorium should last until we have determined which Rorschach scores are valid and which ones are invalid (Garb, 1999, p. 316)».

Before the publication of Garb's moratorium, over the course of 4 years, the controversy on the Rorschach has included 15 publications (8 proponents and 7 opponents), so allocated: 5 papers[5], 5 comments to the papers[6], 4 replies to the comments to the papers[7], 1 reply to the replies to the comments to the papers (Meyer, 1997b). After the moratorium, the dispute has produced 49 publications (27 proponents and 22 opponents), such as allocated: 6 opponent papers[8], 23 proponent comments to the papers[9] 13 opponent replies to the comments to the

---

[5] Nezworski, & Wood, 1995; Wood, Nezworski, & Stejskal, 1996a; Burns & Viglione, 1996; Ganellen, 1996; Burns, & Viglione, 1997.

[6] Exner, 1995; Exner, 1996; Meyer, 1997a; Garb, Florio, & Grove, 1998; Parker, Hunsley, & Hanson, 1999.

[7] Wood, Nezworski, & Stejskal, 1996b; Wood, Nezworski, & Stejskal, 1997; Garb, Florio, & Grove, 1999; Wood, J. M., Nezworski, Stejskal, Garven, & West, 1999.

[8] Grove, & Barden, 1999; Lilienfeld, Wood, & Garb, 2000; Wood, Teresa, Garb, & Lilienfeld, 2001a; Hunsley, & Giulio, 2001; Wood, Nezworski, Stejskal, & Garven, 2001; Wood, Garb, Lilienfeld, & Nezworski, 2002; Wood, Nezworski, Lilienfeld, & Garb, 2003; Wood, Nezworski, & Garb, 2003; Garb, Wood, Lilienfeld, & Nezworski, 2005; Mihura, Meyer, Dumitrascu, & Bombel, 2013.

[9] Acklin, 1999; Weiner, 1999; Stricker, & Gold, 1999; Viglione, 1999; Hiller, Rosenthal, Bornstein, Berry, & Brunell-Neuleib, 1999; Acklin, McDowell, Verschell, & Chan, 2000; Garfield, 2000a; 2000b; Weiner, 2000; Lerner, 2000; Widiger, 2001; Aronow, 2001; Exner, 2001; Meyer, 2001a; Weiner, 2001; Rosenthal, Hiller, Bornstein, Berry, & Brunell-Neuleib, 2001; Viglione, & Hilsenroth, 2001; Ganellen, 2001; Gacono, Loving, & Bodholdt, 2001; Bornstein, 2001; Hamel, Gallagher, & Soares, 2001; Westen, & Weinberger, 2004; Bornstein, & Masling, 2005.

papers[10], 4 proponent replies to the replies to the comments to the papers[11]. All of which in 4 years.

In 2005 the Society of Personality Assessment (SPA) decided that it was enough. With the publication of *The Status of Rorschach in Clinical and Forensic Practice: An official Statement by the Board of Trustees of the Society for Personality Assessment*, the SPA asserts that the validity and the reliability of the Rorschach Inkblot test are similar to other generally accepted instruments used in the assessment of personality and psychopathology, and that its responsible use in clinical and forensic field is appropriate. Moreover, the SPA drew attention to the importance of taking into account the followings:

- As part of standard clinical care, Rorschach-based inferences, as with inferences from all psychological tests, should be integrated with information from other sources, such as clinical interview and collateral material.
- Clinicians should recognize factors specific to Rorschach testing that may affect or modify interpretation of its scores, such as how engaged a client was with the task.
- The importance of standardized administration and scoring cannot be overstressed. Atypical administration and scoring can lead to incorrect inferences and risk misinterpretation of Rorschach findings.
- It is important to attend to the research literature to ensure Rorschach inferences are consistent with the evidence.

Paradoxically, the fragility of the Rorschach has been converted in its strength: the ambiguity, the lack of structure, and the uniqueness of its stimuli have led who worked on the CS to pay a particular attention to the above

---

[10] Wood, Nezworski, & Stejskal, 1996b; Wood, Nezworski, & Stejskal, 1997; Garb, Florio, & Grove, 1999; Wood, Nezworski, Stejskal, Garven, & West, 1999; Wood, & Lilienfeld, 1999; Hunsley, & Bailey, 1999; Dawes, 1999; Wood, Lilienfeld, Garb, & Nezworski, 2000a; 2000b; Wood, Teresa, M., Garb, & Lilienfeld, 2001b; Garb, Wood, Nezworski, Grove, & Stejskal, 2001; Hunsley, & Bailey, 2001; Wood, Lilienfeld, Nezworski, & Garb, 2001; Wood, Nezworski, Stejskal, & McKinzey, 2001; Schulz, & Waldinger, 2005; Wood, & Nezworski, 2005; Garb, & Grove, 2005.

[11] Meyer, 1997b; Meyer, 2000; Meyer, 2001b; Meyer, & Archer, 2001; Westen, & Weinberger, 2005.

mentioned, more than it has been done with other instruments. What the SPA recommended was nothing more than the basic ethical standards of any good diagnostic practice. It appeared clear how the final diagnostic outcome could be influenced by the quality of the clinical practice, and the CS administration, coding and interpretation was built strictly on this principle.

The next year, John E. Exner died.

In 2013, Mihura and colleagues (Mihura, Meyer, Dumistrascu, & Bombel, 2013) published a systematic review and meta-analyses of the CS, with the aim to impose some order on the Rorschach literature and to gain a more accurate picture of it. After 8 years of quiet, the reply from Wood and colleagues was, again, as rapid as the following of the former authors. In their reply, Wood et al. (Wood, Garb, Nezworski, Lilienfeld, & Duke, 2015) retracted the moratorium of the 2005, and confirmed that there were abundant scientific evidences to justify the use of some CS scores. However, they recommended the use of the new International Norms for the CS (Meyer et al., 2007) and not the previous CS Norms proposed by Exner (Exner, 2003). The answer by Mihura and colleagues (Mihura***) was newly based on methodological issues and erroneous procedures made by the opponents.

Moreover, in 2016 appeared on the scene also Czopp and Zeligman (2016), who judged inappropriate the interpretations of 13 variables, and the criteria for judging their validity, in the meta-analyses of Mihura and colleagues (2013).Currently, the last chapter of this neverending story is the response of Mihura and colleagues (Mihura, Meyer, Dumistrascu & Bombel, 2016), to the Czopp and Zeligman's reply. Given the several errors committedin reporting the meta-analytic methodology used by Mihura and colleagues, they invited Czopp and Zeligman to do a new systematic review and meta-analyses by their own.

When Exner passed away in 2006, he left no explicit instructions for how the Research Council would enact developments while copyright and ownership of his works remained with his family. Thus, for 3 years the members of the Research Council(e.i. Philip Erdberg, John Exner, Christopher Fowler, Roger

Greene, Gregory Meyer, Joni Mihura, and Donald Viglione) explored ways in which the RRC could prepare an updated version of the CS . Ultimately, however, Exner's heirs decided that they wanted to honor his memory by preserving the CS as he left it, with no further changes (Erard, Meyer, & Viglione, 2014, p. 166)».

Therefore, the RRC decided that, after twenty years of updates, the Rorschach method needed an upgrade. Thus, it was born the Rorschach-Performance Assessment System (R-PAS; Mihura, Meyer, Dumitrascu, Bombel, 2011).

2.4. RORSCHACH-PERFORMANCE ASSESSMENT SYSTEM™

A performance test consists in to observe what a person does (not just how well) when is involved in a standardized task, capable to provoke a specific type of response. The main features of this kind of measures are four:

- Standardization, the task situation is strongly controlled, reproducible, and applicable in a way that is almost the same for every subject;
- Fiction, the subject is led to believe that he/she understood the aim of the task, while the examiner is interested in some other aspects of the performance;
- Specific interest: the examiner is interested in studying one or more traits, carefully defined in order to distinguish them clearly;
- Multivariate data: the examiner registers the procedure used by the subject to solve the task, not only the result of the performance (Chronbach, 1949).

In the R-PAS, the core of the performance is represented by the instruction given at the beginning of the administration: the subject is presented the first card and is asked the question «*What might this be?* ». This request induces the respondent to a "misperception" of the image, and forces him/her to avoid the unique, obvious, right response "an inkblot". The reason why it is so important

the role of the instruction is, indeed, in the fact that it generates a condition of problem solving. The respondent has to put in place a pool of strategies in order to execute the task. Effectively, those strategies show the person's way of functioning, namely which personality traits are prominent in his/her ordinary life, especially in situations not familiar, not structured or ambiguous (as the inkblot is). The coding is the translation of the strategies used by the respondent in quantitative variables.

The switch from the previous method (i.e. CS) was made by taking firmly into account six foundation principles:

I. Selecting and highlighting those variables with the strongest empirical, clinical, and response process/behavioral representational support, while eliminating those with insufficient support.

II. Comparing test takers' scores to a large international reference sample, using a graphic array of percentiles and standard score equivalents.

III. Providing a simplified, uniform, and logical system of terminology, symbols, calculations, and data presentation, in order to reduce redundancy and increase parsimony.

IV. Describing the empirical basis and psychological rationale for each score that is to be interpreted.

V. Providing a statistical procedure to adjust for the overall complexity of the record and a graphical illustration of its impact on each variable.

VI. Optimizing the number of responses given to the task in order to ensure an interpretable and meaningful protocol, while drastically reducing both number of times the task needs to be re-administered because of too few responses and the likelihood of inordinately long and taxing administrations because of too many responses.

VII. Developing new and revised indices by applying contemporary statistical and computational approaches.

VIII. Offering access to a scoring program on a secure, encrypted web-platform any device that can inter-face with the Internet (e.g. PC, Notebook, iPad) (Meyer, Erard, Erdberg, Mihura, & Viglione, 2011).

2.4.1.   FROM CS TO R-PAS

Other than a consistent amount of papers, the controversy on the CS actually provided awareness about some gaps in the method. Thus, the transition from CS to R-PAS was characterized by four primary considerations: 1) selected variables should have empirical support in the literature, 2) behavioral representation of the inferred characteristic in the coded response behaviors, 3) perceived utility based on the experience of 246 practitioners, 4) parsimony (Meyer & Eblin, 2012).

All variables existing in R-PAS were already present in others previous methods. Those variables that did not have an empirical support, or were redundant, were eliminated or simplified (Gacono, Bannatyne-Gacono, Meloy, & Baity, 2005; Katko, Meyer, Mihura, & Bombel, 2010; Bornstein, 1996; Mihura et al., 2012; Diener, Hilsenroth, Shaffer, & Sexton, 2011; Graceffo, Mihura, & Meyer, 2012; Bandura, 1954a, b; De Koninck & Crabbe-Decleve, 1971; Dumitrascu, Mihura & Meyer, 2010; Nelson, 1954; Stein, 1973). In doing this, some of these variables were included with different names. Moreover, in order to improve reliability, R-PAS has provided more complete and clear definitions of them. Finally, all indices in R-PAS have a continuous distribution, also those that in CS were dichotomous (e.g. HVI, S-CON). This last change was necessary for two reasons. First, discrete measures can only assume a limited number of values, and do not use all of the available variance, indeed recent evidences show that in psychopathology continuous measures are more representative and valid than discrete measures (Makron, Chmielewski, & Miller, 2011). Additionally, the last version of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2013) proposes a dimensional scheme in clinical assessment. So there are both empirical as well as theoretical reasons for using continuous rather than dichotomous indices also with the Rorschach.

Furthermore, in R-PAS the names of indices and variables are more intuitive than with CS (e.g., *Special Scores* are simply named *Cognitive Codes,* the

percentage of *FQ-* responses is named *FQ-%* instead of *X-%*). In addition, the interpretation procedure is simpler, and more accurate: in CS, there was a *step-by-step* procedure (Exner, 1991), while the interpretation in the R-PAS is based on the observation of the data, which are grouped into eight clusters: *Resources and Controls*, *Affect*, *Ideation*, *Mediation*, *Information Processing*, *Interpersonal Perception*, and *Self-perception*. In each cluster, the variables and indices are presented into a graphic, with percentiles and standardized normative data. Moreover, variables and indices with strong empirical support are presented in Page 1, those whose validity and interpretative significance are more tentative are presented in Page 2 (Meyer, & Eblin, 2012). The normative data lead more accurate information and are more representative of the nonclinical population of CS normative data. With regard to the reference samples, however, they are quite representative of the cultures and languages in the Western Hemisphere, whereas the eastern countries, cultures and languages are under-represented.

The main innovation related to interpretative process is the variable named "Complexity." As it could be inferred by the name, Complexity represents how articulated and elaborated is the response process, taking into account the number of responses, the space representation, the ideas that come into mind, and the respondent's ability to describe the characteristics of the inkblot which determined the response.

The R-PAS administration is a bit different from the CS one. The guidelines for the clarification phase were updated, becoming clearer than the previous ones. A more rigorous and standardized clarification phase brings with it two advantages: a lower variability across the examiners, and more informative protocols. An examiner who does not investigate determinants when it is necessary to do so ends up with losing some basic information; conversely, an examiner who overly investigates certain components of the verbalization when it is not necessary to do so, tends to produce a more complex protocol than it would be if clarification were conducted more optimally.

In his *Psychodiagnostic* (1922), Herman Rorschach declared that a protocol with less than 14 responses is useless. Indeed, in the CS, to achieve

information from a protocol, the minimum number of responses (R) was 14 and there was not a maximum limit. Furthermore, Exner noted that protocols with 14 to 17 responses tend to be poorly informative (Reese, Viglione, & Giromini, 2014). Conversely, too many responses can lead to overpathologize. To fix these issues, R-PAS introduced the R-Optimized administration: Before initiating the administration, the respondent is asked to give two or maybe three responses per card. Next, during the spontaneous response phase, if the respondent provides only one response to any given card, the examiner asks for a second one (prompt, Pr); if the respondent goes on further the third response and gives a fourth response to any given card, the examiner asks for the card back (pull; Pu). In this way, R is a number that cannot exceed 40, and typically is comprised between 20 and 30. The minimum allowed number of responses per protocol is 16 (the minimum threshold was raised).

At the end of this process, the R-PAS coding was based on the 13 categories:

Coding Sequence Question	Coding Category	Codings
1. What is the card angle?	Card orientation	> < v @
2. Where is it seen?	Location	W, D, Dd
3. Is white space used? If so, how?	Space	SI, SR
4. What is seen?	Content Class	H, (H), Hd (Hd), A, (A), Ad, (Ad), An, Art, Ay, Bl, Cg, Ex, Fi, Sx, NC
5. Are at least two objects meaningfully related?	Synthesis	Sy
6. Are the objects all vague?	Vagueness	Vg
7. Are the objects identical?	Pair	2
8. How well does it fit the blot?	Form Quality	o, u, -, n

9. Do many people see it?	Popular	P
10. What makes it look like that?	Determinants	M, FM, m (a, p, a-p), FC, CF, C, C', Y, T, V, FD, r, F
11. Are there issues with thought processes?	Cognitive codes	DV1/2, DR1/2, PEC, INC1/2, FAB1/2, CON
12. What critical themes are present?	Thematic Codes	ABS, PER, COP, MAH, MAP, AGM, AGC, MOR, ODL
13. Were steps taken to manage R?	R-Optimized	Pu, Pr

Card Orientation: when the response is given without turning the card, no card orientation is needed. All others behaviors should be coded as < or >, respectively if the card is turned on the left or on the right side, and as @ if the card was turned at least 90 degrees an then putted again in the upright position.

Location and space: where on the inkblot or in the white space the response is seen.

W	Whole	Response uses the entire inkblot
D	Common Detail Area	Response uses one or more frequently used detail areas
Dd	Unusual Detail	Response uses one or more rarely used detail areas
	White Space	
SR	White Space Reversal	Non-inked or background area on card is a focal percept such that the traditional figure and ground become reversed
SI	White Space Integration	Non-inked or background area on card is integrated with inked areas

Content Codes: what is seen in the card, the examiner code all the categories that apply but only code each category once per response.

H	Whole human, including realistically described religious or historical figures, also code figures described as humans but with non-human parts.
(H)	Imaginary, fictional, quasi-, or supernatural whole human.
Hd	Human detail, for an incomplete human form.
(Hd)	Imaginary or fictional human detail.
A	Whole animal.
(A)	Imaginary, fictional, or cartoon whole animal.
Ad	Animal detail, for an incomplete animal form.
(Ad)	Imaginary or fictional animal detail.
An	Anatomy, for internal body parts and structures that are not visible from the outside. Also for perceptions of anatomy from medical imaging devices.
Art	Art, for objects of art, or for objects that are, or are described as being, decorative or ornamental.
Ay	Anthropology, for references to a specific historical or cultural context.
Bl	Blood.
Cg	Clothing.
Ex	Explosion, bomb blast, volcanic eruption, and fireworks.
Fi	Fire, flames, embers, or smoke.
Sx	Sexual organs, activity, or clothing.
NC	Objects and contents that are not classified in other categories, including abstractions like depression and sensory experiences.

Object qualities: codes the peculiar features of what is seen in the inkblot.

Sy	Synthesis	Distinct and separate objects in a relationship
Vg	Vagueness	Objects with vague or indistinct outline or boundary
2	Pair	Identical objects based on the symmetry of the blot

Form Quality: how well the object fits the location used by the subject and how frequently it is identified.

o	Ordinary	Form fit that is relatively frequent and accurate
u	Unusual	Form fit that is of intermediate frequency or accuracy or both
-	Minus	Form fit that is infrequent and inaccurate
n	None	Response does not contain an object with definite form or outline

27

Popular Responses: those responses that are given by at least one of every three people are codified as P, namely a specific area and a specific contents have to match in the Popular Response.

Determinants: those perceptional characteristics of the card that the respondent takes in account in order to interpret the stimulus.

Human Movement	M	Human (or superhuman) movements such as activity, experience, sensation, and emotion. Includes animals involved in exclusively human activities.
Animal Movement	FM	Movements performed or emotional expression that are suitable for animals.
Inanimate Movement	m	Inanimate movements in many forms, including mechanical or inorganic movement, natural forces such as gravity and wind, or even rigor mortis of a human body.
Active vs Passive	a p	The passive to active continuum measures the amount of effort or force incorporated in a movement. Passive movements are marked by relatively less effort and force than active movements.
Form Color	FC	Color contributes to a response object but form is dominant
Color Form	CF	Color is dominant in a response object but form contributes
Color	C	Color determines a response object without form playing a role
Achromatic Color	C'	Is coded for presence or absence of black, grey, or white.
Diffuse Shading	Y	It involves the perception of shading as a part of the response.
Vista	V	Shading determines dimensionality or perspective
Texture	T	Shading determines tactile impression
Form Dimension	FD	The blot outlines generate a perception of depth or dimensionality
Reflection	r	The symmetry of the card is interpreted as an objects and its reflection.
Pure Form	F	The form of the inkblot is the only determinant.

Cognitive codes: behaviorally, cognitive codes capture disrupted or illogical thought processes that are indicative of thought disturbance and confusion: level 1 codes characterize mild to modest examples of thinking problems and level 2 codes represent moderate to severe disruptions in thinking.

Language & Reasoning		
DV1/2	Deviant Verbalization	A mistaken or inappropriate word or phrase to communicate or to describe a response
DR1/2	Deviant Response	Distortions or confused, rambling, or circumstantial language that drifts from the task
PEC	Peculiar Logic	Peculiar, strained, confused, or overly concrete reasoning
Perceptual		
INC1/2	Incongruous Combination	Implausible or impossible attributes are ascribed to an object
FAB1/2	Fabulized Combination	Implausible or impossible relationships between two or more distinct response objects
CON	Contamination	Two mutually exclusive response objects that are visually superimposed on each other in the same blot area

Thematic Codes: identify meaningful content characteristics. Unlike most other codes, some Thematic Codes are based on spontaneous elaborations and characterizations that are not obvious visual features of the responses themselves.

ABS	Abstract Representation	Concrete blot features are representational and symbolize an abstract, higher order construct or concept.
PER	Personal Knowledge Justification	Personal knowledge or experience is used to justify a response.
COP	Cooperative Movement	Cooperative, positive, or pleasant interactions are occurring between two objects.
MAH	Mutuality of Autonomy-Health	Two objects are mutually and autonomously engaged in a reciprocally interactive activity.
MAP	Mutuality of Autonomy-Pathology	An agent or object intentionally compromises the autonomy or integrity of another object or is destructive to it.
AGM	Aggressive Movement	Aggressive or hostile activity, intent, or ideation is occurring.
AGC	Aggressive Content	Response content involves an aggressive, dangerous, harmful, injurious, malevolent, or predatory element.
MOR	Morbid Content	Objects are damaged or states of distress or dysphoria are attributed to them.

| ODL | Oral Dependency Language | Response Phase verbalizations linked to oral activity and content or interpersonal passivity and dependence. |

III. EMPATHY IN PERSONALITY ASSESSMENT

3.1. DO WE HAVE AN EMPATHY INDEX?

Empathy, together with all its features (i.e. as emotion recognition, emotion contagion, capability to assume the perspective of others, compassion, sympathy), has recently come to light in the form of peculiar components of some psychopathologies present in the DSM 5 (APA, 2013), such as Narcissistic Personality Disorder (NPD), Antisocial Personality Disorder (AsPD), Schizophrenia Spectrum (SS), Autism Spectrum Disorder (ASD).

More in general, the construct of empathy has received increasing attention over the last few decades. But, what can we say about the instruments, in the field of personality assessment? Do they have any specific scales or indices targeting this construct?

The American Psychological Association website, under heading "Specialized Knowledge" of Personality Assessment (American Psychological Association, 23/10/2017), cites the MMPI-2 and the PAI as major self-reported inventories and the Rorschach Inkblot Test as best known performance-based measures.

The Minnesota Multiphasic Personality Inventory-2, Restructured Form (MMPI-2-RF; Ben-Porath, 2012) includes 10 Restructured Clinical (RC) Scales: Demoralization, Somatic Complaints, Low Positive Emotions, Cynicism, Antisocial Behavior, Ideas of Persecution, Dysf Negative Emotions, Aberrant Expriences, Hypomanic Activation. The Antisocial Behaviors scale (RC4) contains the items about thefts, uninhibited sexual behavior, drugs or alcohol use, bad companies, troubles with law, misconducts at school, and impulsiveness.

Ultimately, RC4 items are more related to unethical behaviors, or actions against society, pair groups or family, thus more focused on lack of morality rather than of empathy.

However, it is possible to identify some items that intuitively might have to do with empathy (or the lack of it), as enjoying hurting people, being worried about having injured other people's feelings, and having poor reactions to animals' suffering. Surprisingly, those items do not converge in a scale about empathy, or in a scale that has to do with related psychopathologies. That is, even if those items are much more focused on the pain of someone else -especially when the suffering is voluntary caused- they do not converge in a scale of empathy or of a personality disorders related to its lack.

In the Personality Assessment Inventory (PAI), there is the Antisocial Features scale (ANT), which is subdivided into three subscales: Antisocial Behaviors (ANT-A), Egocentricity (ANT-E), Stimulus-Seeking (ANT-S). Compared to the RC4, the ANT-E is closer to that kind of empathy typically lacking in Antisocial or Narcissistic Personality Disorders. Indeed, it «was conceptualized to include the pathological egocentricity and narcissism often thought to lie at the core of this disorder» (Morey, 2007, p. 112). However, even in this case, the majority of items investigates morality and the common sense about what is right or wrong. Again, as it was for the MMPI-2-RF, some other PAI items can be related to empathetic behaviors but that are not taken into account for the assessment of NPD and APD. Those items investigate being helpful to other people, try to be sympathetic with others, and remorse and guilty conscience.

In its current version, available R-PAS variables do not address antisocial behavior or narcissism directly. However, there are variables that assess how the respondent comes into relations with others.

- SR (Space Reversal): the "this" of the question "*What might this be?*" is typically understood to refer to the inkblot itself. Therefore, when the respondent uses the white background to organize and structure his/her response, rather than the inked area, he/she complies by giving a response,

but does so in a different or unexpected way. In particular, this often represents oppositionality, individualism, inventiveness or creativity, an atypical structuring to perception, or even being sensitive to interpersonal pressure so as to resist it.

- H (Human Content): involves a more accurate, integrated, and complete view of people and suggests the potential for a cognitively sophisticated and realistic view of oneself and others. Most H responses involving also the elaboration of the environment where they are (objects, intentionality, narrative implications, movement, and relationship) thus, in turn, entail more elaborated and informed schemas for understanding people. .

- R (reflection): is a representation that suggests the respondent may be self-focused in everyday life, including relationships, and other's experiences.

- M: historically, the M responses are considered indices of higher cognitive functioning, ability to imagine (the inkblots do not move), and empathy (Giromini, Porcelli, Viglione, Parolin, & Pineda, 2010; Pineda, Giromini, Porcelli, Parolin, & Viglione, 2011; Andò, Salatino, Giromini, Ricci, Pignolo, et al., 2015; Porcelli, & Kleiger, 2016; Giromini, Viglione, Brusadelli, Zennaro, Di Girolamo, et al. 2016; Giromini, Viglione, Pineda, Porcelli, Hubbard, et al. 2017). The cognitive abilities associated with giving an M response include the ability to imagine and to envision action or emotion, to reflect on life experience, and a degree of developmental maturity. M is therefore a type of mentalization that contributes to the capacity for identification with others.

- COP (Cooperation): the propensity to have cooperative, collaborative, synchronized interactions suggests a person who has a generally positive template for envisioning relationships.

- MAH vs MAP (Mutuality of Autonomy Healthy vs Pathological): MAH suggests attention to relationships in one's environment, and thus a healthy and productive understanding of others. MAP, conversely, are present in persons that have envisioned a controlling, malevolent, and/or destructive relationship.

- GHR/PHR: the first is coded when human content or human activity is accompanied by codes indicating an accurate, logical, adaptive, and non-

malevolent percept. The structural features of the response suggest an ability to envision the self and relationships with others in an adaptive or positive way. It implies an intact understanding of self and other. The second, is automatically assigned if one or more problematic codes accompany responses involving human representations or human activity. Therefore, PHR indicates that human images, which are analogs to self or interpersonal schemas, have been viewed in a structurally or thematically problematic way that suggests a propensity to misunderstand others, relationships, and/or the self.

- AGC (Aggressive Content): aggressive, powerful, dangerous, predatory, or threatening images coded by AGC is a behavioral indication that these themes are on the person's mind.
- ODL (Oral Dependent Language): An elevated frequency of these terms or images identifies respondents who are implicitly motivated by dependent needs, related to an underlying dependent trait or a state.

Trying to summarize, in the MMPI-2-RF and in the PAI there are no items or scales that assess lack of empathy (or empathy itself) directly. Rather, there are items and scales that assess characteristics and behaviors usually related with the lack of empathy. More exactly, there are items and scales that assess characteristics and behaviors usually present in people who often are devoid of empathy. Concerning the R-PAS, its condition is similar to the other two tools, perhaps with the exception of the M variable, which is deemed to also assess the empathy capabilities of the respondent.

3.2. EMPATHY ASSESSMENT TOOLS

While the most widely used, personality assessment instruments do not specifically focus on the measurement of empathy –perhaps because of that – during the last four decades many measures for its evaluation have been developed. The big majority of these instruments are self-report questionnaires,

followed by some tasks about expressions recognitions, perspective taking assumption, and understanding of ambiguous stories. Furthermore, the recent development of the neuroimaging techniques made it possible the combined use of these tasks and neuroimaging instruments (Guste, Di Girolamo, Giromini, 2017).

3.2.1. SELF-REPORTED QUESTIONNAIRES

Interpersonal Reactivity Index (IRI; Davis M.H., 1980, 1983). The IRI is probably the most famous instrument for the empathy assessment. Given that for a lot of time it was the only tool available, it is also the most used one in research. It was designed to measure different empathic tendencies, through 28 items split in 4 subscales on a seven-items Likert scale: a) Perspective Taking (PT); b) Fantasy (FS); c) Empathic Concern (EC); d) Personal Distress (PD). The Cronbach's αs coefficients are ranging from .68 to .79.

Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004). The second well-known self-reported instrument is the EQ. Its author is Baron-Cohen, an expert on the empathy field, notably in autistic and Asperger diseases. The EQ is composed by 60 items, on a scale from 0 to 2. The EQ measures three different empathy features: cognitive empathy, emotional reactivity and social skills (Lawrence et al., 2004). The Cronbach's α varies from the language the questionnaire is translated to, but more or less is always in the range of .85.

Empathy-Based Stories (MES; Eskola, 1998). The aim of this instrument is to see different points of view of different people, and their personal experiences in emotionally heavy situations. This instrument works by giving to the participants a sheet of paper with a short story wrote on one side. The task consists into finish the story, writing it on the other side of the sheet. Those stories might be very different, can represent different fields, and depend on the examiner. The advantage of this tool is that people are completely free to write everything they want. To analyze the data of this instrument, matrixes are used.

In every study, matrixes are different, thus it depends completely on the examiner what he wants testing or improving.

Basic Empathy Scale (BES; Jolliffe & Farrington 2006). BES is a self-report 20-items instrument, which was developed to measure affective and cognitive empathy. The questionnaire is divided into two types of subscales, on a Likert scale based on 5 points: 9-item *Cognitive Empathy Subscale* and 11-item *Affective Empathy Scale*. The Cronbach's α for the affective empathy subscale is 0.81, and for the cognitive empathy subscale it is 0.84.

Questionnaire Measure of Emotional Empathy (QMEE; Mehrabian & Epstein, 1972).The QMEE is a self-report that measures emotional empathy. It has 33 items that has to be answered by a 9 point ratings from -4 (strong agreement) to 4 (strong disagreement). QMEE has 7 subscales: a) vulnerability to emotional contamination, b) appreciation of the feelings of unfamiliar people; c) strong emotional response, d) tendency to be emotionally touched by other people negative emotional capabilities, e) sympathetic capacities, g) wish to be in contact with people that have problems (Mehrabian & Epstein, 1972). The reliabilities of the subscales are medium: the Cronbach's αs vary from 0.63 to 0.80 (Lyons & Hazler, 2002).

Toronto Empathy Questionnaire (TEQ; Spreng et al., 2009). TEQ is a 16-items self-report instrument that was developed to measure different types of empathy. It has to be answered by a Likert type scale of 5 points. Its reliability is medium (Cronbach's α is 0.88).

Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers et al., 2011). The QCAE is composed of 31 items, rated on a 4-point Likert scale: strongly agree (1), slightly agree (2), slightly disagree (3), and strongly disagree (4). Scores from each item were added to produce two subscales: Cognitive Empathy and Affective Empathy. These subscales are, in turn, composed of subcomponents. The subcomponents of Cognitive Empathy are Perspective Taking (PT) and Online Simulation (OS). PT measures the capability to put oneself in another person's shoes, while OS assesses attempt to put oneself in another

person's position by imagining what that person is feeling and is likely to be used for future intentions. The subcomponents of Affective Empathy are Emotion Contagion (EC), Proximal Responsivity (PrR) and Peripheral Responsivity (PeR). EC is focused on the automatic mirroring of other's feelings. PrR is a measure of the emotional responsiveness to the feelings of others who are close within the social or affective subject's context. PeR is similar to PrR, however, its context is detached, such as experiencing empathy with protagonists in a film or a novel.

3.2.2. EMPATHY TASKS AND NEUROIMAGING

Functional Magnetic Resonance Imaging (fMRI) activation paradigm (Vollm et al., 2006). The fMRI activation paradigm is used to expose activation areas related with empathy processing. It is a visual activation paradigm, which consists of a series of cartoons with different short stories in every block of pictures. The cartoons can contain a two type stories: Physical and Empathy. At the beginning of the procedure, a series of questions are given to participants. The scenarios of each series are made in the way that the characters of cartoons continue their story in upcoming pictures, the participant has to put him/herself in the protagonist's shoes. Each of the blocks is shown for only 4 seconds in the upper part of the screen, then, for other 4 seconds, the possible endings of that story are shown at the bottom part of the screen. The participant has to choose the answer to the question that is given before, and only one of the two possibilities is right: the more correct answers a person gives, the higher level of empathy he/she has. Therefore, by doing fMRI study, it is seen which part of the brain is active while doing a part regarding empathy (Kim et al., 2010).

Electromyography (EMG). This instrument is used to capture the electrical activity of facial muscles, because it is related with the exposition to facial expressions (Ekman & Friesen, 1975; Tomkins, 1991). The aim of the tool is to catch all the facial reactions in the face-to-face situation. Moreover, the main variable is the time of the reaction: faster a person shows his reaction, a better

understanding of other's feelings that the person should have. However, in the previous studies it was found that reactions can be learned or controlled by the person, so it is not always a good way to measure empathic responses (Sonnby-Borgstrom et al., 2003).

Reading the Mind in the Eyes-Test (RME-T; Baron-Cohen et al., 1997; Baron-Cohen, Wheelwright, Hill et al., 2001). The RME-T is used to assess emotion recognition; it includes 36 still pictures of the eye region, the person has to choose among four options about what emotion the pictures could represent. The aim of the test is to measure the frequency of matching a semantic definition to its expression in the picture, and the score is calculated by the sum of correct responses.

Electroencephalogram (EEG) activity. The EEG works by showing different types of pictures on the computer to the participant. After the subject is connected to EEG apparatus, for 3' the person is recorded in the resting state with eyes closed. Than a series of pictures are shown to him/her: 1'30'' positive emotions, 1'30'' neutral stimuli, for 1'30'' erotically pictures, for 1'30'' negative pictures. Lastly, the EEG is recorded again in the resting state for 3' with the eyes closed. In between of each series some grey-colored pictures with meaningless context are presented for 1'30''. After that EEG was recorded, to the participant is asked to value every block of pictures that they have seen before by the scales of 9 points (the most positive, negative or neutral, and the most activating, relaxing or neutral). Persons that received most points on emotionally active pictures were considered that are more empathic than those who got lower scores. Again, by doing EEG measure, it is possible to see which part of the brain activates when a person watches different pictures that represents different stimuli (Demidova et al., 2013).

Performance-Based Measure of Empathy (Derntl et al., 2009). This is a computer task that includes: facial affect perception, affective responsiveness, emotional perspective taking. Two forced-choice responses are given, so an accuracy of the responses are also considered. The person has to recognize the

emotion in the picture and choose the answer as fast as possible (the reaction time is one of the measure of the test). The less time he/she takes, the better accuracy a person gets, than a higher level of empathy and of understanding emotional states of others is considered that the participant has.

Multifaceted Empathy Test (MET; Dziobek & Heekeren, 2008). The MET is a rating scale that measures cognitive and affective empathy. It consist of different series of the photographs, where are represented some emotionally stimulating situations. Cognitive empathy is measured asking for a person to try to name the mental states of the persons in the photographs. After this part, a person is informed about the correct answers that he gave. Later on, for the emotional empathy, it is asked for the participants to rate their personal emotional reactions to the given pictures. The MET consists of 23 pairs of different pictures. For the level of excitement by using Self-Assessment Manikin (SAM; Lang et al., 1997). It is a visual-analogue scale that has a rating scale from 0 to 9. The reliability of the MET varies from medium to high. Cronbach's α differs from 0.71 to 0.92.

Story-Based Empathy Task (SET; Dodich et al., 2015). SET is a non verbal task that was developed to measure an intention and emotion acknowledgment. This task is based on original cartoons, where is asked to identify intentions (SET-IA) and emotional states (SET-EA). The control condition is the inference of causality reaction (SET-CI), based on the personal knowledge of the individual of physical properties of objects and human bodies (Dodich et al., 2015, p.1908). Each of these conditions are composed of six pictures and then it is asked to choose a possible ending for that story given in the picture. Each of the parts can be valued maximum of 6 points. More points a participant gets, higher level of understanding about other's feelings it is considered that he/she has.

Social Relations Model (SRM; Kenny & La Voie, 1988). SRM is a rating scale type of measure that was developed to describe dyadic relationships when the components are assessed on a continuous scale. SRM is divided into three components: perceiver (how the participant sees other people), target (how the person him/herself is seen by other people), relationship (how a perceiver sees the target) (Kenny, 2001). Two ways to use this model can be used: *round robin*

or *block*. In *round robin*, every member of the group has to rate or judge every other person in the group. In *block*, a group is divided into two subgroups, and each person from each subgroup has to rate or judge every person from other subgroup. SRM is an instrument that helps a participant to understand, if other people do understand his/her emotions and helps to see, if the participant him/herself understands other individual's emotions.

IV. A CONTRIBUTION TO THE ITALIAN VALIDATION OF QCAE

When this research started (in 2014), the Questionnaire of Cognitive and affective Empathy (QCAE; Reniers, et al., 2011) was the most recent, self-reported questionnaire for the assessment of both components of empathy. Furthermore, the QCAE had already been used in an Italian study (Giromini et al., 2016). Thus, we decided to use it as criterion variable to select R-PAS variables related to empathy and work on the development and validitation of a Rorschach Empathy Index.

In the research of Giromini and colleagues (2016), the authors used a paper-and-pencil version of the QCAE, but this instrument was developed as an online administration test. Given that, we inspected its validity and reliability in an online sample, too.[12]

The QCAE was developed in 2011, based on the most recent conceptualizations of empathy, viewing it as a multidimensional construct. More in detail, Reniers et al. (2011) administered a number of widely used self-report measures of empathy—including the IRI and EQ—and factor analyzed their results to obtain a pool of items measuring either cognitive or affective empathy. Compared to other similar empathy measures, the QCAE thus offers the advantage of providing separate, reliable scores for the cognitive and affective components of empathy. Furthermore, these Cognitive Empathy and Affective Empathy scales are, in turn, composed of subcomponents. The subcomponents of Cognitive Empathy are perspective taking (PT) and online simulation (OS). PT measures the capability to put oneself in another person's shoes, whereas OS assesses attempts to put oneself in another person's position by imagining what

---

[12] All the information present in this chapter are taken from the paper Marzia Di Girolamo, Luciano Giromini, Christina L. Winters, Colinda M. B. Serie & Corine de Ruiter (2017): The Questionnaire of Cognitive and Affective Empathy: A Comparison Between Paper-and-Pencil Versus Online Formats in Italian Samples, Journal of Personality Assessment, DOI: 10.1080/00223891.2017.1389745.

that person is feeling and is likely to be used for future intentions. The subcomponents of Affective Empathy are emotion contagion (EC), proximal responsivity (PrR), and peripheral responsivity (PeR). EC is focused on the automatic mirroring of other's feelings. PrR is a measure of the emotional responsiveness to the feelings of others who are close within the social or affective subject's context. PeR is similar to PrR, but its context is detached, such as experiencing empathy with protagonists in a film or a novel.

AIM OF THE STUDY

Because we intended to use the QCAE as the criterion variable for our Rorschach Empathy Index development, prior to testing the association of R-PAS variables to QCAE scores, we inspected the psychometric properties of an Italian version of the QCAE. The goal of the study described in this chapter was thus to contribute to the Italian validation of the QCAE.

4.1. METHODS

This study used two data sets derived from two research projects, which used different methods of administering the QCAE (paper-and-pencil vs. online). The first project aimed at investigating an interpersonal competence measure, and used the QCAE in its paper-and-pencil version, to investigate convergent validity (Giromini et al., 2016). The second project aimed to examine an online format for QCAE administration.

4.1.1. PARTICIPANTS

Paper-and-pencil data set: The paper-and-pencil data set consists of data from a study conducted by Giromini et al. (2016). After translating the questionnaire to Italian, using the translation–back translation method, the authors administrated the QCAE in paper-and-pencil format. The original sample

size of Giromini et al. (2016) study consisted of 408 students from an Italian university, ranging in age from 18 to 57 (M D 22.6, SD D 4.6); 74% were women. However, one of the participants did not fill out the QCAE, so our final sample was reduced to 407. Although the authors inspected central tendency, dispersion, and internal consistency, Giromini et al. (2016) did not present detailed analyses on the reliability and validity of the Italian QCAE.

Online data set: The online data set was collected to evaluate the psychometric properties of the QCAE obtained by means of online administration. The same Italian QCAE, translated into Italian by Giromini et al. (2016), was used in this study. A number of other self-report and performance-based measures were administered to examine convergent validity. For this study, we only used those instruments that have previously been validated for use within the Italian context. These instruments are detailed later.

The original sample size of our online data set was 287. We decided to exclude two participants: one because she was 17 years old, and one because she was non-Italian and resided in Italy less than 10 years (2.7 years). Our final sample included 285 participants from 18 to 68 years old (M D 26.4, SD D 7.0), 224 of whom were women (78.6%). About 60% were university students (n D 166) and the other 40% was comprised of individuals with various occupations or unemployed.

4.1.2. PROCEDURE

Paper-and-pencil data set: These data were collected at two Italian universities, located in Milan and Rome. Prospective participants had been personally invited in class by the research assistant to volunteer for a study on psychology and interpersonal relationships. Inclusion–exclusion criteria were (a) Italian citizenship, (b) fluent in the Italian language, and (c) not receiving psychiatric medications.

Online data set: This sample was recruited using flyers, social networking, and word of mouth; the volunteers were informed about a research study on the capability to recognize others' emotions. The data were collected using the Google Form service. Inclusion–exclusion criteria were the same as in the paper-and-pencil study.

4.1.3. MEASURES

Both data sets: All participants were administered the QCAE along with a number of other psychological scales. Ideally, to evaluate the convergent validity of a new empathy measure, one should try to use the most widely accepted measures of empathy; that is, the IRI (Davis, 1980) and the EQ (Baron-Cohen & Wheelwright, 2004). Although both had been validated in Italy (Albiero, Ingoglia, & Lo Coco, 2006; Preti et al., 2011), 21 items of the QCAE were derived exactly from these two instruments (6 from the IRI and 15 from the EQ; Reniers et al., 2011). For this reason, convergent validity was tested by focusing on constructs only close to empathy; that is, interpersonal competence, openness, extraversion, agreeableness, well-being, emotional regulation, and emotion recognition. The translation of the QCAE into Italian was made in accordance with the classical translation-back translation procedure (Geisinger, 2003): First, a bilingual individual translated the English original version into Italian, and then a second bilingual individual who was blind to the original QCAE version back translated the Italian version into English to identify potential discrepancies. The final, Italian QCAE version was eventually approved by two expert researchers who speak both Italian and English fluently.

PAPER-AND-PENCIL DATA SET

In addition to the QCAE, participants included in this sample also completed the following questionnaires. Interpersonal Competence

Questionnaire. The Interpersonal Competence Questionnaire (ICQ; Buhrmester et al., 1988) is composed of 40 items measured on a 5-point Likert scale. The items make up five subscales: (a) the ability to initiate relationships, (b) the ability to assert displeasure with others, (c) the ability to disclose personal information, (d) the ability to provide emotional support and advice, and (e) the ability to manage interpersonal conflict. Reliability and validity of ICQ scores was demonstrated by Buhrmester, Furman, Wittenberg, and Reis (1988) and Giromini et al. (2016) for the original and the Italian versions, respectively. In our sample, Cronbach's alphas were .86 (initiation relationship), .77 (emotional support), .77 (negative assertion), .81 (disclosure), and .78 (conflict management).

NEO Five-Factor Inventory: The NEO Five-Factor Inventory (NEO-FFI; McCrae & Costa, 2004) is a short version of the Revised NEO Personality Inventory (NEO PI–R; Costa & McCrae, 1992), an instrument that measures personality traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism. It is made up of 60 items, all measured on a 5-point Likert scale. The Cronbach's alpha values in our sample were .68 (openness), .72 (conscientiousness), .63 (extraversion), .62 (agreeableness), and .75 (neuroticism), which were similar to the ones reported by McCrae and Costa (2004) and those found in the Italian validation study of the NEO-FFI (Caprara, Barbaranelli, Hahn, & Comrey, 2001).

Psychological General Well-Being Index: The Psychological General Well-Being Index (PGWBI; Dupuy, 1977, 1984) is a 20-item self-report scale that assesses psychological well-being. Each item is measured on a 6-point Likert scale and the total score is broken down into six subscales: absence of anxiety, absence of depressed mood, positive wellbeing, self-control, general health, and vitality. In this study we used the Italian version validated by Grossi, Masconi, Groth, Nievo, and Apolone (2002). Cronbach's alphas were .84 (absence of anxiety), .76 (absence of depression), .82 (positive well-being), .56 (self-control), .59 (general health), .69 (vitality), and .92 (total PGWBI score). Since the introduction of the original version of the PGWBI (Dupuy, 1984), many studies

have used this instrument and provided support for its validity (e.g., Badia, Guti érrez, Wiklund, & Alonso, 1996).

Difficulties in Emotion Regulation Scale: The Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004) is a measure of difficulties in emotion regulation. It includes 36 items measured on a 5-point Likert scale, differentiating six areas of emotion regulation problems: (a) nonacceptance of emotional responses, (b) difficulties in engaging in goal-directed behavior, (c) difficulties in controlling impulses, (d) lack of emotional awareness, (e) limited access to emotion regulation strategies, and (f) lack of emotional clarity. Previous studies have demonstrated excellent psychometric properties for the DERS both in Italian (de Campora, Giromini, Larciprete, Volsi, & Zavattini, 2014; Giovannini et al., 2014; Giromini, Brusadelli, Di Noto, Grasso, & Lang, 2015; Giromini, Velotti, De Campora, Bonalume, & Zavattini, 2012) and foreign studies (e.g., Dan-Glauser & Scherer, 2013; Miguel, Giromini, Colombarolli, Zuanazzi, & Zennaro, 2017; Ruganci & Genc€oz, 2010). In our sample, Cronbach's alphas were .86 (nonacceptance), .86 (goals), .87 (impulse), .72 (awareness), .90 (strategies), .88 (clarity), and .95 (total DERS score).

ONLINE DATA SET

In addition to the QCAE, participants in this study were also administered the following tests.

Toronto Alexithymia Scale–20: The Toronto Alexithymia Scale–20 (TAS–20; Bagby, Parker, & Taylor, 1994) is a self-report questionnaire including 20 items, rated on a 5-point Likert scale. In addition to the total TAS–20 score, three subscale scores are typically used: difficulty identifying feelings (F1), difficulty describing feelings (F2), and externally oriented thinking (F3). In the original study by Bagby et al. (1994), the TAS–20 demonstrated acceptable internal consistency both for the total score (Cronbach's a D .81), and for each factor (F1 D .78, F2 D .75, F3 D .66). The Italian version of the TAS–20 (Bressi et al., 1996) also showed Cronbach's alpha values ranging from .52 to .77 in a nonclinical

sample. In our sample, Cronbach's alphas were .85 (difficulty identifying feelings), .79 (difficulty describing feelings), .67 (externally oriented thinking), and .85 (total score).

Reading the Mind in the Eyes-Test: The Reading the Mind in the Eyes-Test (RME–T; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, & Jolliffe, 1997) is used to assess emotion recognition. It includes 36 still pictures of the eye region, and the person has to choose among four emotions that the pictures could represent. A link with the standardized glossary of the RME–T was present in the online administration. The aim of the test is to measure the frequency of matching a semantic definition to its expression in the picture, and the score is calculated by the sum of correct responses. The Italian version of the RME–T was introduced by Vellante et al. (2013), who reported information on internal consistency, factor structure, and test–retest reliability of the Italian adaptation. The results of their study support the reliability of the Italian RME–T, although this instrument has produced low internal consistency indices in other studies (e.g., Olderbak et al., 2015). In our study, Cronbach's alpha was relatively low (.32).

4.2. STATISTICAL ANALYSES

Samples homogeneity. Both samples had a similar percentage of men and women, Phi = .05, p = .18. In both samples about three quarters were women. Conversely, a statistically significant difference emerged when examining the mean age of the two samples: Participants in the online dataset were significantly older than those in the pa-per-and-pencil sample, t (454.7) = 8.70, p < .01, d = .67 and mean ages were 22.6 vs. 26.4, respectively. We checked in the combined sample, whether age correlated with QCAE scores, which it did not: |r| ≤ .068, p ≥ .074. Furthermore, when we performed additional analyses (i.e., ANCOVA's) aimed at controlling for the impact of age on the mean differences between the pa-per-and-pencil and online formats, the results were virtually identical to those

we obtained when the variable age was not controlled for. Likewise, because the paper-and-pencil sample only included university students while the online sample also included non-student participants, we performed additional analyses controlling for this possible confounding factor. After excluding all non-students from the combined dataset, we obtained similar results to those obtained when analyzing all available data. Thus, the comparison between the paper-and-pencil versus online administration was not notably affected by sample composition in terms of age or being a student or not.

Composition of the Samples.

	Paper-and-pencil Dataset	Online Dataset	Combined Dataset
Gender, ϕ = .05, p = .18			
Male	105 (25.9%)	61 (21.4%)	166 (24.0%)
Female	301 (74.1%)	224 (78.6%)	525 (76.0%)
Age, $t(454.7) = 8.70$, $p < .01$, $d = .67^*$			
M	22.55	26.41	24.16
SD	4.61	7.01	6.03

[*] Because homoscedasticity could not be assumed, the Welch–Satterthwaite method was used to adjust degrees of freedom.

Reliability and Validity Analyses. For both paper-and-pencil and online datasets, we examined internal consistency and construct validity of QCAE scales. More in detail, QCAE scores' reliability was inspected via examination of Cronbach's alpha and item-scale correlations. Construct validity was tested by performing a confirmatory factor analysis (CFA) and by correlating QCAE scores to empathy-related constructs, such as interpersonal competence and personality traits such as agreeableness and openness (convergent validity). The comparison between paper-and-pencil vs. online QCAE scores was performed via t-test statistics, after testing CFA measurement invariance between the two

formats. For both correlational and t-test analyses, Holm-Bonferroni correction (Holm, 1979) was applied to correct for multiple testing.

As for the convergent validity, based on the previous literature on empathy and on our theoretical considerations, we expected that the Italian QCAE would correlate positively with psychological well-being, extraversion, openness, and agreeableness, but negatively with neuroticism and emotion dysregulation (Henry, Bailey, & Rendell, 2008). Indeed, extraverted individuals tend to be well-disposed and comfortable in human interactions (Costa & McCrae 1992). Conversely, neuroticism, anxiety, and depression probably decrease openness to social interactions, and the skills associated with facilitating them (Riemann & Allgöwer 1993).

4.3. RESULTS

Internal Consistency. Internal consistency of QCAE scores was estimated for both samples separately and for the combined sample. Within the paper-and-pencil sample, internal consistency was adequate, with Cronbach alpha's ranging from .58 (Proximal Responsivity and Peripheral Responsivity) to .87 (Perspective Taking) for the subcomponents, and ≥ .77 for the Cognitive Empathy and Affective Empathy subscales and the Total Score. Similarly, within the online dataset, Cronbach alpha's ranged from .69 (Peripheral Responsivity) to .84 (Perspective Taking) for the subscales, and was ≥ .81 for the two subscales and total score. For Proximal Responsivity and Peripheral Responsivity, in the paper-and-pencil dataset Cronbach alpha's were .64 and .58, respectively; while in the online datasets, Cronbach alpha's were .69 for both subscales.

Internal Consistency of QCAE.

QCAE Scale	No. of items	Paper-and-pencil Dataset		Online Dataset		Combined Dataset	
		α	item-total correlations	α	item-total correlations	α	item-total correlations
PT	10	.87	.62 – .71	.84	.54 – .72	.86	.60 – .71
OS	9	.83	.52 – .75	.78	.50 – .69	.82	.51 – .72
EC	4	.73	.69 – .79	.76	.69 – .81	.74	.69 – .80
PrR	4	.64	.57 – .77	.69	.60 – .83	.67	.59 – .80
PeR	4	.58	.45 – .80	.69	.48 – .85	.63	.47 – .82
CE	19	.89	.44 – .68	.84	.33 – .65	.87	.40 – .64
AE	12	.77	.26 – .66	.81	.31 – .69	.79	.29 – .67
Total	31	.87	.26 – .60	.86	.31 – .61	.87	.29 – .60

PT = Perspective Taking; OS = Online Simulation; EC = Emotion Contagion; PrR = Proximal Responsivity; PeR = Peripheral Responsivity; CE = Cognitive Empathy; AE = Affective Empathy.

Factor Structure. To factor analyze our QCAE data, we used Lisrel 8.72 (Jöreskog & Sörbom, 2005). Because these analyses aimed at testing whether the factor structure identified by Reniers et al. (2011) would also fit our Italian data, the same methodological approach utilized by Reniers et al. (2011) was used in this study, too. That is, we specified five latent variables (the five scales of the QCAE) and used the same item parcels utilized by the authors (for using item parceling rather than individual items in CFA, see Hall, Snell, & Foust, 1999; Little, Cunningham, Shahar, & Widaman, 2002; Nasser & Wisenbaker, 2003). Then, the same two models proposed – and tested via CFA – by Reniers et al. (2011) were tested. More specifically, in model 1 (M1), the five latent variables (i.e., the five QCAE subscales) were allowed to correlate with each other; in model 2 (M2), a hierarchical structure was tested, with Cognitive and Affective Empathy serving as second order factors. Additionally – and differently from Reniers et al. (2011) – our study also tested a unidimensional model (UM) to provide us with a baseline referent model, to better evaluate M1 and M2.

Goodness of Fit Indices for a Univariate Model, and for Models 1 and 2.

	Paper-and-pencil Dataset			Online Dataset			Combined Dataset		
	UM	M 1	M 2	UM	M 1	M 2	UM	M 1	M 2
χ^2	1400.55	257.67	292.85	1192.13	268.99	288.89	2489.13	419.85	477.78
df	90	80	85	90	80	85	90	80	85
χ^2 p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
χ^2/df	15.56	3.22	3.45	13.26	3.36	3.40	27.66	5.25	5.62
RMSEA	.19	.07	.08	.21	.09	.09	.20	.08	.08
RMSEA 90% CI	.18–.20	.06–.08	.07–.09	.20–.22	.08–.10	.08–.10	.19–.20	.07–.09	.08–.09
SRMR	.13	.07	.08	.15	.08	.09	.13	.07	.08
CFI	.78	.96	.95	.70	.93	.92	.76	.95	.95
NNFI	.75	.94	.94	.65	.91	.91	.72	.94	.94
AIC	1460.55	337.67	362.26	1252.13	348.99	358.89	2549.13	499.85	547.78

UM = Univariate model; M1 = Model 1; M2 = Model 2; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = compared fit index; NNFI = non-normed fit index; AIC = Akaike's Information Criterion.
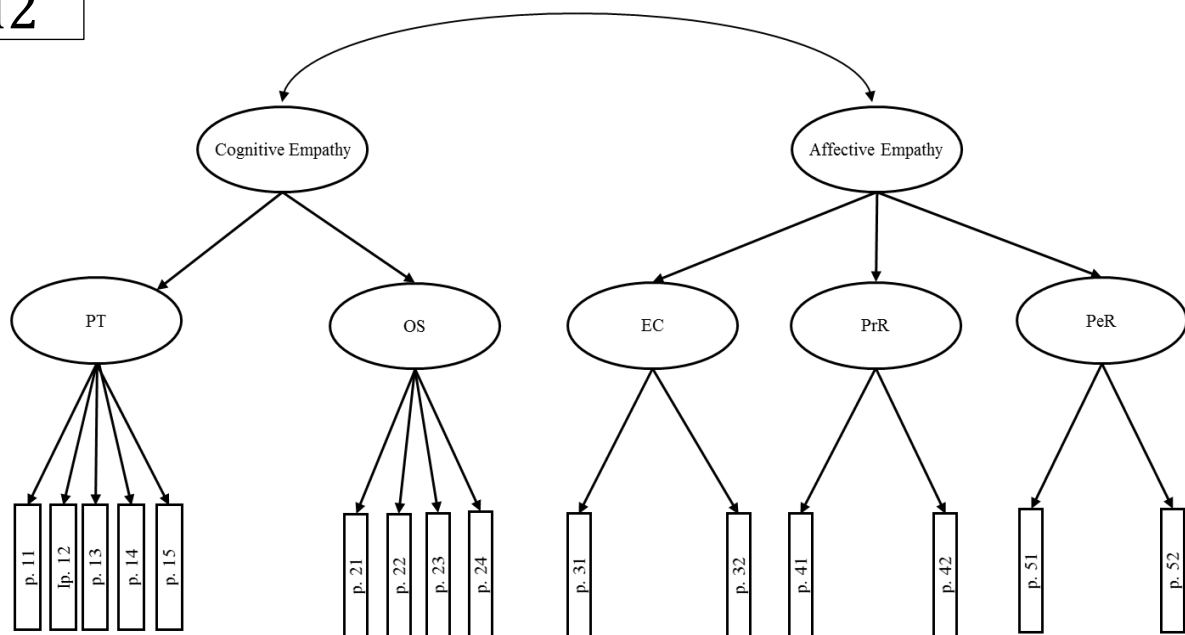
PT = Perspective Taking; OS = Online Simulation; EC = Emotion Contagion; PrR = Proximal Responsivity; PeR = Peripheral Re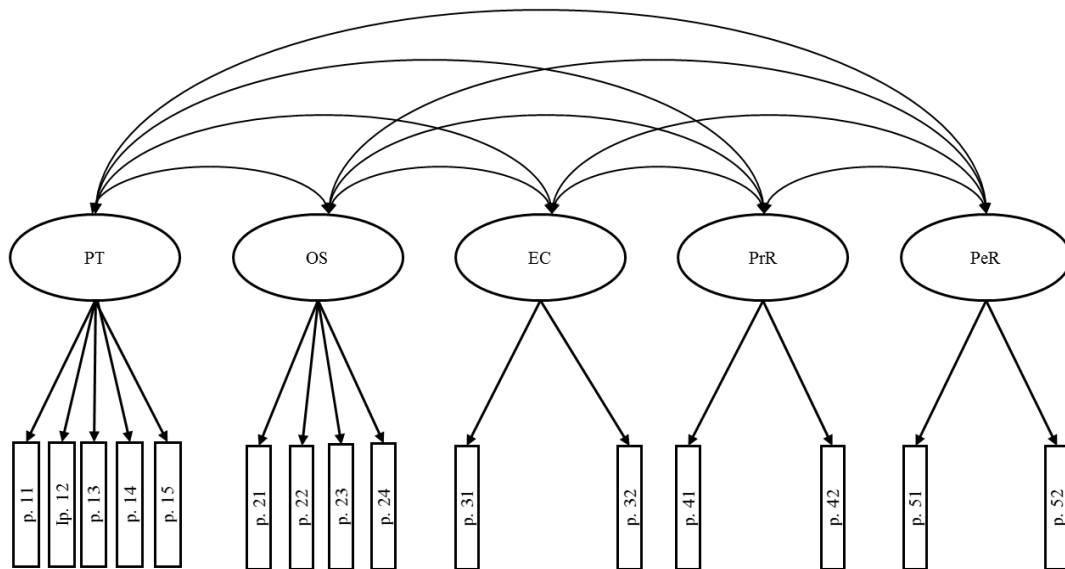sponsivity; CE = Cognitive Empathy; AE = Affective Empathy. The labels "P 11, P 12, …, P 52" refer to the same item parcels utilized by Reniers et al. (2011).

The following goodness of fit statistics were taken under consideration, for all these three models. First, we looked at the χ^2, its associated p-value, and, most importantly, at the ratio between the χ^2 and its degrees of freedom (χ^2/df). According to Watkins (1989), a χ^2/df close to 2 reflects a good fit, and values lower than 5 indicate a quite promising fit. Next, we inspected the root mean square error of approximation (RMSEA) and its 90% confidence interval. Based on Browne and Cudeck (1993), we considered RMSEA values close to .05 to indicate a close fit, values close to .08 to indicate a fair fit, and values close to .10 to indicate a marginal fit. We then inspected the standardized root mean square residual (SRMR), whose values were expected to be close to or lower than .08 to indicate of a good fit (Hu & Bentler, 1980). Moreover, the comparative fit index (CFI) and non-normed fit index (NNFI) were inspected too, with their values being expected to be .90 or higher to indicate a good fit (Bentler & Bonett, 1980). Lastly, we also considered the Akaike's Information Criterion (AIC), whose values may be used to compare different models, as the lower the AIC, the better the fit of the model (Akaike, 1973).

Based on the criteria described above, the UM did not provide an adequate fit, and therefore it was discharged. Conversely, both M1 and M2 fit relatively well our data. For example, both M1 and M2 produced RMSEA values below .10, SRMR values close to or lower than .08, and CFI and NNFI values above .90, in all samples under consideration. We thus compared M1 versus M2 by using the χ^2 test (i.e., by testing the difference between the two χ^2 values) and by examining their AICs.

Comparison between Model 1 and Model 2.

	Paper-and-pencil Dataset	Online Dataset	Combined Dataset
$\Delta(M2 - M1)\,\chi^2$	35.18	19.9	57.93
$\Delta(M2 - M1)\,df$	5	5	5
p-value $\Delta(M2 - M1)\,\chi^2$	< 0.001	< 0.001	< 0.001
$\Delta(M2 - M1)\,AIC$	24.59	9.90	47.93

M1 = Model 1; M2 = Model 2.

The results of these additional analyses, indicate that M1 provided a significantly better fit than did M2, $\chi^2 \geq 19.9$, $p < .001$, and produced notably lower AIC values. Accordingly, it was concluded that M1 offered the best fit for our data.

Factor Loadings for Reniers et al.'s (2011) QCAE Parcels, as Obtained from our CFA M1

Variable	Paper-and-pencil Dataset	Online Dataset	Combined Dataset
Perspective Taking			
P 11	.73	.80	.76
P 12	.77	.75	.77
P 13	.77	.69	.74
P 14	.78	.69	.76
P 15	.75	.74	.75
Online Simulation			
P 21	.72	.72	.72
P 22	.80	.80	.79
P 23	.77	.72	.76
P 24	.76	.62	.71
Emotion Contagion			
P 31	.71	.68	.69
P 32	.71	.78	.75
Proximal Responsivity			
P 41	.66	.68	.68
P 42	.69	.71	.70
Peripheral Responsivity			
P 51	.59	.71	.62
P 52	.72	.83	.79

Note. The labels "P 11, P 12, ..., P 52" refer to the same item parcels utilized by Reniers et al. (2011)

Convergent validity. Convergent validity with the ICQ, NEO-FFI, PGWBI, and DERS was calculated for the paper-and-pencil sample, while convergent validity with the TAS-20 and RME-T was calculated for the online sample. Below we discuss correlations that were statistically significant after Holm-Bonferroni correction (Holm, 1979).

Convergent Validity Analyses.

	PT	OS	EC	PrR	PeR	CE	AE	Tot
Paper-and-pencil Dataset								
ICQ (*n* = 407)								
IR	.35**	.15	-.10	.17	.03	.30**	.04	.24**
ES	.52**	.45**	.01	.45**	.19*	.57**	.27**	.56**
NA	.41**	.17	-.12	.15	.01	.35**	.01	.26**
DC	.26**	.13	.03	.18*	.12	.23**	.15	.24**
CM	.26**	.43**	.03	.25**	.03	.40**	.13	.36**
NEO-FFI (*n* = 407)								
Neuroticism	-.25**	-.22**	.41**	.10	.21**	-.27**	.32**	-.05
Extraversion	.29**	.14	-.06	.19*	.07	.26**	.08	.23**
Openness	.31**	.19*	-.05	.25**	.22**	.30**	.17	.31**
Agreeableness	-.02	.31**	.07	.24**	.11	.16	.18*	.21**
Conscientiousness	.29**	.29**	.01	.15	.09	.33**	.11	.30**
PGWBI (*n* = 407)								
AA	.08	.11	-.23**	-.04	-.15	.11	-.19*	-.01
AD	.16	.16	-.14	.04	-.07	.19*	-.08	.10
PWB	.11	.09	-.19*	.01	-.07	.12	-.11	.03
SC	.17	.18*	-.19*	.02	-.07	.20**	-.11	.10
GH	.16	.13	-.11	.03	-.08	.17	-.07	.09
VIT	.08	.09	-.21**	-.07	-.14	.10	-.18*	-.02
Total	.15	.15	-.23**	-.01	-.13	.17	-.17	.05
DERS (*n* = 407)								
Nonacceptance	-.16	-.16	.24**	.01	.04	-.19*	.13	-.08
Goals	-.04	-.09	.25**	.10	.08	-.07	.19*	.04
Impulse	-.23**	-.32**	.23**	-.05	.06	-.32**	.11	-.19*
Awareness	-.34**	-.26**	-.06	-.29**	-.19*	-.35**	-.23**	-.37**
Strategies	-.21**	-.25**	.29**	-.02	.08	-.26**	.16	-.12
Clarity	-.30**	-.22**	.20**	-.04	.01	-.30**	.08	-.19*
Total	-.28**	-.30**	.28**	-.05	.03	-.34**	.12	-.19*
Online Dataset								
TAS-20 (*n* = 285)								
DIF	-.16	-.15	.29**	.03	.00	-.18	.14	-.04
DDF	-.27**	-.13	.06	-.17	-.14	-.25**	-.10	-.22*
EOT	-.33**	-.38**	-.05	-.31**	-.27**	-.43**	-.26**	-.43**
Total	-.31**	-.27**	.15	-.17	-.16	-.35**	-.06	-.27**
RME-T (*n* = 282)								
Total	.16	.13	.00	.08	.08	.18	.07	.16

PT = Perspective Taking; OS = Online Simulation; EC = Emotion Contagion; PrR = Proximal Responsivity; PeR = Peripheral Responsivity; CE = Cognitive Empathy; AE = Affective Empathy. ICQ = Interpersonal Competence Questionnaire; IR = Initiation Relationship; ES = Emotional Support; NA = Negative Assertion; DC = Disclosure; CM = Conflict Management; NEOFFI = NEO Five-Factor Inventory; PGWBI =

Psychological General Well-Being Index: AA = Absence of Anxiety; AD = Absence of Depression; PWB = Positive well-being; SC = Self-control; GH = General Health; VIT = Vitality; DERS = Difficulties in Emotion Regulation Scale; TAS-20 = The Toronto Alexithymia Scale-20: DIF = Difficulties Identifying Feelings; DDF = Difficulties Describing Feelings; EOT = Externally Oriented Thinking; RME-T = Reading the Mind in the Eyes-Test.
* Significant at $\alpha \leq .05$ after Holm-Bonferroni correction significance testing;
** Significant at $\alpha \leq .01$ after Holm-Bonferroni correction significance testing.

As for the correlations of QCAE to ICQ, it is interesting to note that Cognitive Empathy and the Total QCAE scores correlated positively with all ICQ scales, $r \geq .23$. Moreover, Affective Empathy correlated with the ICQ Emotional Support scale only, $r = .27$. Of all QCAE scales, Emotion Contagion was the only one that did not correlate with any of the ICQ scales.

Similarly, the correlations between the QCAE and NEO-FFI revealed a different pattern for Cognitive versus Affective Empathy. Cognitive Empathy correlated positively with Extraversion ($r = .26$), Openness ($r = .30$), and Conscientiousness ($r = .33$), and negatively with Neuroticism ($r = -.27$). Furthermore, Affective Empathy correlated only with Neuroticism ($r = .32$). Noteworthy, the effect size of the relationship between Emotion Contagion and Neuroticism was medium to large, i.e., $r = -.41$. All other correlations between the QCAE and NEO FFI consisted, at maximum, of a medium effect sizes.

The total well-being score (PGWBI Total) did not correlate with Total QCAE score, but correlated positively with Cognitive Empathy ($r = .17$), and negatively with Affective Empathy ($r = -.17$). Because of the two correlations are exactly the opposite, they cancel each other out in the final correlation between the Total QCAE score and the PGWBI Total. Again, when looking at the QCAE subcomponents, Emotion Contagion produced the strongest correlation with the total well-being score, $r = -.23$.

The Total QCAE score produced significant correlations with the total DERS ($r = -.19$) and total TAS-20 ($r = -.27$) scores. However, while Cognitive Empathy correlated $r = -.34$ with the total DERS score and $r = -.35$ with the total TAS-20 score, Affective Empathy did not correlate with these two. Emotion

Contagion was the only one that produced positive correlations with DERS and TAS-20 scales and/or subscales.

Finally, the QCAE did not produce statistically significant correlations with the RME-T. Comparison between QCAE scores from Paper-and-Pencil and Online administrations. Prior to comparing QCAE scores from the paper-and-pencil versus the online administration, we tested CFA measurement invariance between the two formats. Because M1 provided the best fit our data, M1 only was analyzed for structural invariance. These analyses were performed across four steps, in line with previous research in the field assessment (e.g., Beaujean, Freeman, Youngstrom & Carlson, 2012). That is, first, configural invariance assessed if the factor model was invariant across the two groups/formats. Next, metric invariance investigated if the factor loadings for QCAE parcels were the same in both groups/formats. Third, scalar invariance was tested by constraining all the scales' origins (i.e., intercepts) across the two groups/formats. Finally, invariant unique variance analyzed the invariance of the unique residual variances across the two groups/formats. The results of these analyses suggest that the paper-and-pencil and online formats were structurally invariant. Indeed, all models were adequate, and no notable differences from one step to another were observed.

Structural Invariance of the QCAE (M1) between the Paper and Pencil and Online Administrations

Model 1	χ^2	df	p	$\Delta\chi^2$	Δdf	p	AIC	RMSEA (90% CI)	SRMR	CFI	NNFI
1 Configural Invariance	526.7	160	<.001	–	–	–	686.7	0.08 (0.07-0.09)	0.08	0.95	0.93
2 Metric Invariance	538.1	170	<.001	11.4	10	0.33	678.1	0.08 (0.07-0.09)	0.08	0.95	0.93
3 Scalar Invariance	538.1	180	<.001	0.0	10	1.00	718.1	0.08 (0.07-0.08)	0.08	0.95	0.94

	4 Invariant Unique Variance										
	565.7	195	<.001	27.6	15	0.02	715.7	0.07 (0.07-0.08)	0.08	0.95	0.94

AIC = Akaike's Information Criterion; RMSEA = root mean square error of approximation; CI = Confidence Interval; SRMR = standardized root mean square residual; CFI = compared fit index; NNFI = non-normed fit index.

Accordingly, we next performed a series of t-tests to compare the scores produced by the two formats. The online version produced statistically significantly higher QCAE scores than the paper-and-pencil version, with a small or small to medium effect size (Cohen, 1988). This difference was more evident with the female than with the male samples. Also noteworthy, although Reniers et al. (2011) did not report detailed, descriptive statistics concerning their samples' QCAE scores, they did report the average scores (and relative standard errors) of men and women on the Cognitive Empathy and Affective Empathy QCAE scales. More specifically, men had a mean of 56.1 (SE = .5) on Cognitive Empathy and a mean of 32.3 (SE = .3) on Affective Empathy, and women had a mean of 59.4 (SE = .3) on Cognitive Empathy and a mean of 36.8 (SE = .2) on Affective Empathy. These values – which were obtained by Reniers et al. (2011) via online administration – are markedly similar, nearly identical to those observed in this study, when considering the online sample data.

Comparison between QCAE Scores from Paper-and-Pencil and Online Administrations

	Paper-and-pencil Dataset		Online Dataset		*t*	*df*	*Uncorr. p*	*d*
	M	*SD*	*M*	*SD*				
PT								
Men	28.0	5.1	29.7	4.6	-2.12	164	.04	-.34
Women	30.0	5.0	31.4	4.7	-3.21	522	<.01*	-.28
Tot	29.5	5.1	31.0	4.7	-3.99	690	<.01**	-.31
OS								
Men	26.0	4.4	27.1	3.8	-1.60	164	.11	-.26
Women	27.3	4.8	28.0	4.2	-1.88	552	.06	-.17
Tot	27.0	4.7	27.8	4.2	-2.53	690	.01	-.20
EC								
Men	9.7	2.6	9.4	2.5	.83	164	.41	.13

	M	SD	M	SD	t	df	p	d
Women	10.9	2.4	11.4	2.7	-2.53	522	.01	-.22
Tot	10.6	2.5	11.0	2.8	-2.16	690	.03	-.17
PrR								
Men	10.6	2.3	11.1	2.3	-1.35	164	.18	-.22
Women	12.3	2.1	12.8	2.2	-2.84	522	<.01	-.25
Tot	11.9	2.3	12.5	2.3	-3.48	690	<.01*	-.27
PeR								
Men	10.3	2.5	9.8	2.4	1.12	164	.26	.18
Women	12.0	2.2	12.4	2.5	-1.86	436.5[a]	.06	-.17
Tot	11.5	2.4	11.8	2.7	-1.54	562.1[a]	.12	-.12
CE								
Men	54.1	8.3	56.8	6.2	-2.27	164	.03	-.36
Women	57.3	8.3	59.4	7.5	-3.03	522	<.01*	-.27
Tot	56.5	8.4	58.9	7.3	-3.90	690	<.01**	-.30
AE								
Men	30.6	5.4	30.3	5.1	.33	164	.74	.05
Women	35.2	4.9	36.6	5.6	-3.17	444.7[a]	<.01*	-.29
Tot	33.9	5.5	35.3	6.1	-3.02	567.8[a]	<.01*	-.24
Total								
Men	84.7	10.1	87.2	8.7	-1.61	164	.11	-.26
Women	92.4	11.0	96.1	10.7	-3.79	522	<.01**	-.33
Tot	90.4	11.2	94.2	10.9	-4.38	690	<.01**	-.34

[a] Because homoscedasticity could not be assumed, the Welch–Satterthwaite method was used to adjust degrees of freedom

* Significant at $\alpha \leq .05$ after Holm-Bonferroni correction significance testing.

** Significant at $\alpha \leq .01$ after Holm-Bonferroni correction significance testing.

4.4. DISCUSSION

Over the last 20 years, the empathy construct has been refined, notably the distinction between cognitive and affective components of empathy. In line with these refinements, Reniers et al. (2011) developed the Questionnaire of Cognitive and Affective Empathy (QCAE), a 31-item self-report measure of cognitive and affective empathy. The main purpose of our study was to examine internal consistency, factor structure, and convergent validity of an Italian version of the QCAE. Furthermore, we also compared the average scores obtained by a paper-and-pencil version with online administration of the QCAE.

In terms of internal consistency, all QCAE scales from both paper-and-pencil and online versions produced Cronbach's alpha values above .70, except for Proximal Responsivity and Peripheral Responsivity, which produced Cronbach's alpha scores between .58 and .69. Thus, both paper-and-pencil and online formats produced similar internal consistency results, and these results are comparable to those reported by Reniers et al. (2011). It is noteworthy that in Reniers et al.'s (2011) study Proximal Responsivity and Peripheral Responsivity also demonstrated the lowest Cronbach's alpha values of all QCAE subscales (.70 and .65, respectively). Although many statisticians criticize the idea that Cronbach's alpha values below .70 reflect lack of reliability for the scales under investigation (e.g., John & Soto, 2007; Sijtsma, 2009), future studies should pay particular attention to the reliability of scores from these two subscales. Perhaps, a tentative explanation for these relatively low reliability indices may be that these two sub-components, along with Emotion Contagion, are the ones with the lowest number of items (i.e., 4 each).

The results of our confirmatory factor analysis suggest that both data from the paper-and-pencil and those from the online dataset fit the models proposed by Reniers et al. (2011) relatively well, and certainly better than did the unidimensional model. Our study thus suggests that the QCAE shows structural validity and factorial stability across different Western cultures and languages, regardless of administration format. On the other hand, since the QCAE has only been studied in Western populations, additional validation research in non-Western samples is necessary. Furthermore, and perhaps more importantly, because our goal was to test whether the model(s) proposed by Reniers et al. (2011) would also fit our Italian data, we decided to use the same methodological approach that they used in their study. That is, we decided to perform our CFAs on item parcels rather than on individual items. On one hand, this approach ensured that any potential discrepancies in the CFA results of our vs. Reniers et al.'s (2011) studies could not be due to the analytic strategies being different from one study to another. On the other hand, however, because the QCAE items are scored on a 4-point Likert scale, future studies performing CFAs on all items (e.g.,

by using an appropriate estimator with robust standard errors) would probably be beneficial.

Our convergent validity analyses revealed some interesting patterns. Cognitive Empathy correlated positively with interpersonal competence (ICQ), amiable personality traits of the NEO-FFI (i.e., extraversion, openness, and conscientiousness), and psychological well-being (PGWBI), and negatively with alexithymia (TAS-20), neuroticism (NEO-FFI), and difficulties with emotion regulation (DERS). Those correlations were anticipated, as they are largely in line with recent findings that associate dysfunction in social functioning with risk of a variety of psychopathological conditions, such as psychosis (Henry, Bailey, & Rendell, 2008).

Conversely, Affective Empathy produced a more complex pattern of correlations, which cannot be understood without looking at the correlations produced by its subscales: While Proximal Responsivity correlated positively with interpersonal competence and amiable personality traits (i.e., extraversion, openness, and agreeableness), Emotion Contagion – and to a lesser extent Peripheral Responsivity – correlated negatively with well-being (PGWBI), and positively with alexithymia (TAS-20), neuroticism (NEO-FFI), and difficulties in emotion regulation (DERS). Given the negative correlations between QCAE Affective Empathy and well-being, considering affective empathy as a resource does not appear to be a foregone conclusion. In fact, high levels of emotion contagion had recently been associated with some pathological conditions. Weisbuch et al. (2011) found an increased risk for eating disorders in young women susceptible to emotion contagion. Also high levels of emotion contagion were found in patients with schizophrenia (Horan et al., 2015) and with difficulties in emotional regulation (Miguel et al., 2016). Combined with previous literature, our findings thus suggest that emotion contagion could be associated with psychological vulnerability.

Taken together, our convergent validity results also support the cross-cultural applicability of the QCAE. Indeed, the QCAE total score produced positive and statistically significant correlations with instruments measuring constructs

related to empathy (i.e., interpersonal competence and amiable personality traits) and negative correlations with difficulties in recognition and regulation of emotions (i.e., alexithymia and emotional dysregulation). Conversely, the correlation of the QCAE to emotion recognition was nonsignificant. Given that emotion recognition and empathy are only partially overlapping constructs, the relatively weak correlation between QCAE and RME-T is not unexpected, but rather suggests that emotion recognition is probably necessary, but not sufficient to empathize with others. Furthermore, it should be noted that while the QCAE is a self-report measure, the RME-T is rather a performance-based instrument. As such, it is not too surprising that the two instruments do not correlate strongly with each other (Mihura et al., 2013).

One of the most interesting results of our study, in our opinion, is that when compared to the standard, paper-and-pencil format, the online administration format produced significantly higher QCAE scores. Based on our post-hoc analyses controlling for age and student status (i.e., being a student vs. not being a student), it is unlikely that these differences may be accounted for simply by demographic heterogeneity across the two samples. Perhaps, a better explanation for these findings may be ascribed to self-selection bias (e.g., participants had not been personally invited by the research assistant to volunteer for the study) and under-coverage in online surveys (Bethlehem, 2010). Said differently, it is possible that those who decided by themselves to volunteer in the online group were more interested in knowing about emotions and empathy compared to the paper-and-pencil sample (who was explicitly asked to volunteer by an assistant) – a characteristic, that is typical of empathetic individuals. Future QCAE research might further inspect whether online administrations produce higher scores than paper-and-pencil format.

Although our findings provide initial support for the cross-cultural applicability of the QCAE, some of our study's limitations deserve mentioning. First, our two samples are far from being representative of the general Italian population, and some demographic and sample size differences between the paper-and-pencil and online samples make it difficult to rule out that the two

samples scored differently on the QCAE for some uncontrolled reasons. For these reasons, our findings still need to be further replicated with other samples too. Second, but somehow related to this first point, future studies should attempt to control for many other variables that we could not control for in our study, such as socioeconomic status, marital status, etc. Third, one of the instruments we used to test convergent validity, the RME-T, had a very low internal reliability. As such, the generalizability of its results to other studies is difficult to evaluate. Fourth, we did not examine divergent validity or test-retest stability, which are important to better estimate the validity and reliability of our QCAE scores.

Despite these limitations, our study is the first to investigate the reliability and validity of the QCAE in Italy, and to compare QCAE scores obtained with paper-and-pencil versus online administration formats. We found evidence that the Italian version of the QCAE has sound psychometric properties. We showed the QCAE had adequate internal reliability, factorial stability and convergent validity. As such, the instrument holds promise as an easy to administer self-report tool for the assessment of the cognitive and affective components of empathy.

V. RORSCHACH EMPATHY INDEX (DEVELOPMENT AND VALIDATION)

Given that the lack of empathy is a key feature of several psychopathological conditions, the construct of empathy has become more and more important in the literature, and so has its assessment. Moreover, as it was discussed above, the most widely utilized personality inventories (i.e. MMPI-II-RF and PAI) do not investigate this trait specifically.

Furthermore, R-PAS assesses the respondent's behavior, and this could be helpful for the investigation of empathy. More precisely, empathic people are aware of the impact that their behavior has on other people. By consequence, someone who is lacking on this capability should not be aware, or concerned, about the role that his/her behavior has in relationships. Given that the confidence of the information derived from self-reported questionnaires is closely connected with the subject's awareness and honesty (Meyer, 2017a), using this type of instruments could not be sufficient to assess empathy, as they obviously are at risk for social desirability.

On the other side, having a specific empathy index in a personality assessment test, could give information about this capability without necessary administer others instruments. Surely, the Rorschach is not the fastest test for the assessment of empathy, and not the most specific one. Nevertheless, it would be helpful for a clinician to have a first idea about what the empathetic attitudes of the respondent could be, to decide if it is necessary to explore them deeper.

Finally, it would be interesting to investigate empathy capabilities independently from the remaining NPD or APD features. Usually, items that investigate lacking of empathy are incorporated in indices that also comprise others Narcissistic's (e.g. grandiosity, haughty attitudes, need of admiration, envy) or Antisocial's (e.g. violation of rules, irritability, irresponsibility) behaviors, which not necessarily are related to empathy. R-PAS does not assesses

personality through patterns of symptoms, but rather through the respondent's ways of functioning and approaching in the daily life. Thus, R-PAS could give a more representative frame about the role played by empathy in the personality functioning.

5.1. RORSCHACH EMPATHY SCALE (RES)

Despite the presence of variables related to interpersonal abilities and to the construct of empathy, R-PAS codes do not investigate empathy directly, nor do they do so by encompassing its whole complexity. Thus, we tried to create a new pool of variables in the effort to gather more empathy aspects. The group of new variables was named Rorschach Empathy Scale (RES).

5.1.1. DEVELOPMENT

The RES variables came from four paths: the QCAE items, theoretical hypotheses, administration experience, and some suggestions offered by the R-PAS manual. Below, are summarized these efforts, by presenting all RES items investigated in this research project.

The first version of the RES included 18 types of behaviors (RES-18).

1. The respondent verbally expresses an emotional arousal due to the exposure to the card (e.g., *"This is sad"*; *"This scares me"*; *"Oh, this is so cheerful!"*; *"Ahhh, a skinny poor small b.f."*)	YES (1) NO (0)
2. The object of the response (e.g., a person, an animal, an object) is described as feeling or expressing emotions or feelings (e.g., *"A sad face"*; *"two women who are washing clothes… they seem happy!"*; *"An angry dog!"*)	YES (1) NO (0)
A. The emotional or mental state of the object (e.g., a person, an animal, an object) is incongruent with the context	YES (1) NO (0)

	(e.g., "*A person dressed in black, like at a funeral, he is smiling*"; "*A party… but everyone seems sad*")	
	B. The respondent is confused about what emotion an object (e.g., a person, an animal, an object) experiences or expresses (e.g.,"*I can't understand if he's happy or angry*", "*She seems to be smiling …or screaming*")	YES (1) NO (0)
	C. The respondent is unable to explain the reasons why an object of the response experiences an emotion (e.g., "*I don't know why but I think he's sad*"; "*She seems happy, even if I don't understand why*")	YES (1) NO (0)
	D. The respondent attributes different behaviors and/or emotional or mental states to two objects that are perceptually identical based on the symmetry of the blot (e.g. "*Two men sitting at a table. The one on the left is happy, the right one is sad.*")	YES (1) NO (0)
3.	The response includes behaviors, actions, emotions or feelings that are about to happen or be experienced (e.g., "*It's about to catch this thing in the middle*"; "*These two ladies have a luggage, they're leaving, but they're gonna miss home*")	YES (1) NO (0)
4.	The object of the response (e.g., a person, an animal, an object) is described as having intentions, wishes, and/or desires (e.g., "*Here is the kid, he would like to have some toys to play with*"; "*That's the lion, it'd like to catch some pray*")	YES (1) NO (0)
5.	The object of the response (e.g., a person, an animal, an object) is described as being waiting for something, or thinking of someone or something (e.g., "*A person who is waiting for the bus*"; "*A guy who is thinking of his wife*")	YES (1) NO (0)
6.	The respondent emphasizes his/her own perspective, ignoring that different people might see different things (e.g., "*This is obviously a butterfly!*"; "*I'm a musician, so I see it perfectly: I'm sure that this is a guitar!*"; "*Clearly a sword!*")	YES (1) NO (0)
7.	The words used by the respondent in AGM or MAP responses suggest a positive mood or emotion (pleasure, satisfaction, cheerfulness).	TRUE (1) FALSE (0) N.A. (999)
8.	The words used by the respondent in AGM or MAP responses suggest a negative mood or emotion (sadness, nuisance, contempt, disgust, angriness).	TRUE (1) FALSE (0) N.A. (999)

9. Emotional valence of the response (e.g., "*A butterfly with broken wings*"; "*An evil creature with the horns*" = negative; "*A butterfly*"; "*A person bending*" = neutral; "*A smiling face!*"; "*Two bears high-fiving e.o.*" = positive)	NEGATIVE (-1) NEUTRAL (0) POSITIVE (1)
10. During the CP, how many questions does the examiner need to make in order to clarify the response? (e.g., ERR: R: "There are the wings and there is the head, it's backward!" E: "What makes it look backwards?" R: "*It's little… instead of the wings, that are bigger*!" →Number of questions = 1)	N. of Questions
11. During the CP, the respondent does not accurately clarify his/her response, so that the examiner needs to make multiple questions about the same feature/component in order to obtain the needed information. (e.g., R: "*This looks soft!*" E: "*What makes it look soft?*" R: "*It seems like this…*" E: "*What in the inkblot makes it look soft*? N. = 2)	N. of Questions
12. The respondent expresses a concern to help the examiner beyond of what was asked in the instructions. (e.g., R: "Am I going too fast? Do I need to slow down?" R: "Is it too much for you to type?" R: "Do you need me to spell it for you?")	YES (1) NO (0)
13. The respondent resists to cooperate with the examiner (e.g., R: "*Yes, it's right there…can't you see it?*"; R: "*Well… it's not so difficult to understand!*" ERR.; R: "*correct.*").	YES (1) NO (0)
14. The respondent sees a human face (Hf) in the response	YES (1) NO (0)
15. The respondent sees gesture of compassion, help, support or attempts to save a living being (e.g., "*These are two arms that are holding this other arm*", "*This is a person who is trying to help this one*")	YES (1) NO (0)

As noted above, we tried to convert the QCAE items in possible Rorschach behaviors. Obviously, many QCAE items overlap, especially those that are in the same scale, because they assess similar attitudes. Thus, we did not convert each

item in a single Rorschach behavior. Moreover, some QCAE items describe more complex behaviors than how it is practicable to detect in the Rorschach task (e.g. item 20, *I am quick to spot when someone in a group is feeling awkward or uncomfortable*), thus it was not possible to convert all the items. What we did, was trying to recognize the basic empathy feature of each item and supposing how this capability could emerge in the Rorschach test context.

RES 1: the behavior that we tried to identify with this variable was a process similar to the emotion contagion, and an emotional sensitivity to the environment. QCAE items that had inspired this variable were the 2, 7, 8, 9, 10, 11, 12, 13, 14, and 29. These items are the most part of the QCAE Affective Empathy scale, with the exception of the item 17 and 23: the first was included in another variable, and the second is related to kinds of situation that are difficult to adapt to the Rorschach request.

RES 2 (a, b, c, d): in a Rorschach-task point of view, decoding the inkblot as an expression, or a mood, could suggest an attitude to pay attention to other's feelings. QCAE items that contributed to the RES 2 were the 19, 20, 21, 22, 24, and 26. In the QCAE, those items represent the big part of the Perspective Taking scale. However, for this variable, we kept just those items related to the decoding of other's expressions. Moreover, we identified four further codes related to the emotional or mental state of the object in the card. RES 2a, b, and c were an attempt to detect inconsistency between the object emotional condition and the context described in the card. The three subitems were supposed to detect a failure into relate emotional condition of people to the context (something similar to the QCAE 17, but not quite the same). The RES 2d, came from administration practice; sometimes the respondent sees different expressions or moods in symmetric details of the card (e.g. CARD VII, "*Two women, this one on the left is sad*"). Considering that – despite the symmetry – the two sides of the cards are often not exactly the same, this could represent a high attitude to attention to details related to others emotional state.

RES 3, 4, and 5: these three variables were referred to a behavior quite close to the cognitive empathy inference process. These items were created on

two basis. Starting from the QCAE, we found some items of the Perspective Taking scale that assessed the capability to anticipate other's intentions or desires (i.e. items 15, 16, 18, 25, 27), not only related to the emotional state. In addition, it is well known that the movement (M, FM, m) is considered a complex elaboration of the stimulus, because the stimulus does not move (Rorschach, 1921; Exner, 1993). In particular, the Human Movement (M) is considered the most important variable to detect empathy skills, because of its relation with the Mirror Neuron System activity (Giromini, Porcelli, Viglione, Parolin, & Pineda, 2010; Pineda, Giromini, Porcelli, Parolin, & Viglione, 2011; Andò, Salatino, Giromini, Ricci, Pignolo, et al., 2015; Porcelli, & Kleiger, 2016; Giromini, Viglione, Brusadel-li, Zennaro, Di Girolamo, et al. 2016; Giromini, Viglione, Pineda, Porcelli, Hubbard, et al. 2017). However, the M code does not take into account if a movement is present or is anticipate, namely in both cases the determinant is the M (Meyer, Erard, Erdberg, Mihura, & Viglione, 2011). Based on the theory under the empathy capabilities (Ritter, Dziobek, Preißler, Rüter, Vater, et al. 2011), it could be different the level of cognitive empathy skills if a movement is seen or is inferred. When a subject gives an M response, the movement is just described (e.g. "*A butterfly with open wings*"); when the movement is anticipated, there is in some way an inference about mental states, intentions or wishes of the object (e.g. "*A butterfly with open wings, as if it is going to fly*"). Initially, the description of an inferred behavior seemed simple, and the RES 4 variable just detected the anticipated movement. However, in the protocols we found some other descriptions that could represent a mental state inference. Therefore, the other two variables (RES 4, and 5) were added.

RES 6: this variable was based on the items 1, 3, 4, 6, 18, and 28 of the QCAE, all of which are present in the Online Simulation subscale, and are related to the capability to have different points of view. Since the ambiguity of the Rorschach stimuli, someone who suggests that there is just one way to decode the inkblot could be accustomed to have this approach also in daily life. For this item, we expect a negative correlation with cognitive empathy.

RES 7, and 8: These variables were suggested by the interpretation of the AGM in the R-PAS manual. The AMG code could represent the respondent's attitude towards the aggression and may include pleasure and interest, or fear and detachment. It may be possible to determine this attitude by his/her verbalizations while giving the response (Meyer, Erard, Erdberg, Mihura, & Viglione, 2011, p. 345). Thus, we tried to keep a record of the mood of the AGM responses.

RES 9: with a hypothesis similar to the previous one, we considered helpful to take in account the emotional valence of responses, in particular the ones where were present thematic or cognitive codes.

RES 10, 11, 12, and 13: these items derived mostly from the administration practice. In particular, they try to take into account oppositional, compliant, or supportive aspects in the relationship with the examiner. A similar interpretation is given to the Card Turning, the Prompt, and the Pull when a protocol has an high frequency of them.

RES 14: the R-PAS manual suggest that most Human Detail responses are faces, which are defined as a compelling interpersonal stimulus (p. 335). However, human faces are coded as a simple human detail, which – in terms of interpretation – suggests a limited and less cognitively sophisticated conceptualization of others. If the human face is an interpersonal stimulus, interpreting it as something that implies a low conceptualization of others could be counterintuitive, or maybe incorrect.

RES 15: this item came from theory, and from administration practice. Gesture of compassion, help, or support are not so rare in the Rorschach, but in the R-PAS there are not codes for this kind of responses. This variable was developed following the "behavioral representation" principle, as others were (i.e. COP, MOR). Namely, the situational described in the inkblot is closely parallels to the real-life behavior that the code is supposed to investigate (Weiner, 1977; Meyer, Viglione, & Giromini, 2014).

5.1.2. CODING PRACTICE

The main selection of the RES variables was made through a work of blind "cross-coding" between two examiners in 25 Rorschach protocols, in order to test how the scoring was coherent between them.

One of the two examiners knew nothing about the rationale behind the RES, this condition was helpful to better understand the issues related to the new variables, which indeed became more evident.

Some items were excessively dependent on the subjective examiner's judgment (RES 2a, RES 7, RES 8, RES 9)[13], and this could compromise the reliability of the scores. Some others (RES 2b, RES 2c)[14] were not coherent with the Rorschach request, like trying to understand what kind of expression is the one in the card, or to explain the reason why an object has a particular mood. When someone sees an expression in the card, probably he/she has already identified and decoded it, and further elaborations are not required in the instructions. Moreover, R-PAS recommends to avoid the word "why" during the R-PAS administration, thus generally the respondent is dissuaded to explain "why"[15]. The RES 2d was considered redundant with the main item RES 2, because the information brought is quite the same, namely an attitude to pay attention to those details that could suggest emotion or expressions. The RES 10 and RES 11[16] could be influenced by too much different factors (e.g. anxiety, examiner's clarity in explaining the request, the respondent's comprehension of the task, and the elaboration of the response). Thus, we cannot be sure that this

---

[13] RES 2a) The emotional or mental state of the object (e.g., a person, an animal, an object) is incongruent with the context; RES 7) The words used by the respondent in AGM or MAP responses suggest a positive mood or emotion; RES 8) The words used by the respondent in AGM or MAP responses suggest a negative mood or emotion; RES 9) Emotional valence of the response.

[14] RES 2b) The respondent is confused about what emotion an object; RES 2c) The respondent is unable to explain the reasons why an object of the response experiences an emotion.

[15] If in the Clarification Phase the respondent asks "*Do I have to say he reason why*?", the examiner has to take back him to the request "*what about the inkblot made it look like that to you*".

[16] RES10) During the CP, how many questions does the examiner need to make in order to clarify the response?; RES 11) During the CP, the respondent does not accurately clarify his/her response, so that the examiner needs to make multiple questions about the same feature/component in order to obtain the needed information.

kind of behaviors are oppositional (and the frustration felt by the examiner was not sufficient to assess the respondent's empathy skills!). Finally, during the administration, cooperative or oppositional behaviors are often nonverbal (e.g. the respondent waits for the examiner to finish writing, or does not); furthermore, the verbalized ones – which usually sound like comments – are often not write down by the examiner. Therefore, given that the information leaded by the RES 12/13[17] was easy to get lost, it was removed.

This was the second version, of the RES.

1. The respondent verbally expresses an emotional arousal due to the exposure to the card. (e.g., "*This is sad*"; "*This scares me*"; "*Oh, this is so cheerful!*"; "*Ahhh, a skinny poor small b.f.*")	YES (1) NO (0)
2. The object of the response (e.g., a person, an animal, an object) is described as feeling or expressing emotions or feelings. (e.g., "*A sad face*"; "*two women who are washing clothes… they seem happy!*"; "*An angry dog!*")	YES (1) NO (0)
3. The response includes behaviors, actions, emotions or feelings that are about to happen or be experienced. (e.g., "It's about to catch this thing in the middle"; "These two ladies have a luggage, they're leaving, but they're gonna miss home")	YES (1) NO (0)
4. The object of the response (e.g., a person, an animal, an object) is described as having intentions, wishes, and/or desires. (e.g., "Here is the kid, he would like to have some toys to play with"; "That's the lion, it'd like to catch some pray")	YES (1) NO (0)
5. The object of the response (e.g., a person, an animal, an object) is described as being waiting for something, or thinking of someone or something. (e.g., "A person who is waiting for the bus"; "A guy who is thinking of his wife")	YES (1) NO (0)
6. The respondent emphasizes his/her own perspective, ignoring that different people might see different things. (e.g., "This is obviously a butterfly!"; "I'm a musician, so I see it perfectly: I'm sure that this is a guitar!"; "Clearly a sword!")	YES (1) NO (0)
7. The respondent sees a human face (Hf) in the response.	YES (1)

---

[17] RES 12) The respondent expresses a concern to help the examiner beyond of what was asked in the instructions; RES 13) The respondent resists to cooperate with the examiner during the RP, beyond prompts (Pr) and pulls (Pu).

	NO (0)
8. The respondent sees gesture of compassion, help, support or attempts to save a living being. (e.g., "These are two arms that are holding this other arm", "This is a person who is trying to help this one")	YES (1) NO (0)

5.2. METHODS

This study has benefited from four master's theses projects. In general, the approach was administering instruments that assess features related to the empathy capabilities, for the convergent validity, together with the criterion variable (QCAE) and the R-PAS. In all the samples, the RES was coded.

The first study was conducted from January to December in 2016. The additional instruments for the convergent validity were the Reading the Mind in the Eyes Test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, & Jolliffe, 1997) the Toronto Alexithymia Scale Bagby, Parker, & Taylor, 1994), the Interpersonal competence Questionnaire (Buhrmester et al., 1988). This study was focused on the empathy features related with the relationships and the emotions.

The second study, was conducted between October and December 2016. The R-PAS and the QCAE were integrated with the Difficulty in Emotions Regulation Scale Gratz & Roemer, 2004). After our contribution on the Italian validation of the QCAE, we were interested to the relation between empathy and emotion regulation in the R-PAS.

In the last one, the ICQ, the RME-T, the TAS-20, and the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005) were used in order to study how the Rorschach variables were connected with different levels of dominance and arousal in relational situations. It started in July and finished in September 2017.

73

Studies	M	F	Age	Examiners	Convergent Validity
1° (n = 173)	56	117	18-33 (M 23.41, SD 3.24)	IB, MDG, SR, DG, FA, CC, GG	ICQ, TAS, RME
2° (n = 54)	14	40	18-28 (M 22.30, SD 2.36)	AR, ER	DERS
4° (n = 60)	15	44	18-38 (M 23.17, SD 3.95)	FM, GA	ICQ, TAS, RME, IAPS

To minimize the influence of examiners' competence and any within-study idiosyncrasies, while maximizing the generalizability of the findings, we decided to analyze all data combined. So, our initial data set was comprised of 287 protocols. Next, we split this data set into two sub-sets of data, to select the variables to be included in our Rorschach Empathy Index (Developmental sample, n = 143) and then cross-validate this index with an independent sample (Validation sample, n =144). The allocation of protocols to the Developmental versus Validation samples was conducted at random, after ensuring that each examiner contributed equally to each of the two samples. In doing so, we equally scattered the differences among the examiners across the two samples.

5.2.1.   PARTICIPANTS

Inclusion–exclusion criteria were 1) being Italian citizen 2) not receiving psychiatric medications[18], 3) being 18 years old or more, 4) be unaware about the Rorschach test, 5) do not be colorblind.

The two samples were balanced for age and educational levels. The percentage of male and female in the two samples was marginally different, but this difference was not statistically significant.

---

[18] Informaion about the medications were asked in the sociodemographic questionnaire.

Composition of the Samples.

	Sex		Age	Educational Level			
	M	F		Secondary education	High school	Bachelor's degree	Master's degree
Development (n = 143)	38 (26.6 %)	104 (72.7 %)	18-35 (M 23.21, SD 3.22)	2 (1.4 %)	72 (50.3 %)	50 (35.0 %)	17 (11.9 %)
Validation (n = 144)	47 (32.6 %)	97 (67.4 %)	18-38 (M 23.09, SD 3.35)	4 (2.8 %)	72 (50.0 %)	51 (35.4%)	16 (11.1 %)

5.2.2.   PROCEDURE

The ethics committee of the University of Turin gave its approval prior to beginning data collection in all of four studies.

The recruitment was made with a snowball sampling, using flyers, social networks, personal knowledge, and word of mouth. In all the studies, the volunteers were informed about a study on self and other's emotions. All the participants were volunteers, any incentive was offered (e.g., monetary compensation) for their participation and they do not receive restitution about clinical results of the tests.

The volunteers were told that participation was voluntary, that they could interrupt or end their participation at any time, and that all the data were anonymous. In line with the Helsinki Declaration, all of them were asked to read and sign an informed consent statement prior to participating in the study. In the informed consent there were listed the aim of the study, the instruments, the average time necessary to complete the tests, and possible collateral effects (e.g. tiredness). Until the presentation of the informed consent, they did not know in details what kind of tests they were going to do, especially in the case of the Rorschach test, because it is easily available on internet.

5.2.3.   MEASURES

In order to avoid any possible priming effects, the Rorschach test was always the first test administered. The sequence of all the other tests was randomized.

An ID code was given to the participants, they had to use it for the self-report questionnaires, which were all administered online, using the Google Form platform. The links were sent by e-mail to the volunteers after the Rorschach test, and in the e-mail it was asked to fill out the instruments in a condition where they were alone, quiet, and concentrated.

The online administration choice was made for two reason: the first one, was to maintain the original administration of the QCAE (Reniers et al, 2011), so we adapted all the instruments to the online administration. The second one, was because of the time requested for the Rorschach administration, which is from one to two hours. With the online administration, the participants could do the tests when they were more in comfort.

Some of the measures used in this study were the ones already present in the contribution to the Italian validation of the QCAE (Di Girolamo, Giromini, Winters, Serie & de Ruiter, 2017): the QCAE (Reniers et al., 2011), the ICQ (Buhrmester et al., 1988), the TAS-20 (Bagby, Parker, & Taylor, 1994), the RME-T (RME–T; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, & Jolliffe, 1997), the DERS (Gratz & Roemer, 2004). The other was the International Affective Pictures System (IAPS; Lang, Bradley, & Cuthbert, 2005).

Rorschach Performance Assessment-System (R-PAS; Meyer et al., 2011). All the eleven examiners were formed through two steps. First of all, they attended the R-PAS classes at the University of Turin, where it was possible learning all about the scoring. Then, when they were interested in one of the research projects conducted within the lab, they were trained in the administration procedures. All examiners were asked to make a video of their first administration attempts, and they continued to do so until they were prepared to

begin data collection. Moreover, during data collection, the R-PAS trainer supervised the administration and scoring of all protocols.

R-PAS administration was conducted in accordance to the R-PAS guidelines, in a well-lit and silent room, without any peculiar stimulus such as paints, photos, or food. The participants were asked to turn off the phone. 56 of these protocols contributed to a Rorschach inter-rater reliability study (Pignolo, Giromini, Andò, Ghirardello, Di Girolamo, Ales, Zennaro, 2016).

Below I detail the main hypotheses we formulated, concerning the association of R-PAS variables to empathic capabilities:

Productivity and complexity: In general, productivity and complexity are associated with involvement in the task, and high cognitive resources. Empathy – in its whole complexity – requests a variety of cognitive resources, so a positive correlation was expected.

Administration behaviors: the behaviors related to the administration are representative of the interactions with the examiner and the task. High scors on Prompt, Pull or Card Turning (Pr, Pu, or CT) could be related to oppositionality tendencies. Thus, we anticipated a negative correlation between these variables and the QCAE scores.

Localization: we had no specific hypotheses about the relationship between localization variables and QCAE scores.

White space: as regard to the use of the white space, we focused on Space Reversal (SR), which represents creativity, originality, but also to oppositionality. Thus, we presumed a negative correlation between SR and the QCAE scales. Conversely, we did not have any specific expectations concerning the association of Space Integration (SI) to QCAE.

Contents: our hypotheses concerning R-PAS content variables mostly focused on the Human Contents (H) where we expect a positive correlation with empathy. Moreover, Art and Antropology (Art and Ay) should present a negative correlation with the empathy, because they represent a way to move away from

emotions and one's own inner world, and a high intellectualization of the daily experiences.

Organization and Perception: the capability to synthetize (Synthesis; Sy) information that comes from the environment should be positively related to the empathy features, as empathy, like Sy, involves some sophisticated, cognitive skills.

Formal Quality: in general, the capability to see the world as the other people do is an index of a shareable view of the world and of the events. Thus, the Ordinary Formal Quality (FQo) and the Popular (P) responses should be positively related with empathy capabilities.

Determinants: among the determinants, we expected a positive correlation between the QCAE scores and the Human Movement (M) variable, especially with the cognitive scale, because M is presumably related to the activity of the Mirror Neuron System (Giromini, Porcelli, Viglione, Parolin, & Pineda, 2010; Pineda, Giromini, Porcelli, Parolin, & Viglione, 2011; Andò, Salatino, Giromini, Ricci, Pignolo, et al., 2015; Porcelli, & Kleiger, 2016; Giromini, Viglione, Brusadelli, Zennaro, Di Girolamo, et al. 2016; Giromini, Viglione, Pineda, Porcelli, Hubbard, et al. 2017). Moreover, the Reflection (r) should be negatively correlated with the empathy, given that this variable indicates a self-centered way to process information.

Cognitive Codes: given that in the psychotic disorders is present a lack of empathy (Weiss, Kohler, Nolan, Czobor, Volavka et al., 2006; Abu-Akel and Abushua'leh, 2004; Majorek et al., 2009), and that cognitive codes reflect an increased risk for thought disorders, these variables should be negatively related to the empathy levels.

Thematic Codes: in this category there are different kinds of variables. Abstraction (ABS), together with Ay and Art, is present in the IntCont, thus it should have a negative correlation with empathy, in that it reflects a tentative to step away from one's own deeper emotions and feelings. Cooperation (COP) and Mutuality of Autonomy Health (MAH) are related to cooperative, positive,

78

benevolent, mature, and healthy representations of others. Thus, their correlation to the empathy scales should be positive. Conversely, Mutuality of Autonomy Pathological (MAP) reflects poor object relations, and difficulties interacting with others in mature and mutually enhancing, supportive, and autonomous ways; as such, we expected a negative correlation between MAP and empathy capabilities.

Rorschach Empathy Scale. In general, our hypotheses were that RES variables would correlate positively with the items they were derived from. More in details, RES 1 should correlate positively with the Affective Empathy Scale. RES 2, being composed of cognitive empathy items, should correlate positively with the QCAE Cognitive Empathy scale. The same can be said for the RES 3/4/5. RES 6 is the only variable from which we expected a negative correlations with empathy capabilities, because it describes behaviors related to a lack in perspective taking. RES 7 was designed to inspect whether seeing a human face would be considered an interpersonal stimulus; if that was the case, a positive correlation with the cognitive empathy should occur. RES 8 was supposed to detect a complex, and composite behavior, maybe the core of what is commonly identified as "an empathic gesture". Given that, we expected a positive correlation between RES 8 and both types of empathy.

Questionnaire of Cognitive and affective Empathy (QCAE; Reniers et al., 2011). The Italian version of the QCAE was the same used in our previous study (Di Girolamo, Giromini, Winters, Serie & de Ruiter, 2017). In the samples of this research, the levels of Cronbach's Alphas were from acceptable to excellent (Nunnally, 1978).

Cronbach's Alphas

| QCAE Scales | No. of items | Sample | | |
		Development	Validation	Total
Cognitive Empathy	19	.87	.86	.87
Perspective Taking	10	.86	.81	.84
Online Simulation	9	.81	.83	.82
Affective Empathy	12	.80	.82	.81

Emotion Contagion	4	.76	.70	.73
Proximal Responsivity	4	.67	.68	.68
Peripherical Responsivity	4	.70	.72	.71
Total	31	.87	.85	.86

The Interpersonal Competence Questionnaire (ICQ; Buhrmester et al., 1988; Giromini, de Campora, Brusadelli, D'Onofrio, Zennaro, et al., 2016). In our sample, Cronbach's Alphas were from good to excellent (Cohen, 1988). The hypothesis related to this test was that there would be positive correlations between our Rorschach empathy indices, especially with the cognitive one, and the ICQ scores.

Cronbach's Alphas

		Sample		
ICQ Scales	No. Of items	Development (n. = 83)	Validation (n. = 81)	Total (N. = 164)
Initiate Relationships	8	.87	.89	.88
Emotional Support	8	.86	.86	.86
Negative Assertion	8	.79	.74	.77
Self-Disclosure	8	.86	.83	.84
Conflict Management	8	.76	.81	.78
Total	40	.92	.92	.92

Toronto Alexithymia Scale-20 (TAS-20; Bagby, Parker, & Taylor, 1994). In our sample, Cronbach's Alphas were from good to excellent (Cohen, 1988). In literature, alexithymia is negative related to the empathy (Jonason & Krause, 2013), so we expected negative correlation between both Rorschach empathy indices, cognitive and affective, with the TAS-20 scores.

Cronbach's Alphas

		Samples		
TAS-20 Alphas	No. Of items	Development (n. = 90)	Validation (n. = 89)	Total (N. = 179)
Difficulties Identifying Feelings	7	.87	.80	.84
Difficulties Describing Feelings	5	.75	.77	.76
External Oriented Thinking	8	.65	.67	.66
Total	20	.86	.86	.86

Difficulties in Emotion Regulation Scale-16 (DERS; Bjureberg, Ljótsson, Tull, Hedman, Sahlin, et al., 2016). In this study, we used the brief version of the DERS, the one with 16 items, which were selected from the original ones. DERS-16 maintained a high internal consistency, excluded items with the most error variance, and reduced the redundancy. In the development study of the DERS-16, Cronbach's Alphas ranged from .92 to .95. In our samples, the Cronbach's Alpha for the total scale is from .90 to .91, and in general were all from good to excellent except the "nonacceptance" one, in the Validation sample, which was in the poor range.

The relation between empathy and emotion regulation is debated. While it is true that a contact with emotions is essential to understand other's emotional state – indeed the alexithymia is negative related with empathy – there are some evidences (Henry, 2008; Di Girolamo, Giromini, Winters, Serie & de Ruiter, 2017) about a positive correlation between emotion dysregulation and affective empathy. Thus, we expected a positive correlation between the Rorschach affective index and the DERS score, and a negative or a non significant correlation with the Rorschach cognitive empathy index.

Cronbach's Alphas

DERS-16 Scales	No. Of items	Samples		
		Development (n. = 27)	Validation (n. = 27)	Total (N. = 54)
Nonacceptance	3	.60	.44	.50
Goals	3	.82	.76	.79
Impulse	3	.79	.89	.83
Strategies	5	.79	.86	.82
Clarity	2	.90	.84	.88
Total	16	.91	.90	.90

Reading the Mind in the Eyes-Test (RME-T; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, & Jolliffe, 1997). In this research, Cronbach's Alpha was .53 – in the development sample – and .60 – in the validation sample. A link with the standardized glossary of the RME–T was

present in the online administration. Given that some RES items were about the recognition of other's expressions or emotions, we chose the RME-T as for the convergent validity. Thus, we expected a positive correlation with the Rorschach empathy indices, in particular with the cognitive one.

International Affective Pictures System (IAPS; Lang, Bradley, & Cuthbert, 2005). The IAPS is a large set of standardized, emotionally evocative, stimuli that includes contents across a wide range of semantic categories. The aim was to create standardized materials for the study of emotion and attention. The IAPS were developed and distributed by the NIMH Center for Emotion and Attention (CSEA), at the University of Florida.

In the IAPS, three emotional dimensions are taken into account: valence, arousal, and dominance. The Valence is about the positive or the negative value of the situation presented by the picture, and it ranges from pleasant to unpleasant; the Arousal assesses the level of physical activation due to the exposure to the picture, and ranges from calm to excited. The third dimension is the Dominance, and it represents how the person feels him/herself involved in the presented situation, it ranges from 0 to 9 (Lang, Bradley, & Cuthbert, 2005). For this study, we chose 15 pictures: 5 where people were not involved in interactions or nothing was going to happen, 5 with positive interactions among people and others (included animals), and 5 with negative interactions or where there was happening something bad.

The assessment of three dimensions is based on the Self Assessment Manikin, a sheet with three set of five human figures, where Valence, Arousal, and Dominance are graphically represented. The subject can place the X both within the human figures and in the middle of two of them, thus the evaluation can be made on a 9 points Likert scale.

In this study, the administration of the IAPS was made using Google Form platform: the IAPS picture was showed first, just below it there was the image of the SAM, and then a 9-point Likert scale, where the subjects could place their choice. The IAPS pictures were randomized.

Valence, Arousal and Dominance are three dimensions that could contribute to empathic behaviors, because each of them plays a role in the subject's mood, feelings and reaction. Our hypothesis was that the Rorschach empathy indices correlate positive with the arousal and dominance levels, and with the accuracy of the valence attribution.

S.A.M.
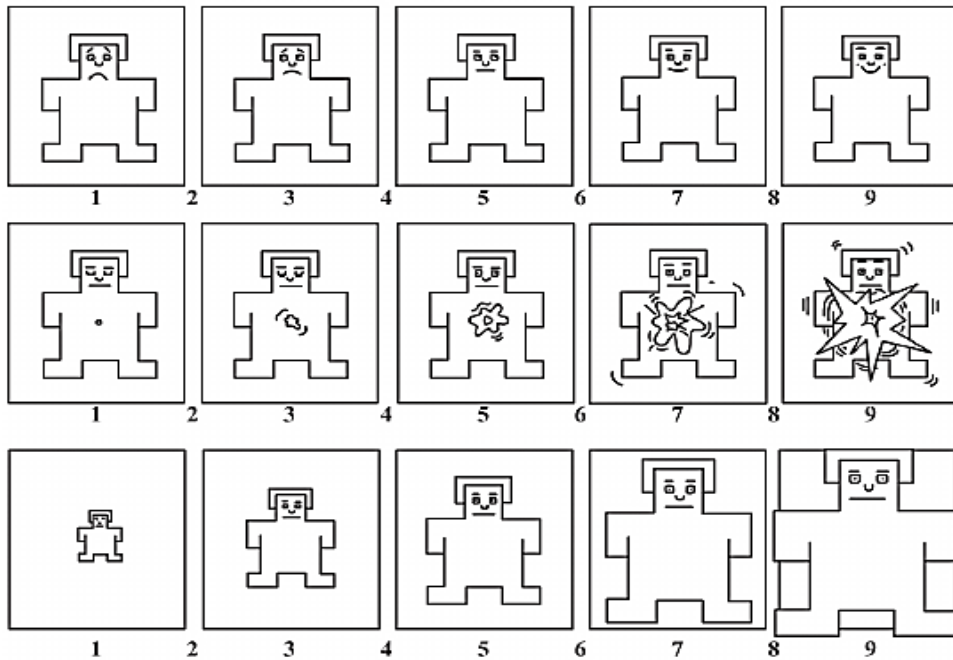
Study sample IAPS averages compared with the normative IAPS data.

	Male				Female				Total			
	Sample (n= 15)		Normative Sample		Sample (n= 42)		Normative Sample		Sample (N= 58)		Normative Sample	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
No Interactions												
2190												
Valence	4.67	1.75	4.73	1.25	4.22	1.57	4.90	1.31	4.36	1.62	4.83	1.28
Arousal	4.17	2.38	2.27	1.72	3.74	2.16	2.50	1.86	3.84	2.19	2.41	1.80
Dominance	4.17	2.77	5.48	2.08	3.22	2.13	6.23	1.91	3.44	2.31	5.92	2.01
2320												
Valence	5.72	1.87	5.45	1.33	6.41	1.81	6.82	1.37	6.26	1.83	6.17	1.51
Arousal	3.56	2.01	2.56	1.68	5.28	2.06	3.20	2.03	4.84	2.15	2.90	1.89
Dominance	3.78	2.49	6.55	1.75	5.50	2.29	6.77	1.87	5.04	2.43	6.66	1.81
2480												
Valence	4.22	2.07	4.76	1.23	4.65	2.15	4.77	1.90	4.58	2.13	4.77	1.64
Arousal	4.67	2.47	2.80	1.81	5.48	2.29	2.55	1.76	5.32	2.35	2.66	1.78
Dominance	4.72	2.19	5.61	1.91	5.04	2.50	5.11	2.21	4.97	2.40	5.33	2.09
2038												
Valence	5.50	1.38	5.08	1.19	5.56	2.02	5.11	1.49	5.53	1.86	5.09	1.35
Arousal	3.89	1.78	2.98	1.66	4.20	2.00	2.84	2.09	4.11	1.93	2.94	1.93
Dominance	4.28	2.14	6.21	1.80	4.93	2.36	6.51	1.90	4.79	2.31	6.36	1.85
2200												
Valence	4.44	1.82	4.64	1.18	4.95	1.56	5.03	1.39	4.01	1.79	4.79	1.38
Arousal	3.44	2.25	2.33	1.78	4.03	2.22	4.93	1.65	3.44	2.14	3.18	2.17
Dominance	2.89	2.30	5.97	1.98	4.90	2.10	5.32	1.77	2.89	2.16	5.44	2.17
Positive Interactions												
2550												
Valence	7.11	1.94	7.37	1.20	8.11	1.67	8.14	1.53	7.88	1.77	7.77	1.43
Arousal	6.00	2.20	4.15	2.03	6.54	2.04	5.16	2.67	6.37	2.09	4.68	2.43
Dominance	5.06	2.46	5.89	1.65	5.69	2.71	6.52	2.15	5.55	2.64	5.89	1.65
2091												
Valence	7.00	1.64	6.99	1.40	7.28	2.09	8.26	1.17	7.23	1.98	7.68	1.43
Arousal	5.00	2.63	4.20	2.03	6.28	2.10	4.77	2.47	5.99	2.29	4.51	2.28
Dominance	4.22	2.78	6.60	1.70	5.65	2.36	6.41	1.94	5.30	2.51	6.60	1.70
2360												
Valence	6.78	2.18	6.98	1.76	7.72	1.77	8.20	1.59	7.48	1.90	7.70	1.76
Arousal	4.61	2.03	3.65	2.02	6.24	2.33	3.67	2.52	5.78	2.38	3.66	2.32
Dominance	4.22	1.83	6.46	2.07	5.48	2.77	7.24	2.05	5.16	2.59	6.46	2.07
4599												
Valence	6.72	1.99	7.02	1.28	7.52	1.72	7.23	1.66	7.34	1.81	7.12	1.48
Arousal	5.61	2.52	5.73	1.93	6.31	2.41	5.64	1.96	6.15	2.42	5.69	1.94
Dominance	5.78	2.56	6.54	1.73	6.59	2.48	6.43	1.86	6.42	2.51	6.54	1.73

2655

Valence	7.00	1.85	6.62	1.47	7.81	1.68	7.06	2.42	7.63	1.74	6.88	2.09
Arousal	6.17	1.86	4.15	1.99	6.52	2.06	4.87	2.28	6.45	2.00	4.57	2.19
Dominance	5.56	2.28	6.09	2.21	6.00	2.78	6.18	2.44	5.89	2.64	6.09	2.21

Negative Interactions

6570

Valence	2.28	1.64	2.29	1.84	1.76	1.94	2.10	1.61	1.88	1.86	2.19	1.72
Arousal	5.28	2.67	6.06	2.44	6.65	2.36	6.43	1.85	6.33	2.48	6.24	2.16
Dominance	3.94	2.18	4.02	2.73	4.26	2.50	4.04	2.33	4.21	2.41	4.03	2.52

2410

Valence	4.17	2.20	4.72	1.46	4.15	1.99	4.54	1.94	4.16	2.02	4.62	1.72
Arousal	4.61	2.35	3.83	2.16	4.96	2.31	4.40	2.39	4.86	2.30	4.13	2.29
Dominance	4.94	2.34	5.46	2.03	5.50	2.64	4.58	2.46	5.33	2.56	5.00	2.29

9926

Valence	3.44	1.62	4.16	1.56	4.11	2.91	3.56	1.53	3.95	2.63	3.85	1.59
Arousal	5.00	2.14	5.02	1.71	6.02	2.34	4.65	2.13	5.71	2.35	4.83	1.95
Dominance	4.11	2.30	4.63	1.99	4.31	2.60	4.14	2.00	4.23	2.51	4.37	2.00

4621

Valence	3.00	1.37	3.83	1.65	3.11	2.05	2.71	1.38	3.08	1.88	3.19	1.59
Arousal	4.89	2.45	4.49	1.97	5.22	2.38	5.24	2.39	5.11	2.38	4.92	2.24
Dominance	4.06	2.46	5.53	2.04	4.85	2.37	3.51	2.31	4.63	2.40	4.37	2.41

6315

Valence	2.44	1.58	2.94	1.89	1.74	1.68	1.72	1.23	1.92	1.66	2.31	1.69
Arousal	5.39	2.43	6.04	2.16	6.24	2.37	6.69	2.57	6.03	2.38	6.38	2.39
Dominance	3.22	1.86	4.64	2.62	4.57	2.89	3.00	2.52	4.23	2.70	3.79	2.69

Total Sample Standard Scores Descriptive Statistics (n = 58) [19]

	Min.	Max.	Mean	SD	Skew.	Kurtosis
No Interactions						
2190						
Valence	55.12	148.87	94.52	19.72	0.11	0.36
Arousal	88.25	146.58	112.39	18.71	0.31	-1.13
Dominance	63.28	122.99	81.43	17.63	0.75	-0.59
2320						
Valence	58.58	128.11	100.02	18.62	-0.33	-0.61
Arousal	84.92	148.41	114.61	16.79	0.09	-0.95
Dominance	53.09	119.39	86.24	19.64	-0.07	-1.00
2480						
Valence	65.52	138.69	97.37	19.63	0.20	-0.89
Arousal	86.01	153.43	121.46	19.38	-0.14	-0.86
Dominance	68.92	126.34	96.27	16.65	-0.05	-0.93
2038						
Valence	65.67	143.44	104.36	20.73	-0.04	-0.49
Arousal	84.92	147.10	107.97	15.54	0.42	-0.44
Dominance	56.54	121.41	85.62	18.97	-0.04	-1.21
2200						
Valence	58.80	145.76	90.85	19.89	0.13	0.08
Arousal	84.93	140.23	100.66	14.23	0.84	0.07
Dominance	69.31	124.61	81.23	14.23	1.35	1.18
Positive Interactions						
2550						
Valence	39.48	112.90	99.70	19.80	-1.61	1.70
Arousal	83.46	126.67	110.17	13.30	-0.44	-0.90
Dominance	55.55	128.27	96.77	23.73	-0.20	-1.23
2091						
Valence	29.93	113.85	95.58	20.28	-1.24	1.03
Arousal	76.91	129.54	109.80	14.53	-0.47	-0.73
Dominance	50.59	121.18	88.47	21.36	-0.19	-1.00
2360						
Valence	42.90	111.08	96.39	17.25	-1.22	0.81
Arousal	82.80	134.53	113.01	15.48	-0.23	-1.15
Dominance	60.43	118.41	90.04	18.72	0.04	-1.24
4599						
Valence	48.11	119.05	101.58	18.97	-1.15	0.53
Arousal	63.74	125.59	103.60	18.48	-0.62	-0.77
Dominance	51.97	121.33	98.61	21.78	-0.72	-0.72

[19] IAPS Standard Scores of Valence, Arousal and Dominance were obtained based on the normative data from the IAPS manual (Lang, Bradley, & Cuthbert, 2005).

2655						
Valence	72.15	115.22	104.94	12.40	-1.28	0.89
Arousal	75.55	130.34	111.68	14.33	-0.71	-0.33
Dominance	65.45	119.75	97.28	18.30	-0.42	-1.12
Negative Interactions						
6570						
Valence	89.62	159.39	98.04	17.89	2.45	5.19
Arousal	63.61	119.17	100.49	16.87	-0.67	-0.64
Dominance	81.96	129.58	99.82	14.10	0.35	-1.10
2410						
Valence	68.43	138.20	95.04	17.17	0.60	-0.47
Arousal	79.50	131.90	104.23	14.52	-0.15	-1.05
Dominance	73.80	126.20	101.36	15.94	-0.20	-1.24
9926						
Valence	73.11	148.58	99.14	24.19	0.67	-0.90
Arousal	70.54	132.08	106.35	18.26	-0.33	-1.14
Dominance	74.73	134.73	98.78	18.52	0.54	-0.79
4621						
Valence	79.34	145.38	97.88	18.10	0.90	0.02
Arousal	73.75	127.32	101.11	16.08	-0.14	-1.04
Dominance	79.02	128.82	101.13	15.34	0.16	-1.12
6315						
Valence	88.37	159.38	96.79	15.46	2.31	5.22
Arousal	66.23	116.44	96.97	15.09	-0.32	-1.22
Dominance	84.44	129.05	102.23	14.68	0.43	-1.24

5.3. ANALYTICAL APPROACH

The aim of this study was to identify Rorschach variables that could assess the empathy in its two dimensions, cognitive and affective. As noted above, we split our combined sample into two halves: half was used to select the variables to be included in our empathy index (Developmental sample), the other half was used to cross-validate the resulting index (Validation sample).

5.3.1. DEVELOPMENT PHASE – STUDY 1

In the development phase, we inspected the developmental sample only, and correlated selected, nonredundant R-PAS variables (see below) and RES variables with the Cognitive Empathy and the Affective Empathy scales of the QCAE. This phase aimed at selecting those R-PAS and RES variables that could contribute to the assessment of empathy, and that could thus contribute to the development of a Rorschach empathy index.

R-PAS output includes 392 variables (divided in coding variables, and indices derived from the percentage, the sum or the proportion of coding variables). Each of these variables is calculated in Raw Scores, Raw Standard Scores, and Complexity Adjusted Standard Scores. In our data analysis, we did not use the indices, in order to avoid redundancy with the variables they are composed of. The complexity-adjusted scores were excluded because the complexity correction is not available for all the variables (e.g. Card Turning, Pull or Prompt), and the standard scores were chosen because they tend to normalize the distribution. Thus, we used the Raw Standard Scores and then we standardized the RES scores, based on the mean and the standard deviation of the development sample.

Moreover, we used a unique variable for the R-PAS Cognitive Codes (namely we ignored the distinction between level 1 and level 2, for DV, DR, INC,

and FAB), and for the A, Ad contents (namely we added the A and the Ad all together), because those variables assess similar constructs but, taken independently, have a low base rate. Moreover, in doing this we controlled the effect of possibly spurious outliers (Hastings, Mosteller, Tukey, & Winsor, 1947).

The Rorschach does not have strong correlations with the self-reported measures, as shown by Meyer (1997b) and Mihura (2012). Moreover, this was an explorative study and the RES variables were new, and never tested before. To be sure that we were not putting away useful variables, in the developmental study we decided to select those R-PAS and RES variables that have correlation greater than or equal to .10 (which is indicative of a *small* effect size; Cohen, 1988).

At the end of the developmental phase, the selected variables were pooled in indices for the assessment of the cognitive empathy, for the affective empathy and for the empathy in general. Those indices were therefore correlated with the criterion variable in the development sample.

5.3.2.   VALIDATION PHASE – STUDY 2

The correlations between the Rorschach empathy indices and the QCAE were replicated in the validation sample. As a further confirmation of the Rorschach empathy indices validity, we compared the averages scores of the obtained Rorschach empathy indices between the group with higher QCAE scores, and the group with lower QCAE scores. The subjects in the two groups were pooled basing on scores above and below one standard deviation from the averages in the QCAE scales. The QCAE standard scores were obtained using the mean and the standard deviation values derived from the online dataset in Di Girolamo, Giromini, Winters, Serie & de Ruiter (2017).

The convergent validity was tested with the correlations between the Rorschach empathy indices and the scales of the DERS, TAS-20, ICQ, RME-T, and IAPS.

The IAPS dimensions (i.e. Valence, Arousal and Dominance) were treated differently based on the meaning of the SAM scales. Indeed, in the case of the Arousal and the Dominance, the SAM scales provide a quantitative continuum (from the absence to a high level of occurrence), whereas the Valence SAM scale provides a qualitative continuum. This means that the two former dimensions range like the empathy construct ranges, while the latter does not. Thus, we needed to establish a way to identify the correct valence attribution. In order to do this, we created three dichotomous variables, based on the quality valence of the three groups. As regard to the Neutral group, the value of 1 was associated if the score given by the subject was comprised between -1 SD and +1 SD relative to the mean valence of that picture. For the Positive group, the value of 1 was assigned if the score given by the subject was in the range between the average and +1SD, the value of 2 if it was above 1SD. For the Negative group, the value of 1 was assigned if the score was between the average and -1SD, the value of 2 if it was below -1SD. In all the other cases, it was assigned the value of 0. To summarize, the value of 1 was assigned to all the scores that formed part of the "positive" or "negative" range using the averages as cut-off, the value of 2 was assigned to all the scores that were above the positive or the negative average more than 1SD. The rationale under this approach was that who is able to give a score at least in the average range, should have a good skills to evaluate the valence of the daily life situation. Moreover, who scores a positive picture with a value over the average, namely "more positive", should be more sensitive to the emotional quality of relational situations (idem for the negative pictures).

5.4. DEVELOPMENT SAMPLE RESULTS – STUDY 1

The Rorschach Cognitive Empathy Index (R-CEI) was composed by the correlations between the R-PAS variables and Cognitive Empathy QCAE scale, by Pearson's r greater than or equal to .10. The selected R-PAS variables were: number of responses (R = .103), responses to the 8th, 9th, and 10th card (R8910;

r = .111), Abstraction (ABS; r = -.189), Fire (Fi; r = .164), Human Movement (M; r = .103), Inanimate Movement (m = .117), Pair Responses (2; r = .113), Achromatic Color (C' = .143), Pure Color (C; r = -.105), Feral Movement (FM; r = .159), Synthesis (Sy; r = .169), and Cooperation (COP; r = .189).

The Rorschach Affective Empathy Index (R-AEI) was composed by the R-PAS variables that correlated with the QCAE Affective Empathy Scale for Pearson's r greater than or equal to .10. The selected R-PAS variable were: Card Turning (CT; r = -.179), Art content (Art; r = -.204), Anthropology (Ay; -.142), Explosion (Ex; r = .115), Fire (Fi; r = .146), Synthesis (Sy; r = .154), Pair (2; r = .158), Color Form (CF; r = -.204), Feral Movement (FM; r = .177), Deviant Responses (DR; r = -.118), Cooperation (COP; r = .135).

The Rorschach Empathy Index (R-EI) was composed by the sum of the variables present in the R-CEI and in the R-AEI. Those variables included in both indices were double weighted (e.g. COP, Fi, Sy).

Correlations between R-PAS variables and QCAE scales (N = 143).

R-PAS	QCAE COG		QCAE AFF		QCAE TOT	
	r	Sig.	r	Sig.	r	Sig.
Productivity						
R	.103	.222	.022	.795	.083	.325
R8910	.111	.186	.028	.737	.092	.274
Administration behaviors						
Pr	-.088	.297	-.009	.918	-.066	.437
Pu	.071	.397	-.064	.448	.017	.845
CT	-.083	.327	-.179	.033	-.150	.074
Localization						
W	.047	.579	.011	.896	.038	.651
D	.007	.931	.015	.856	.013	.877
Dd	.032	.707	.011	.895	.028	.741
Use of the white space						
SR	.038	.648	.033	.693	.044	.602
SI	.068	.422	-.041	.629	.026	.758
Contents						
H	.089	.291	.050	.553	.088	.297
SumH	.074	.377	.005	.954	.054	.520
NPH	.038	.651	-.053	.528	-.001	.990

SumA	.132	.117	.080	.340	.133	.113
An	.000	.999	.014	.869	.007	.931
Art	-.073	.383	-.204	.014	-.157	.062
Ay	-.057	.496	-.142	.091	-.113	.178
Bl	.042	.618	-.031	.709	.013	.878
Cg	.065	.442	-.033	.699	.028	.738
Ex	.082	.329	.115	.171	.117	.165
Fi	.164	.050	.146	.082	.189	.023
Sx	.043	.612	.042	.617	.052	.541
NC	.024	.774	.006	.943	.020	.813
Organization and perception						
Sy	.169	.043	.154	.066	.197	.018
Vg	-.086	.309	.011	.895	-.054	.523
Pair	.113	.179	.158	.059	.160	.056
Formal Quality						
FQo	.070	.404	.072	.394	.086	.307
FQu	.022	.793	-.031	.710	-.001	.992
FQm	.054	.518	.016	.849	.046	.584
FQn	-.070	.404	-.002	.985	-.050	.556
M-	.052	.537	-.073	.386	-.002	.985
Popular	.086	.310	.070	.409	.095	.257
Determinants						
M	.103	.219	.097	.249	.122	.147
FM	.159	.058	.177	.034	.202	.016
m	.117	.165	-.043	.607	.059	.485
FC	-.074	.381	-.008	.926	-.055	.511
CF	.030	.725	-.204	.014	-.085	.313
C	-.105	.211	-.086	.308	-.118	.162
C'	.143	.088	.020	.808	.110	.191
Y	.075	.376	-.025	.764	.039	.645
T	-.087	.304	.045	.593	-.037	.662
V	.049	.561	.027	.747	.048	.568
refl	-.013	.881	-.081	.338	-.050	.549
FD	.025	.764	.043	.613	.040	.638
F	-.098	.244	-.012	.883	-.075	.376
Complexity						
Blend	.094	.266	.022	.791	.077	.363
CBlend	-.015	.860	-.023	.785	-.022	.792
Cognitive Codes						
DV	.032	.703	-.033	.698	.005	.948
DR	-.099	.242	-.118	.161	-.129	.123
PEC	-.027	.745	-.037	.659	-.038	.650
INC	.083	.326	.042	.622	.079	.348
FAB	.057	.502	-.074	.380	.001	.989

Thematic Codes							
ABS	-.189	.024	-.050	.556	-.157	.061
PER	-.039	.640	.009	.913	-.023	.788
COP	.189	.023	.135	.108	.201	.016
MAH	-.003	.975	.067	.426	.033	.697
AGM	.066	.433	.094	.264	.095	.261
AGC	-.057	.496	.032	.703	-.023	.782
MOR	.092	.277	.081	.334	.106	.209
MAP	.006	.947	.080	.341	.045	.590
ODL	-.038	.656	.016	.854	-.018	.830

The same procedure was adopted for the RES. The selected variables were added to the three Rorschach empathy indices, and named with the same names but the addition of "Comp" (because they are "composed" of both R-PAS and the RES variables).

Correlations between RES variables and QCAE scales (n = 143).

	QCAE COG		QCAE AFF		QCAE TOT	
	r	Sig.	r	Sig.	r	Sig.
RES 1	.027	.746	.056	.503	.048	.568
RES 2	.041	.629	.052	.536	.055	.511
RES 3	-.004	.964	-.004	.963	-.005	.956
RES 4	.186	.026	.045	.590	.153	.068
RES 5	.076	.368	.093	.267	.101	.230
RES 6	-.114	.176	.031	.710	-.063	.456
RES 7	.057	.543	-.045	.627	.016	.862
RES 8	.053	.531	.167	.046	.123	.143

The Rorschach Cognitive Empathy Index Composite (R-CEI-Comp) included the R-CEI, and the RES variables correlated with the QCAE Cognitive Empathy scale: the object of the response is described as having intentions, wishes, and/or desires (RES 4; r = .186), and the respondent emphasizes his/her own perspective, ignoring that different people might see different things (RES 6; r = -.114).

The Rorschach Affective Empathy Index Composite (R-AEI-Comp) included the R-AEI, and the RES variable correlated with the QCAE Affective

Empathy Scale: the respondent sees gesture of compassion, help, support, or attempts to save a living being (RES 8; r = .167).

The Rorschach Empathy Index Composite (R-EI-Comp) consisted of the sum of the all variables selected from the R-PAS and the RES.

5.5. VALIDATION SAMPLE RESULTS – STUDY 2

The obtained indices were correlated with the criterion variable in the development sample, and then in the validation sample.

Pearson's Correlations between the Rorschach empathy indices and the QCAE scales, in the Developmental (n = 143) and in the Validation (n = 144) sample.

	QCAE COG		QCAE AFF		QCAE TOT	
	Dev.	Val.	Dev.	Val.	Dev.	Val.
R-CEI	.282**	.202*	.185*	.172*	.292**	.243**
R-CEI-Comp	.324**	.202*	.179*	.191*	.318**	.254**
R-AEI	.268**	.150	.416**	.234**	.401**	.241**
R-AEI-Comp	.259**	.116	.423**	.235**	.399**	.218**
R-EI	.312**	.196*	.330**	.220**	.388**	.266**
R-EI-Comp	.338**	.185*	.323**	.231**	.402**	.265**

** Correlation is significant at the 0.01 level. * Correlation is significant at the 0.05 level.

In general, in the development sample all the correlations were statistically significant, while in the validation sample they slightly shrank.

In the development sample, the correlation between the QCAE Cognitive Empathy Scale and the R-CEI was almost in the medium range (r = .284), but in the validation sample the correlation is low (r = .202). Similarly, the correlation of the R-AEI with the QCAE Affective Empathy Scale was in a medium range (r = .416) in the development sample, and in a low range in the validation one (r = .234). The final index, the R-EI, had a medium correlation with the total QCAE score (r = .388) in the develop sample, and a small to medium correlation in the validation sample (r = .266).

The indices with the RES variables followed the same trend. In the development sample, the correlation between R-CEI-Comp and the QCAE Cognitive Empathy Scale was r = .324, the R-AEI-Comp correlated with QCAE Affective Empathy Scale with r = .423, and the total index (R-EI-Comp) had a correlation with the total QCAE score with r = .402. In the validation sample, the correlations between the indices and the corresponding QCAE scale were the same of the previous three indices (R-CEI, R-AEI, R-EI), namely r = .202 for the cognitive empathy, r = .235 for the affective empathy, and r = .265 for the total empathy.

As a further confirmation of the Rorschach empathy indices validity, we conducted additional analyses. First, we compared the Rorschach empathy indices from high and low QCAE scores, in the validation sample; then, we calculated the convergent validity with the other instruments involved in the research; finally, we deepened the negative correlation between the CT variable and the empathy scales.

5.5.1.   QCAE SCORES COMPARISON - LOW VS HIGH

Given that our participants were nonclinical individuals, we were concerned that very few of them had very good or very poor empathic capabilities. With such limited variability, it is not surprising that our Rorschach-based, empathy indices did not correlated strongly with QCAE scores. Indeed, the correlation statistic is a measure of co-variation; as such, if a variable does not have enough variability, it cannot possibly co-vary with other variables, i.e., it cannot correlate strongly with them. Thus, we decided to select those individuals with very high QCAE scores, and those with very low QCAE scores, and compare our Rorschach-based, empathy indices between these two groups. Moreover, to better understand the role played by the RES variables, they were isolated and grouped in RESCOG, for the cognitive empathy, RESAFF, for the affective one, and RESTOT for the total empathy.

Comparison between Rorschach empathy indices from high and low QCAE scores, in the validation sample (n = 144).

	LOW QCAE		HIGH QCAE					
	M	SD	M	SD	t	df	sig.	*d*
R-CEI	67.33	4.98	71.16	5.26	-2.601	52	.01	-.74**
R-CEI-Comp	58.21	4.53	61.80	5.10	-2.566	52	.02	-.73*
RESCOG	-1.09	9.51	0.96	7.73	-0.856	52	.40	-.24
R-AEI	6.58	4.03	9.72	5.19	-2.414	48	.02	-.69*
R-AEI-Comp	14.29	3.76	17.03	4.73	-2.286	48	.03	-.65*
RESAFF	99.07	11.97	97.42	7.50	.566	48	.57	.16
R-EI	75.85	8.95	81.52	8.52	-2.168	43	.05	-.47*
R-EI-Comp	39.42	3.94	42.36	4.15	-2.421	43	.05	-.46*
RESTOT	31.05	7.41	34.42	6.90	-1.574	43	.32	-.23

** The difference is significant at the 0.01 level. * The difference is significant at the 0.05 level.

In general, there were differences in the Rorschach empathy indices scores according to the empathy levels assessed by the QCAE. More in detail, the R-CEI (d = -.74), the R-AEI (d = -.69) and the R-EI (d = -.47) had effect sizes from medium to a medium-large range (Cohen, 1988). However, the RES variables did not contribute significantly to the values Rorschach empathy indices, as it could be seen from the values of the Cohen's *d* in the composite indices, which were lower than the previous. Moreover, there were not differences in the RES scores between the two groups.

5.5.2. CONVERGENT VALIDITY

The convergent validity was calculated with the correlations between Rorschach empathy indices and the DERS, ICQ, TAS-20, RME-T and IAPS, we did not use the split sample but the combined one, because these measures did not contribute to the selection of the Rorschach or RES variables.

Convergent Validity between Rorschach empathy indices and the other instruments.

	R-CEI	R-CEIComp	R-AEI	R-AEI-Comp	R-EI	R-EI-Comp
DERS (n = 54)						
Non.	.044	.056	.164	.106	.108	.084
Goa.	.044	.041	.201	.196	.127	.118
Imp.	.156	.155	.300*	.305*	.241	.239
Str.	.113	.126	.263	.215	.198	.178
Cla.	.114	.119	.270*	.258	.202	.194
TOT	.122	.130	.305*	.271*	.225	.207
ICQ (n = 164)						
IR	.137	.181*	.056	.038	.112	.138
ES	.252**	.279**	.250**	.213**	.282**	.285**
NA	.073	.125	.031	.015	.060	.090
DC	.146	.194*	.137	.139	.159*	.194*
CM	.057	.091	.019	.011	.045	.066
TOT	.181*	.238**	.133	.113	.178*	.211**
TAS-20 (n = 179)						
DIF	-.019	.044	-.041	-.021	-.033	.018
DDF	-.115	-.118	-.128	-.112	-.135	-.131
EOT	-.190*	-.179*	-.177*	-.156*	-.206**	-.192*
TOT	-.122	-.086	-.134	-.110	-.143	-.109
RME-T (n =162)						
	.006	.025	.039	.035	.024	.033
IAPS (n = 58)						
VALENCE						
Tot.	.309*	.273*	.327*	.355**	.346**	.345**
Pos.	.181	.196	.276*	.328*	.244	.283*
Neg.	.093	.063	.096	.096	.103	.087
AROUSAL						
Neu.	-.277*	-.244	-.024	.039	-.176	-.135
Pos.	.021	.030	.232	.298*	.127	.164
Neg.	.078	.094	.271*	.314*	.180	.212
DOMINANCE						
Neu.	-.102	-.073	.023	.009	-.049	-.042
Pos.	.131	.140	.290*	.329*	.222	.248
Neg.	.144	.152	.308*	.328*	.238	.256

DERS = Difficulties in Emotion Regulation Scale: Nonacceptance, Goals, Impulse, Strategies, Clarity; ICQ = Interpersonal Competence Questionnaire; IR = Initiation Relationship; ES = Provide Emotional Support; NA = Negative Assertion; DC = Disclosure; CM = Conflict Management; TAS-20 = The Toronto Alexithymia Scale-20: DIF = Difficulties Identifying Feelings; DDF = Difficulties Describing Feelings; EOT = Externally Oriented Thinking; RME-T = Reading the Mind in the Eyes-Test; IAPS = International Affective Pictures System.

In general, the correlations with the DERS were in line with our expectations. The cognitive indices did not correlate with the DERS scales. However, the R-AEI and the R-AEI-Comp had positive statistically significant correlations with the Impulse, Clarity and Total scales, even if the former had better correlations than the latter. Nonacceptance and Goals subscales did not present significant correlations with any Rorschach empathy indices. Nonetheless, the R-EI and R-EI-Comp did not present significant correlations with the DERS subscales and the total score.

In general, the ICQ presented better correlations with the R-CEI and the R-CEI-Comp, and this was in accordance with our hypothesis. The R-AEI and the R-AEI-Comp did not correlate with the ICQ, except for the PES subscale, but – in general – the R-EI-Comp had a higher effect size than the R-EI. The only scale that did not present significant correlations with the Rorschach empathy indices was the CR subscale.

The TAS-20 was expected to have in general negative correlations with the Rorschach empathy indices. However, the latter did not correlate with the total TAS-20 score and the DDF subscale, while the EOT subscale had negative correlations with all of the six indices, which was expected.

The correlations between the RME-T and all the Rorschach empathy indices were non significant.

The total IAPS Valence presented a significant and positive correlation with all the six indices, as it was assumed, but this correlation seemed to be mostly due to the correlation between the Rorschach affective indices and the Positive Valence score. The correlations between the Arousal dimension and the R-EI/R-EI-Comp was not significant, contrary to the one between the R-AEI/R-AEI-Com and the positive/negative arousal. The Neutral Arousal did not present significant correlations but with the R-CEI, simultaneously the Neutral Arousal was the unique significant correlation between the Arousal dimension and the Rorschach cognitive empathy indices. The Dominance dimension presented significant positive correlation with the two Rorschach affective indices.

5.5.3.   CARD TURNING FOLLOW-UP – STUDY 3

In our first study (n = 121), the Card Turning (CT) had a negative significant correlation with the QCAE (r = -.223, p = .01). Moreover, According to the R-PAS manual, the CT represents two pattern of behaviors, what is more, almost contrasting. On one hand, CT could be a manifestation of open-mindedness, and on the other hand, and indicator of a closed approach to the environment: «Turning the card is a behavioral manipulation of the environment. It can indicate a variety of qualities, including flexibility, independence, curiosity, oppositionality, disinhibition, avoidance of a particular association, or an active or energetic approach to the task» (Meyer, Erard, Erdberg, Mihura, & Viglione, 2011, p. 331).

During the initial instructions, it is not specified if turning the card is allowed. Thus, the subject does not actually know if he can turn the card; someone turns the card because he/she already gave one or two responses, and is looking for the second/third one; someone does it immediately; someone turns the card to 360° and then gives the response with the card in the original position. Sometimes these behaviors are put in place after asking if turning the card is possible or not, and sometimes subjects just turn the card.

In light of this, in our fourth study, we detected four behaviors in order to distinguish when CT indicates flexibility and curiosity, from when it represents oppositionality, avoidance or disinhibition.

1. The subject asks – or not – if turning the card is allowed (CTQUEST vs CTNOQUEST). It is related to the context, and to the disinhibition attitude described in the CT interpretation.
2. The subject turns the card before having explored the card in the right position (1°RCT), so he/she turns the card and then gives the first response. We tried to distinguish when the CT is put in place because the subject helps him/herself looking at the card in another position (in order to give another

response), from when the card is turned without even trying to give the response in the right position, namely the position suggested by the examiner.

3. The subject turns totally the card and then interpret the card in the right position (@). If the subject turns the card but at the end gives the response in the initially position, the CT has been done "unnecessarily". More specifically, this phenomenon could indicate an oppositional manifestation, because the behavior could have been not aimed to the interpretation of the card.

Correlations between QCAE scales and the four behaviors related to CT.

	QCAE COG	QCAE AFF	QCAE TOT
	r	r	r
1°RCT	-.034	-.180	-.107
@	-.074	.063	-.023
CTQUEST	.161	.065	.144
CTNOQUEST	-.279*	-.295*	-.333**

* Correlation is significant at the 0.05 level. ** Correlation is significant at the 0.01 level.

The CTNOQUEST presented significant correlations with all the three QCAE scales, while the other three behaviors did not. However, even if it is not significant, turning the card before having explored the card in the standard position had a low negative correlation with the affective scale of the QCAE.

From the results, the "quest" behavior seemed to be the one that explained the biggest portion of QCAE scores variability. In order to deepen this hypothesis, we made a one way ANOVA between the QCAE scores in the three CT behaviors.

Multiple Comparisons of QCAE scales in NOCT, CTQUEST, and CTNOQUEST (LSD post Hoc).

QCAE scores	CT behaviors		Mean Difference	St. Err.
QCAE COG	CTNOQUEST	NOCT	-5.33	2.94
		CTQUEST	-6.02*	2.84
	NOCT	CTQUEST	-0.69	2.52
QCAE AFF	CTNOQUEST	NOCT	-4.50*	1.90
		CTQUEST	-3.44	1.83
	NOCT	CTQUEST	1.06	1.63

QCAE TOT		CTNOQUEST	NOCT	-9.83*	4.08
			CTQUEST	-9.46*	3.95
		NOCT	CTQUEST	0.37	3.50

* The mean difference is significant at the 0.05 level.

The results showed that there was no difference in QCAE scores between those who did not turn the card and those who turned it after asking if it was appropriate to do so (NOCT vs CTQUEST). Conversely, there were differences in the QCAE scores depending on the different behaviors related to CT. Indeed, those who turn the card without asking, have cognitive empathy scores lower than those who turn the card after asking if they can do so (CTNOQUEST vs CTQUEST), and affective empathy scores lower than those who did not turn the card at all (CTNOQUEST vs NOCT).

From these results, we created a dichotomous variable named CT-II, which was assigned the value "0" when CT occurred without asking if it was allowed to turn the card, and the value "1" in all the other cases (i.e. when they asked if they could turn the card before doing so, and when they just did not turn the card). Thus, we replaced the CT Raw SS variable present in the R-AEI, R-AEI-Comp, and in the R-EI, R-EI-Comp (the CT variable was not present in those selected for the Rorschach cognitive empathy indices) with the CT-II, and then we recalculated the correlations between the QCAE scores and the second version (-II) of the indices.

Pearson's Correlations between the QCAE and Rorschach indices with and without the CT-II (n = 60).

QCAE	R-AEI	R-AEI-II	R-AEI Comp	R-AEI-II Comp	R-EI	R-EI-II	R-EI Comp	R-EI-II Comp
COG	.298*	.353**	.260*	.315*	.261*	.281*	.223	.249
AFF	.476**	.484**	.476**	.493**	.458**	.455**	.465**	.470**
TOT	.430**	.473**	.403**	.450**	.395**	.408**	.372**	.392**

* Correlation is significant at the 0.05 level. ** Correlation is significant at the 0.01 level.

In general, using CT-II rather than CT in the equation to generate the REI notably improved the correlations, with the exception of the QCAE Affective Empathy scale and the REI-II.

Given that these results suggested that the CTNOQUEST could be a form of disinhibition – or inappropriate conduct – we correlated it with the Deviant Response variable (DR), which is the only one closely linked to the adherence to the request, and to adequate behaviors during the task.

Correlations between CT behaviors and DR

	DR1	DR2	DR
CT RawSS	.024	.046	.039
CTNOQUEST	.150	.342**	.255*
CTQUEST	.058	-.105	-.014

* Correlation is significant at the 0.05 level. ** Correlation is significant at the 0.01 level.

The CT Raw SS, which does not distinguish between "asked" or not "asked" permission to turning the card, did not present significant correlations with the DR variables. Conversely, CTNOQUEST did not present a significant correlation with the DR1, but it did with the DR2 – with a medium effect size –, and with the DR behavior in general.

5.6. DISCUSSION

5.6.1. DEVELOPMENT SAMPLE – STUDY 1

The R-PAS variables that correlated exclusively with the Cognitive QCAE scale were R, R8910, M, m, C', C CF, and ABS. Those that were related with the QCAE Affective scale only were CT, Art, Ay, Ex, and DR. Then, there were five variables related to both empathy components: Fi, Sy, 2, FM, and COP[20].

---

[20] R = n. of responses; R8910 = n. of responses to the 8th, 9th, and 10th cards, M = Human Movement; m = Inanimate Movement; C = Pure Color; C' = Acromatich Color; CF = Color Form; ABS = Abstraction; CT = Card Tourning; Art = Art Content; Ay = Anthropology Content; Ex = Explosion Content; DR = Devian Response; Fi = Fire Content; Sy = Synthesis; 2 = Pair; FM = Feral Movement; COP = Cooperation.

The positive correlation of the productivity variables to cognitive empathy capabilities is in line with our hypothesis. Low R can result from cognitive deficits, emotional difficulties, oppositionality; conversely, high R records – without Pulls – may suggest an imaginative, productive, spontaneous, or intelligent person. Moreover, high R8910 may include emotional situations with other people, and a general responsiveness to compelling or vibrant stimuli.

Among the administration behaviors, just CT was related to empathy, the affective one and in a negative direction. High levels of CT could be associated to flexibility, independence, curiosity, disinhibition, avoidance of a particular association, or an active or energetic approach to the task. The negative correlation with affective empathy is coherent with the disinhibition, it is counterintuitive in the other cases. However, this aspect was investigated in a specific study, and will be treated further.

Any of the white space codes resulted related to empathy; a possible explanation about this result could be linked to the meaning of SR. Indeed, SR should represent oppositionality; however, oppositionality is a voluntary and deliberate behavior, usually put in place against social rules or norms. This behavior provides for awareness – I have to know what I am fighting against – but disregard about the context. Lack of empathy is more related to the absence of awareness.

Among the content variables, Fi is positively correlated with both types of empathy. Fi is a content linked to the damage, danger and destruction, but also with passion and energy. The correlation with the empathy could be associated to a sensitivity, or a particularly attention, with the theme of the damage; indeed, this variable is present in Critical Content index (CritCont), which could be related with traumatic experiences. Moreover, the Fi is coded even if the "damage" has already occurred (e.g. "*A piece of paper, it is burnt*"), thus the connection with the cognitive empathy may arise from the capability to figure out something happened in the past.

Also Ex variable is positively related with empathy, but just with the affective one. Ex may indicate feelings of exploding (e.g. "in anger") or being out of control, as well as fear of the same, and could be related to traumatic events. With the affective empathy, there are also two correlations, but negative, namely Art and Ay. These two variables, along with ABS thematic code, form the IntCont index that represents a way of reducing immediate personal salience, emotional impact, or to make them unreal. These results are coherent with the meaning of the variables, as well as with our expectation.

We expected a positive correlation between empathy and Human Contents (H) or the sum of them (SumH). The results did not confirm this hypothesis, however the positive correlation with the M responses suggests that empathy is more related to the ability to reflect about human intentions and interactions, than with the simple perception of them.

As regard the organization and the percept of the inkblot, Sy and Pair have positive correlations with both kinds of empathy. Sy is a variable linked with the capacity to differentiate objects, and then link them in a meaningful connection with each other. In daily life, such sophistication might be manifested by breaking down information, or problems, into parts and then synthesizing them in a more complex organization. This capability is in line with the empathic process (and with our hypotheses), where the different elements of a scenario are combined in a whole and more complex understanding. Conversely, Pair responses are not interpreted on their own, Pair is a relatively obvious way of organizing the response, and it is not linked to important response-process. However, another possible interpretation could be that the correlation between the two types of empathy and COP brought with it both Sy and Pair ones. Indeed, COP is scored when a cooperative, positive, or pleasant interaction (Sy) occurred between two (Pair) objects.

Among the determinants, the positive correlation between M and the cognitive empathy was expected, and in line with the interpretation of this variable, namely the capacity for identification with others, and the capacity to reflect on life experience. Moreover, this result was in line with the previous

literature about the link between M and the Mirror Neuron System activity (Giromini, Porcelli, Viglione, Parolin, & Pineda, 2010; Pineda, Giromini, Porcelli, Parolin, & Viglione, 2011; Andò, Salatino, Giromini, Ricci, Pignolo, et al., 2015; Porcelli, & Kleiger, 2016; Giromini, Viglione, Brusadelli, Zennaro, Di Girolamo, et al. 2016; Giromini, Viglione, Pineda, Porcelli, Hubbard, et al. 2017).

The negative correlations between the cognitive empathy and C and CF variables were unexpected, but they are still coherent. The extent to which form is represented in color responses suggests cognitive organization and control: the less the form contributes to determine the response, the most the emotional experiences is unfiltered or unmodulated. If there is a general difficulty in structuring and elaborating the emotional stimuli, also the comprehension of others' experiences could be lacking (Miguel et al., 2017; Di Girolamo, Giromini, Winters, Serie, & de Ruiter, 2017).

The positive correlations between cognitive empathy and m, FM and C' are harder to interpret. In general, m is related to experiences outside one's control, FM to ideations caused by needs or instincts, while C' may indicate a dampened emotional reactivity or a generally gloomy attitude towards life. Considering these variables singularly could not be informative, however, they are the most frequent variables[21] of the Potentially Problematic Determinants index (PPD), together with Y, V, and T. This index is generally related to cognitive abilities that, under certain conditions, can become a liability; PPD can be indicative of experiences outside one's control in terms of stimulating, irritating, upsetting or pressing needs, feelings, and perturbations, but it is associated with a richness of the internal experience, heightened reactivity to the world, depth and sensitivity. The other variables of this index did not correlate with cognitive empathy scale, but this could be because of they are rarer. Moreover, FM had a positive correlation with the affective empathy too, and this could be related to the instinctive and spontaneous aspects connected with this variable.

---

[21] Averages of the PPD index determinants: FM = 3.3, C' = 1.8, m = 1.6, Y = 1.5, T = 0.7, V = 0.7 (Meyer, Viglione, Mihura, Erard, & Erdberg, 2011).

Prior to analyzing the data, we hypothesized that there would be a negative correlation between the reflection code (r) and empathy. Results showed that this correlation is negative but it is not strong enough to be interpreted, this was probably due to the low base rate of this variable.

Among the cognitive codes, DR presented a negative correlation with the affective empathy. Thought disorders have been correlated with the lack of empathy (Mitchell et al., 2012), both with the affective and with the cognitive. However, the presence of DR not necessary represents a thought disorder. DR is more related to the language, and more specifically to a difficulty to adhere to the request of the task. The corresponding behavior in daily life goes from flexibility and freedom in speaking (DR1), to a significant derailing from the context, and could impede the individuals' ability to effectively communicate with others.

In light of this, the negative correlation between affective empathy and DRs could be coherent. Moreover, these results were in line with the ones related to the CT behaviors, and will be deepen further.

COP is the only one thematic code that correlate with empathy, and it is positively related with both components. COP responses involve attributing cooperative, collaborative, synchronized, teamwork, pleasant, benevolent, or helpful qualities to interactions. High levels of COP suggest a subject who has a positive template for envisioning relationships, and reflect an interest in presenting oneself as cooperative or as responsive to positive social cues.

We hypothesized that Mutuality of Autonomy Health (MAH) would correlate positively with empathy, and that Mutuality of Autonomy Pathological (MAP), would correlate negatively. However, these variable are extremely rare, and perhaps this low base rate could explain the lack of association observed in our research.

To summarize, from the R-PAS variables, the whole empathy seems related to sensitivity to damaged objects, the capability of differentiating objects and linking them in a meaningful and complex organization, and the attitude towards cooperative, positive and pleasant interactions. More specifically, the

cognitive empathy is closely related to high cognitive resources, capability to think about others, richness of internal experience, depth and sensitivity. Alongside, cognitive empathy is in contrast with the failure in modulating and understanding one's own emotions. Finally, affective empathy is linked with instinctive and spontaneous emotions, but it is in contrast with the attitude to be uninhibited, and the tendency to use the intellectualization as defense against emotional involvement.

Among the RES variables, only three were selected: RES 4, the RES 6 and the RES 8. The correlations with the others (i.e., RES 1, RES 2, RES 3, RES 5, and RES 7) did not reach the minimum effect size.

RES 1 is coded when the respondent verbally expresses an emotional arousal due to the exposure to the card. This variable did not reached the minimum effect size with any of the two empathy scale. A possible explanation for this result could be that the emotions elicited by the card could be less related to affective empathy, or a sensitivity towards external situations, and could have to do more with idiosyncratic issues, as personal fears, lived experiences, and views of the world (e.g. I have arachnophobia so I'm scared about spiders in Card X).

RES 2 is coded when the object of the response is described as feeling or expressing emotions or feelings. In the Rorschach inkblots there are not real expressions. The absence of a minimum effect size could suggest, also in this case, that seeing expressions could be more related to the inner world of the subject that to an external input.

The results related to RES 3, RES 4, and RES 5 are mixed and controversial. RES 3[22] and RES 5[23] were intended as a different way to gather, along with RES 4, the variety of inferences about others' mental states. RES 4 is coded when the object of the response is described as having intentions, wishes, and/or desires,

---

[22] RES 3: The response includes behaviors, actions, emotions or feelings that are about to happen or be experienced

[23] RES 5: The object of the response (e.g., a person, an animal, an object) is described as being waiting for something, or thinking of someone or something

and its positive correlation with the cognitive empathy was in line with the QCAE scale from which it derives. Nonetheless, it was not the same for RES 3 and RES 5. What is more, RES 4 was conceived as a sort of "alternative version" of the RES 3, which indeed was initially intended as the closer variable to cognitive empathy capabilities. These results are possibly related to the linguistic expression used by the respondents when they describe the movement. It is not rare that people use *is going to* or *is about to* as an interchangeable manner to say that *is doing something*. Furthermore, who is doing something is already in a state of tension – as the posture is – which is already coded by the R-PAS. Similar explanations could be take into account for the RES 5.

In this view, the meaning of the positive correlation between RES 4 and cognitive empathy became more interesting. Verbalizing others' desires or wishes could be more significant than imaging what someone is going to do, with a view of awareness about others.

RES 6 is coded when the respondent emphasizes his/her own perspective, ignoring that different people might see different things. The results suggested that an excessive emphasis on an interpretation is negatively correlated with cognitive empathy, possibly because of a lack in the perspective taking capability.

RES 7 did not reached the minimum effect size, thus it was not selected. This result partially confirmed that human faces and human details could not have different meanings in terms of relational skills or representation of others. However, not even Hd or H codes presented effectively correlations with the QCAE scales; thus, probably seeing just the human, the whole figure or the details, is not directly linked with the empathy skills.

In RES 8, the respondent sees gesture of compassion, help, support or attempts to save a living being. We expected a correlation with both types of empathy, but results showed a significant positive correlation only with the affective empathy. More in general, this variable should be related to an altruistic behavior oriented to an object (human, animal or object) in difficulties or in danger scenario. If distress of others causes distress in the observer, then offering

108

help could be more oriented to alleviate observer's own distress than just the one of the object (Davis; 1994). The relationship between distress and the motivation to help someone else could explain the stronger correlation between the RES 8 and affective empathy, than with the cognitive one.

5.6.2.   VALIDATION SAMPLE – STUDY 2

The development phase produced six indices, three consisting of the R-PAS variables correlated with the QCAE scales (R-CEI, R-AEI, R-EI), and three composed by the R-PAS variables and the RES variables which correlated with the QCAE scales at r ≥ .10 (R-CEI-Comp, R-AEI-Comp, R-EI-Comp).

In general, Rorschach empathy indices showed a stronger correlation with the criterion variable in the development sample, while in the validation one they were smaller. From one hand, this was expected because variables were pooled based on the correlations in the development sample, thus it is self-evident that correlations were higher than the ones in the validation sample. Despite the *shrinkage*, the Rorschach empathy indices remained statistically significant also in the validation sample, albeit with a smaller effect size. Noteworthy, correlations in the validation sample were very informative about the role played by the RES variables in the empathy assessment: the correlations with the QCAE were equal with or without the RES variables. This pattern of findings is also confirmed by the comparisons of Rorschach empathy indices averages between groups with high and low QCAE scores. Indeed, Cohen's *d*s were in a range from medium to medium-large, which support their benefit in the assessment of high or low levels of empathy. As such, our attempt to develop new items to measure empathy within the Rorschach failed.

On the other hand, by combining existing R-PAS variables, we were able to create a new, composite score that significantly correlated with QCAE scores. Our convergent validity analyses partially confirmed this claim.

The positive correlation between the emotion dysregulation and the Rorschach affective empathy indices is in line with our hypothesis, as well as with previous literature (Miguel et al., 2016Miguel; Di Girolamo, Giromini, Winters, Serie & de Ruiter, 2017). In addition, these correlations could be explained by Ex and Fi variables being present in the R-AEI, in that they represent impulsivity and instinctual emotions. Moreover, the absence of negative correlation with Goals and Strategy scales (namely lacking in capabilities to direct emotions towards specific goals, or manage them through appropriate strategies) suggests that R-AEI is not related to an emotionality out of control, but it is more close to a spontaneous emotional flow.

ICQ is a measure of interpersonal and social skills closer to the "practical side" of the relationships, thus it has few items about the emotional component of empathy. Conversely, it is very focused on the abilities of being assertive, taking decisions, and putting in place concrete behaviors in social/relational situations. In this view, the highest correlations found with the "Comp" Rorschach empathy indices seem reasonable, because ICQ is more focused than the QCAE on the capability to assume perspective taking (RES 6) and pro-active behaviors towards others' needs and wishes (RES 4). The absence of correlations with the Negative Assertion or Conflict Management could be due to the absence of specific-items about those in our criterion variable, and thus in the selected R-PAS variables.

Negative correlations between Rorschach empathy indices and the External Oriented Thinking TAS-20 scale is coherent with the R-PAS selected variables. In fact, as noted above, C', FM, and m, are variables that take part in the PPD index, which represents depth, sensitivity, and attention to the inner world. EOT scale, by contrast, is a measure of a thinking style oriented to material aspects in daily life. The remaining scales, the Difficulty in Identifying Feelings (DIF) and the Difficulty in Describing Feelings (DDF), did not present significant correlations with Rorschach empathy indices, even if they showed r values in the expected direction. Also in this case, these results could be related to the R-PAS

selected variables not being specific for the capability to describe or identifying feelings.

The absence of significant correlations with the RME-T could be explained by the following reasons. First, the RME-T has been chosen because it is classified as a performance-based test, more specifically a perceptive one. However, even if the RME-T is not technically a self-report, the performance requested in the RME-T is quite different from the one requested in the Rorschach task. Second, the RME-T – both in literature and in our studies – has a low internal consistency, which poses questions on its validity too.

As regard the IAPS stimuli, results seemed to be fairly coherent with our hypotheses. The capability to properly consider the quality of relational experiences is directly proportional to both types of empathy capabilities, and in general with the two kinds of Rorschach empathy indices. The negative correlation between neutral arousal and R-CEI/R-CEI-Comp was expected. Indeed, results suggest that the more someone has cognitive empathy skills the less s/he will be activated by affectively neutral situations, and this could be due to an accuracy in decoding social situations as "neutrals". Similarly, a propensity to be emotionally activated and involved in social interactions is coherent with the correlations between affective empathy and arousal, as well as with the ones with dominance. In these results, the RES variables showed stronger correlations compared to the R-PAS ones. This seems to be in line with the idea that the motivation to provide help or support (RES 8) is related to the personal distress (Davis, 1994) in seeing others suffering and this claim is reinforced by the stronger correlation between the R-AEI-Comp and the negative arousal, compared to the correlation between R-CEI and the negative arousal.

The most interesting and coherent result of the whole research was probably the one about to the CTNOQUEST, which notably improved the Rorschach empathy indexes. CT is a variable related simultaneously to two features: on one side, turning the card could reveal flexibility, independence, and compliance to the task; on the other, CT could represent oppositionality, or disinhibition. This study contributes to the understanding of what part of the CT

behavior could be significant in distinguishing the different interpretations of CT. Our hypotheses were based on four distinctive behaviors, two of them were closely linked to the card position, and two to the context.

The results suggested that the oppositional or disinhibited behavior is not so much related to the card position (e.g., the subject received the card in the standard position, then he/she turns the card just to interpret it in a different position). First, because the behaviors just related to the card orientation were not statistically significant, and then because the means differences between CTNOQUEST and the other two behaviors (CTQUEST and the NOCT) were statistically significant. Crucially, there was no difference between the NOCT and CTQUEST. This pattern of findings suggests that those who turn the card after asking if this behavior is allowed have the same level of empathy of those who did not turn the card at all. Those who turn the card without caring if it is an appropriate behavior have significantly lower QCAE scores, and thus empathy level. Moreover, the mean difference between CTNOQUEST and the other two behaviors was not the same for the two types of empathy: the differences in empathy between CTNOQUEST and CTQUEST concerned cognitive empathy, whereas the differences between CTNOQUEST and NOCT concerned affective empathy.

In light of this, CTNOQUEST does not seem to be an oppositional behavior, but something more similar to the *faux pas*, the highest level of Theory of Mind (Stone, Baron-Cohen, Knight, 1998). Indeed, the subject does not fight against any rules or instructions (in that the Rorschach instructions do not specify whether turning the card is allowed or not), he/she simply puts in place a behavior without caring if it was appropriated or not. The situation could be clearer with an example: lighting a cigarette under a "no smoke" sign is an improper behavior; lighting a cigarette during a dinner with strangers *could be* an improper behavior. The first behavior is voluntarily put in place against a clear rule, and the subject lighting the cigarette made a choice; this is more similar to the behaviors related to the lack of morality or respect for the norms assessed by the MMPI or PAI. In the second case, the subject does not know if the smoke could bother the

strangers: does he/she care? If s/he does, then he/she will ask about it, if s/he does not, he/she will just light the cigarette.

Obviously, daily social situations are more complex than a Rorschach task (and lighting a cigarette during a dinner is more annoying than turning a card!) but this example helps us to explain that it is not sufficient that a behavior is unappropriated to assess the lack of empathy. What matters is how much that person is aware about what he/she is doing related to the context. More in general, how much a person is caring about the context when he or she acts.

However, the correlations with the second version of the Rorschach empathy indices (the ones with the CTNOQUEST in place of CT) are stronger than the previous ones, which supports the hypothesis that turning the card without questioning whether that behavior is or is not appropriate does reflect a lack of empathy, to some extent. Likewise, the correlation between CTNOQUEST and DR is in line with this claim, because DR is a gradual derail from the context request, a uninhibited and unchained behavior.

It seems likely that the impulse of turning the card has more to do with the affective empathy (CTNOQUEST vs NOCT) rather than with the cognitive component of empathy. More specifically, this attitude could reflect an energetic approach to the task or the curiosity associated to the CT. However, if CT is not mediated by adequate cognitive empathy capabilities (CTQUEST vs CTNOQUEST), this energetic approach could become a *faux pas*.

Some limitations associated with these studies deserve to be mentioned. First, the samples were not clinical; this limitation reduced the empathy scores variability. Second, Rorschach data were collected by graduate students and not by clinicians expert in R-PAS administration; this could have led to a less sophisticated administration and coding; however, the inter-rater reliability between the examiners was good (Pignolo, Giromini, Andò, Ghirardello, Di Girolamo, et al., 2017) and administration and coding occurred under the supervision of an R-PAS trainer. Third, some of the low correlation indices found in our studies could reflect the low association between self-reported and

performance measures in general. As such, future research with more advanced, performance-based, empathy measures would be beneficial. However, in our literature review, we could not found any solid, fully-validated, sophisticated, performance-based measures of empathy – and this is in fact one of the main reasons why we designed this research.

VI. CONCLUSION

This series of studies attempted to develop and validate a new, composite, Rorschach-based, empathy index. The roots of this line of research may be found in the relationship between the activity of the Mirror Neuron System and the M response (Giromini, Porcelli, Viglione, Parolin, & Pineda, 2010; Pineda, Gi-romini, Porcelli, Parolin, & Viglione, 2011; Andò, Salatino, Giromini, Ricci, Pignolo, et al., 2015; Porcelli, & Kleiger, 2016; Giromini, Viglione, Brusadel-li, Zennaro, Di Girolamo, et al. 2016; Giromini, Viglione, Pineda, Porcelli, Hubbard, et al. 2017). All in all, our attempt to create a Rorschach empathy index turned out to be as challenging as defining empathy is.

Taken together, the pattern of correlations produced by the selected Rorschach variables is fairly coherent with our a-priori hypotheses, as well as with the empathy-related, theoretical background reviewed in the first chapter. However, it is intriguing that empathy capabilities seemed to be scattered among various Rorschach variables and indices, and that the two resultant, Rorschach empathy indices, highly correlated with each other, as well as with the two main components of the QCAE (i.e., cognitive and affective empathy). This is probably due to the specific nature of the Rorschach task, which consists of a problem solving and not an empathy-related task. But empathy itself is a complex process too. It requires a variety of capabilities that play a key role in determining the personality of an individual. As a matter of fact, empathy is one of the four traits selected by the DSM 5 task-force to assess personality function. Thus, the distinction between the two main components of empathy could be helpful to appreciate the sophistication of some behaviors. Nonetheless, in the real world, the dichotomy between cognitive versus affective empathy likely disappears, in that these two components are strongly associated with each other.

The poor contribution of the RES variables could be attributable either to the low correlation between performance based and the self-report tests (Mihura et al., 2013), or to our incorrect operationalization of the QCAE items, and in

general of the empathy-related behaviors, into Rorschach-related, observable, behaviors. Reshaping an explicit statement to an implicit behavior could be affected by a variety of interpretative errors.

Nevertheless, this series of studies does have the potential to significantly advance the science of Rorschach-based, psychological assessment. First of all, the improvement of the Card Turning variable, and its relation with the DR, is an encouragement to deepen this aspect of CT. Aside from the importance for interpretation, CT emerged unexpectedly as the strongest variable related, negatively, to empathy. The Rorschach task is typically understood as something *going on* between the respondent and the inkblots, and the majority of the Rorschach variables have been developed, validated, and established based on the examination of the response process occurring within the mind of the examinee. That is, Rorschach codes typically are interpreted based on how the respondent processes the stimuli. However, if one aims at measuring empathy using the Rorschach, the presence of the examiner, and the interaction with him/her, likely plays a crucial role.

In the last three decades, the importance of the empathy construct has been increasing substantially, up to become one of the four components of the Alternative Dimensional Model for the personality assessment, in the last version of the DSM. However, as discussed in the third chapter, personality assessment tools are not in line with this recent necessity to investigate the patients' empathy capabilities. In return, specific self-report questionnaires for the assessment of cognitive and affective empathy had been put forward, but with the consequence of extending the assessment time, and placing empathy assessment under a high risk of the social desirability bias.

In this framework, the possibility to assess empathy with a personality test could advantage clinicians who need an initial, general, overview of the patient. Moreover, a performance-based test could bypass, or even reduce, the self-reported bias limitation.

From our results, it seems that the Rorschach test gives the possibility to distinguish the two types of empathy. In this perspective, the final evaluation of patients' empathy skills, and thus the personality assessment, could be more accurate and complete. The M variable could be a good example about the contribution of this distinction.

The R-PAS manual reports that the Human Movement (M) is a directly related indicator of the respondent's empathy level. This interpretation is mostly based on the line of research that found a relation between the M response and the Mirror Neuron System activity (Giromini, Porcelli, Viglione, Parolin, & Pineda, 2010; Pineda, Giromini, Porcelli, Parolin, & Viglione, 2011; Andò, Salatino, Giromini, Ricci, Pignolo, et al., 2015; Porcelli, & Kleiger, 2016; Giromini, Viglione, Brusadelli, Zennaro, Di Girolamo, et al. 2016; Giromini, Viglione, Pineda, Porcelli, Hubbard, et al. 2017). Thus, currently, a high number of M responses classifies someone as a very empathic person. However, our research showed that M variable seems to be more related with the cognitive empathy than with the affective one, and these results are fairly in line with the literature on the relation between cognitive empathy and MNS activity (Baird, Scheffer, & Wilson, 2011; Shamay-Tsoory, 2011). Nevertheless, cognitive empathy capabilities are not sufficient to being an actual empathetic subject, therefore the presence of cognitive empathy capabilities is not sufficient to exclude a Narcissistic or an Antosocial Personality Disorder (Dolan & Fullam, 2004; Ritter, Lorig, Laurent, & Matthews, 2004; Stellwagen & Kerig, 2013; Marcoux, Michon, Lemelin, Voisin, Vachon-Presseau et al., 2014). To summarize, at the current state of the Rorschach interpretation, it may happen that someone with a NPD is classified as a person who has a high level of empathy. Of course, the interpretation of the Rorschach protocol is not merely based on one variable, and there are other indexes that can point out the presence of disorders like NPD or AsPD. Despite this, the issue about the empathy remains, and it becomes more crucial when the personality impairment is not so severe to be easily recognizable from specific indices.

Commonly, empathy is considered a "the more, the better" capability. Namely, what clinicians are interested in is the lack of empathy. This is not properly true for the affective component. Recently, high levels of emotion contagion, which is a crucial affective empathy component, had been associated with pathological or malaise conditions (Weisbuch et al., 2011; Horan et al., 2015; Miguel, 2016; Di Girolamo et al., 2017). The possibility to assess affective empathy singularly, and also not just in terms of lacking, could be informative about how much the person gets emotionally involved when is exposed to others' feelings.

We are still far from an actual Rorschach empathy index, or from its ultimate version. However, this research may have paved the road.

REFERENCES

Abu-Akel, A., & Abushua'leh, K. (2004). 'Theory of Mind' In Violent and Nonviolent Pa- Tients with Paranoid Schizophrenia. Schizophrenia Research, 69, 45–53.

Acklin, M. W. (1999). Behavioral science foundations of the Rorschach Test: Research and clinical applications. Assessment, 6, 319-326.

Acklin, M. W., McDowell, C. J., Verschell, M. S., & Chan, D. (2000). Interobserver agreement, intraobserver reliability, and the Rorschach Comprehensive System. Journal of Personality Assessment, 74, 15-47.

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Pub.

American Psychological Association. (23/10/2017). American Psychological Association.

Http://Www.Apa.Org/Ed/Graduate/Specialize/Personality.Aspx

Ando, A., Salatino A., Giromini, L., Ricci, R., Pignolo, C., Cristofanelli, S., Ferro, L., Viglione, D.J., & Zennaro, A. (2015). Embodied Simulation and Ambiguous Stimuli: The Role of the Mirror Neuron System. Brain Research. 1629. 135-142. Http://Dx.Doi.Org/10.1016/J.Brainres.2015.10.025.

Aronow, E. (2001). CS norms, psychometrics, and possibilities for the Rorschach technique. Clinical Psychology: Science and Practice, 8, 383-385.

Badia, X., Gutiérrez, F., Wiklund, I., & Alonso, J. (1996). Validity and Reliability of the Spanish Version of the Psychological General Wellbeing Index. Quality Of Life Research, 5(1), 101–108.

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The Twenty-Item Toronto Alexithymia Scale-I. Item Selection and Cross-Validation of the Factor Structure. Journal of Psychosomatic Research, 38(1), 23-32. Https://Doi.Org/10.1016/0022-3999(94)90005-1.

Baird, A. D., Scheffer, I. E., & Wilson, S. J. (2011). Mirror neuron system involvement in empathy: a critical look at the evidence. Social neuroscience, 6(4), 327-335.

Bandura, A. (1954a). The Rorschach White Space Response and "Oppositional" Behavior. Journal of Consulting Psychology, 18, 17–21. Doi: 10.1037/H0056911.

Bandura, A. (1954b). The Rorschach White Space Response and Perceptual Reversal. Journal of Experimental Psychology, 48, 113– 118. Doi: 10.1037/H0063608.

Baron-Cohen S., Whellwright 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 Syndrom Or High-Functioning Autism. Journal Of Child Psychology. Vol 42(2), 241-251.

Baron-Cohen, S. (2005). The Empathizing System: A Revision of the 1994 Model of the Mindreading System. In: Ellis, B. and Bjorklund, D. (Eds.), Origins of the Social Mind. Guilford, New York, Pp. 468–492.

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(2), 163-175.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does The Autistic Child Have A "Theory Of Mind"? Cognition, 21(1), 37-46.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, A. T. (1997). Is There a" Language of the Eyes"? Evidence from Normal Adults, And Adults with Autism or Asperger Syndrome. Visual Cognition, 4(3), 311-331.

Bateson, C. D. (2009). These Things Called Empathy: Eight Related But Distinct Phenomena. In J. Decety & W. Ickes (Eds), the Social Neuroscience of Empathy (Pp. 3-15). Cambridge (Mass.): MIT Press.

Ben-Porath, Y. S. (2012). Interpreting the MMPI-2-RF. U of Minnesota Press.

Bentler, P. M., & Bonett, D. G. (1980). Significance Tests and Goodness of Fit in the Analysis of Covariance Structures. Psychological Bulletin, 88(3), 588. DOI: 10.1037/0033-2909.88.3.588.

Bethlehem, J. (2010). Selection Bias in Web Surveys. International Statistical Review, 78(2), 161-188. DOI 10.1111/J.1751-5823.2010.00112.

Bjureberg, J., Ljótsson, B., Tull, M. T., Hedman, E., Sahlin, H., Lundh, L. G., Bjärehed, J., Dilillo, D., Messman-Moore, T., Gumpert, C.H., & Gratz, K. L. (2016). Development and Validation of a Brief Version of the Difficulties in Emotion Regulation Scale: The DERS-16. Journal of Psychopathology and Behavioral Assessment, 38(2), 284-296. DOI 10.1007/S10862-015-9514-X.

Blair, R. J. R. (2005). Responding To the Emotions of Others: Dissociating Forms of Empathy through the Study of Typical and Psychiatric Populations. Consciousness and Cognition: An International Journal, 14, 698–718. Https://Doi.Org/10.1016/J.Concog.2005.06.004.

Bohart, A. C., & Greenberg, L. S. (1997). Empathy and Psychotherapy: An Introductory Overview. In A. C. Bohart & L. S. Greenberg (Eds.), Empathy Reconsidered New Directions in Psychotherapy (Pp. 1–39). Washington, DC: American Psychological Association.

Bornstein, R. F. (1996). Construct Validity Of The Rorschach Oral Dependency Scale: 1967–1995. Psychological Assessment, 8, 200–205. Doi:10.1037/1040-3590.8.2.200.

Bornstein, R.F. (2001). Clinical utility of the Rorschach Inkblot Method: Reframing the debate. Journal of Personality Assessment, 77, 39 - 47.

Bornstein, R.F., & Masling, J. M. (Eds.). (2005). Scoring the Rorschach: Seven validated systems. Mahwah, NJ: Lawrence Erlbaum Associates.

Bressi, C., Taylor, G., Parker, J., Bressi, S., Brambilla, V., Aguglia E., and others. Cross Validation of the Factor Structure of the 20-Item Toronto Alexithymia Scale: An Italian Multicenter Study. J Psychosom Res 1996;41:551–9 Brown, M. M., Thibodeau, R. B., Pierucci, J. M., & Gilpin, A. T. (2017). Supporting the Development of Empathy: The Role of Theory of Mind and Fantasy Orientation. Social Development, (November), 1–14.

Brown, M. M., Thibodeau, R. B., Pierucci, J. M., & Gilpin, A. T. (2017). Supporting the Development of Empathy: The Role of Theory of Mind and Fantasy Orientation. Social Development, (November), 1–14. Https://Doi.Org/10.1111/Sode.12232.

Browne, M. W., & Cudeck, R. (1993). Alternative Ways of Assessing Model Fit. Sage Focus Editions, 154, 136-136.

Buhrmester, D., Furman, W., Wittenberg, M. T., & Reis, H. T. (1988). Five Domains of Interpersonal Competence in Peer Relationships. Journal of Personality and Social Psychology, 55(6), 991–1008. Doi:10.1037/0022-3514.55.6.991.

Burns, B., & Viglione Jr, D. J. (1996). The Rorschach Human Experience Variable, interpersonal relatedness, and object representation in nonpatients. Psychological Assessment, 8, 92-99.

Burns, B., & Viglione, D. J. (1997). "The Rorschach Human Experience Variable, interpersonal relatedness, and object representations in nonpatients": Correction. Psychological Assessment, 9, 82.

Caprara, G. V., Barbaranelli, C., Hahn, R., & Comrey, A. L. (2001). Factor Analyses of the NEO-PI-R Inventory and the Comrey Personality Scales in Italy and the United States. Personality and Individual Differences, 30, 217–228. Https://Doi.Org/10.1016/S0191-8869(00)00030-1.

Cho, Y., & Hong, S. (2013). The New Factor Structure of the Korean Version of the Difficulties in Emotion Regulation Scale (K-DERS) Incorporating Method Factor. Measurement and Evaluation in Counseling and Development, 46(3), 192-201. Https://Doi.Org/10.1177/0748175613484033.

Chronbach, L. J. (1949). I Test Psicologici. La Misura Degli Interessi E Della Personalità [Essentials Of Psychological Testing]. Giunti, Firenze.

Cicchetti, D. V. (1994). Guidelines, Criteria, and Rules Of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. Psychological Assessment, 6, 284–290. DOI: 10.1037/1040-3590.6.4.284.

Cikara, M., & Fiske, S. T. (2013). Their Pain, Our Pleasure: Stereotype Content and Schadenfreude. Annals of the New York Academy of Sciences, 1299(1), 52-59.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences, 2nd Ed. Hillsdale, New Jersey: Erlbaum.

Costa, P. T., & Mccrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual. Odessa: Psychological Assessment Resources.

Damasio, A.R. (1994). Descartes' Error: Emotion, Reason, and the Human Brain. Putnam, New York.

Dan-Glauser, E. S., & Scherer, K. R. (2013). Swiss Journal of Psychology, 72(1), 5.

Danziger, K. (1997). Naming the Mind. London: Sage.

Davis M.H. (1983). Measuring Individual Differences In Empathy: Evidence For A Multidimensional Approach. Journal Of Personality And Social Psychology. Vol. 44(1), 113-126.

Davis, M. H. (1980). A Multidimensional Approach to Individual Differences in Empathy. JSAS Catalog of Selected Documents in Psychology 10, 85.

Davis, M. H. (1994). Empathy: A Social Psychological Approach. Boulder, CO: Westview.

Dawes, R. M. (1999). Two methods for studying the incremental validity of a Rorschach variable. Psychological Assessment, 11, 297-302.

De Campora, G., Giromini, L., Larciprete, G., Volsi, V. L., & Zavattini, G. C. (2014). The Impact of Maternal Overweight and Emotion Regulation on Early Eating Behaviors. Eating Behaviors, 15(3), 403-409.

De Koninck, J. M. & Crabbe-Decleve, G. (1971). Field Dependence and Rorschach White-Space Figure-Ground Reversal Responses. Perceptual and Motor Skills, 33, 1191–1194. Doi:10.2466/Pms.1971.33.3f.1191.

Demidova K. Yu., Dubovik V.V., Kravchenko V.I., Makarchouk N.E. (2014). EEG Activity During Viewing Of Neutral And Emotionally Colored Pictures By Men And Women With Different Levels Of Empathy. Neurophysiology. Vol. 46(2).

Derntl B., Finkelmeyer A., Toygar T.K., Hulsmann A., Schneider F., Falkenberg D.I., Habel U. (2009). Generalized Deficit In All Core Components Of Empathy In Schizophrenia. Schizophrenia Research. Vol. 108(1–3), Pp.197–206.

Di Girolamo, M., Giromini, L., Winters, C. L., Serie, C. M., & de Ruiter, C. (2017). The Questionnaire of Cognitive and Affective Empathy: A Comparison between Paper-and-Pencil versus Online Formats in Italian Samples. Journal of personality assessment, 1-12.

Dodich A., Cerami C., Canessa N., Crespi C., Iannaccone S., Marcone A., Realmuto S., Lettieri G., Perani D., Cappa S.F. (2015). A novel task assessing intention and emotion attribution: Italian standardization and normative data of the Story-based Empathy Task. Neurological Science. 36:1907–1912.

Dolan, M., & Fullam, R. (2004). Theory Of Mind And Mentalising Ability In Antisocial Personality Disorders With And Without Psychopathy. Psychological Medicine, 34, 1093–1102. Https://Doi.Org/10.1017/S0033291704002028.

Dumitrascu, N., Mihura, J. L., & Meyer, G. J. (2010, March). Scoring Guidelines and Norms for Space Reversal (Sr) and Space Integration (Si) Rorschach Responses. In S. L. Horn (Chair), Advances In Rorschach Performance Assessment: Part II. Paper Presented At The Annual Meeting Of The Society For Personality Assessment, San Jose, CA, March 27.

Dupuy, H. J. (1977). The General Well-Being Schedule. In I. Mcdowell & C. Newell (Eds.), Measuring Health: A Guide to Rating Scales and Questionnaire (2nd Ed., Pp. 206–213). USA: Oxford University Press.

Dupuy, H. J. (1984). The Psychological General Well-Being (PGWB) Index. In N. Wenger (Ed.), Assessment of Quality Of Life in Clinical Trials of Cardiovascular Therapies (Pp. 170–183). New York: Le Jacq.

Dziobek I., Heekeren H.R. (2008). Dissociation Of Cognitive And Emotional Empathy In Adults With Asperger Syndrome Using The Multifaceted Empathy Test (MET). Journal Of Autism And Developmental Disorders.

Eisenberg, N., & Miller, P. (1987). The Relation of Empathy to Prosocial and Related Behavior. Motivation and Emotion, 14, 113–149.

Ekman P., Friesen W.V. (1975). Unmasking The Face. Englewood Cliffs, New Jersey: Spectrum-Prentice Hall.

Ellenberger, H. (1989). The Life and the Work of Hermann Rorschach (1884 - 1922). In J. E. Ann M. O'Roark, History and Directory: Society for Personality Assessment Fiftieth Anniversary (P. 192). London: Routledge.

Erard, R. E., Meyer, G. J., & Viglione, D. J. (2014). Setting the Record Straight: Comment on Gurley, Piechowski, Sheehan, and Gray (2014) On the Admissibility of the Rorschach Performance Assessment System (R-PAS) In Court. Psychological Injury and Law, 7(2), 165-177.

Exner Jr, J. E. (1991). Wiley series on personality processes. The Rorschach: A comprehensive system: Interpretation.

Exner, J. E. (1969). The Rorschach Systems. New York, NY: Grune & Stratton.

Exner, J. E. (1974). Rorschach: A Comprehensive System, Vol. 1. New York: Wiley-Interscience.

Exner, J. E. (1995). Comment on "Narcissism in the Comprehensive System for the Rorschach". Clinical Psychology: Science and Practice, 2, 200-206.

Exner, J. E. (1996). A comment on "The Comprehensive System for the Rorschach: A Critical Examination". Psychological Science, 7, 11-13.

Exner, J. E. (2001). A comment on "The misperception of psychopathology: Problems with the norms of the Comprehensive System for the Rorschach". Clinical Psychology: Science and Practice, 8, 386-388.

Feather, N. T., & Nairn, K. (2005). Resentment, Envy, Schadenfreude, and Sympathy: Effects of Own and Other's Deserved or Undeserved Status. Australian Journal of Psychology, 57(2), 87-102.

Force, P. T. (2006). Psychodynamic Diagnostic Manual.

Frith, U., & Frith, C. D. (2003). Development and Neurophysiology of Mentalizing. Philosophical Transactions of the Royal Society B: Biological Sciences, 358(1431), 459-473.

Gacono, C. B., Bannatyne-Gacono, L., Meloy, J. R. & Baity, M. R. (2005). The Rorschach Extended Aggression Scores. Rorschachiana: Journal of the International Society for the Rorschach, 27, 164–190. Doi:10.1027/1192-5604.27.1.164.

Gacono, C. B., Loving, J. L., & Bodholdt, R. H. (2001). The Rorschach and psychopathy: Toward a more accurate understanding of the research findings. Journal of Personality Assessment, 77, 16 - 38.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A Unifying View of the Basis of Social Cognition. Trends Cogn. Sci., 8: 396–403.

Ganellen, R. J. (1996). Comparing the diagnostic efficiency of the MMPI, MCMI-II, and Rorschach: A Review. Journal of Personality Assessment, 67, 219-243.

Ganellen, R. J. (2001). Weighing evidence for the Rorschach's validity: A response to Wood et al. (1999). Journal of Personality Assessment, 77, 1-15.

Garb, H. N. (1999). Call for a moratorium on the use of the Rorschach Inkblot Test in clinical and forensic settings. Assessment, 6, 313-317.

Garb, H. N., & Grove, W. M. (2005). On the merits of clinical judgment: Comment. American Psychologist, 60, 658-659.

Garb, H. N., Florio, C. M., & Grove, W. M. (1998). The validity of the Rorschach and the Minnesota Multiphasic Personality Inventory: Results from meta-analyses. Psychological Science, 9, 402-404.

Garb, H. N., Florio, C. M., & Grove, W. M. (1999). The Rorschach controversy: Reply to Parker, Hunsley, and Hanson. Psychological Science, 10, 293-294.

Garb, H. N., Wood, J. M., Lilienfeld, S. O., & Nezworski, M. T. (2005). Roots of the Rorschach controversy. Clinical Psychology Review, 25, 97-118.

Garb, H. N., Wood, J. M., Nezworski, M. T., Grove, W. M., & Stejskal, W. J. (2001). Toward a resolution of the Rorschach controversy. Psychological Assessment, 13, 433-448.

Garfield, S. L. (2000a). The Rorschach test in clinical diagnosis - a brief commentary. Journal of Clinical Psychology, 56, 431-434.

Garfield, S. L. (2000b). The Rorschach test in clinical diagnosis. Journal of Clinical Psychology, 56(3), 387-393.

Geisinger, K. F. (2003). Testing and Assessment in Cross-Cultural Psychology. In J. R. Graham, J. A. Naglieri, & I. B. Weiner (Eds.), Handbook Of Psychology (Vol. 10). Assessment Psychology (Pp.95–118). New Jersey: Wiley.

Gibbons, S. B. (2011). Understanding Empathy as a Complex Construct: A Review of the Literature. Clinical Social Work Journal, 39(3), 243-252.

Giovannini, C., Giromini, L., Bonalume, L., Tagini, A., Lang, M., & Amadei, G. (2014). The Italian Five Facetmindfulness Questionnaire: A Contribution to Its Validity and Reliability. Journal of Psychopathology and Behavioral Assessment, 36, 415–423. Doi:10.1007/S10862-013-9403-0.

Giromini, L., Brusadelli, E., Di Noto, B., Grasso, R., & Lang, M. (2015). Measuring Psychological Mindedness: Validity, Reliability, and Relationship with Psychopathology of an Italian Version of the Balanced Index of Psychological Mindedness. Psychoanalytic Psychotherapy, 29(1), 70–87. Doi:10.1080/02668734.2015.1006666.

Giromini, L., De Campora. G., Brusadelli. E., D'Onofrio. E., Zennaro, A., Zavattini. G. C. & Lang. M. (2016). Validity and Reliability of the Interpersonal

Competence Questionnaire (ICQ): Empirical Evidence from an Italian Study. Journal of Psychopathology and Behavioral Assessment. 38(1). 113–123.

Giromini, L., Porcelli, P., Viglione, D. J., Parolin, L., & Pineda, J. A. (2010). The Feeling of Movement: EEG Evidence for Mirroring Activity during the Observations of Static. Ambiguous Stimuli in the Rorschach Cards. Biological Psychology. 85. 233-241. Doi:10.1016/J.Biopsycho.2010.07.008

Giromini, L., Velotti, P., De Campora, G., Bonalume, L., & Zavattini, G. C. (2012). Cultural Adaptation of the Difficulties in Emotion Regulation Scale: Reliability and Validity of an Italian Version. Journal Of Clinical Psychology, 68(9), 989–1007. Doi:10.1002/Jclp.21876.

Giromini, L., Viglione, D. J., Brusadelli, E., Zennaro, A., Di Girolamo, M., & Porcelli, P. (2016). The Effects of Neurological Priming on the Rorschach: A Pilot Experiment on the Human Movement Response. Rorschachiana. 37. 58-73. Doi:10.1027/1192-5604/A000077.

Giromini, L., Viglione, D.J., Pineda, J.A., Porcelli, P., Hubbard, D., Zennaro, A., & Cauda, F. (2017). Human Movement Responses to the Rorschach and Mirroring Activity: An fMRI Study. Assessment. 0. 1-14. Https://Doi.Org/10.1177/1073191117731813.

Glass, S. J., & Newman, J. P. (2006). Recognition Of Facial Affect In Psychopathic Of- Fenders. Journal of Abnormal Psychology, 115(4), 815–820

Goldman, A. (2005). Mirror systems, social understanding and social cognition. Interdisciplines.

Goldman, A. (2006) Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading. Oxford University Press, New York.

Gordon, R.M. (2005). Simulation and Systematic Errors in Prediction. Trends Cogn. Sci., 9: 361–362.

Gratz, K. L., & Roemer, L. (2004). Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure, and Initial Validation of the Difficulties in Emotion Regulation Scale. Journal of Psychopathology and Behavioral Assessment, 26(1), 41-54.

Grossi, E., Masconi, P., Groth, N., Nievo, M., & Apolone, G. (2002). Il Questionario Psychological General Well-Being. Versione Italiana. Versione 0.2. Milan: Marzo.

Grove, W. M., & Barden, R. C. (1999). Protecting the integrity of the legal system: The admissibility of testimony from mental health experts under Daubert/Kumho analyses. Psychology, Public Policy, and Law, 5, 224-242.

Hall, R. J., Snell, A. F., & Foust, M. S. (1999). Item Parceling Strategies In SEM: Investigating The Subtle Effects Of Unmodeled Secondary Constructs. Organizational Research Methods, 2(3), 233-256.

Hamel, M., Gallagher, S., & Soares, C. (2001). The Rorschach: Here we go again. Journal of Forensic Psychology Practice, 1, 83-92

Harkness, K. L., Jacobson, J. A., Duong, D., & Sabbagh, M. A. (2010). Mental State Decoding In Past Major Depression: Effect of Sad Versus Happy Mood Induction. Cognition and Emotion, 24(3), 497-513.

Harris, S. T., & Picchioni, M. M. (2013). A Review of the Role of Empathy in Violence Risk in Mental Disorders. Aggression and Violent Behavior, 18(2), 335-342.

Hastings, Jr. C., Mosteller, F., Tukey, J.W., & Winsor, C.P. (1947). Low Moments for Small Samples: A Comparative Study of Order Statistics. The Annals of Mathematical Statistics. 413-426.

Hastings, M. E., Tangney, J. P., & Stuewig, J. (2008). Psychopathy and Identification of Facial Expressions of Emotion. Personality and Individual Differences, 44, 1474–1483.

Hatfield, E., Rapson, R.L., & Le, Y.C.L. (2011). Emotional Contagion and Empathy. In Decety, J., & Ickes, W. (2011). The Social Neuroscience of Empathy. MIT Press.

Henry, J. D., Bailey, P. E., & Rendell, P. G. (2008). Empathy, Social Functioning and Schizotypy. Psychiatry Research, 160(1), 15-22.

Hiller, J. B., Rosenthal, R., Bornstein, R. F., Berry, D. T. R., & Brunell-Neuleib, S. (1999). A comparative meta-analysis of Rorschach and MMPI validity. Psychological Assessment, 11, 278-296.

Hoffman, M.L. (1984) Interaction of Affect and Cognition in Empathy. In: Izard, C., Kagan, J. and Zajonc, R. (Eds.), Emotions, Cognition, and Behavior. Cambridge University Press, New York, Pp. 103–131.

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics, 65-70.

Horan, W. P., Reise, S. P., Kern, R. S., Lee, J., Penn, D. L., & Green, M. F. (2015). Structure and Correlates of Self-Reported Empathy in Schizophrenia. Journal of Psychiatric Research, 66, 60-66.

Hu, L. T., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1-55.

Hunsley, J., & Bailey, J. M. (1999). The clinical utility of the Rorschach: Unfulfilled promises and an uncertain future. Psychological Assessment, 11, 266-277.

Hunsley, J., & Bailey, J. M. (2001). Whither the Rorschach? An analysis of the evidence. Psychological Assessment, 13, 472-485.

Hunsley, J., & Giulio, G. D. (2001). Norms, norming, and clinical assessment. Clinical Psychology: Science and Practice, 8, 378-382.

Jahoda, G. (2005). Theodor Lipps and the Shift from "Sympathy" To "Empathy". Journal of the History of the Behavioral Sciences, Vol. 41(2), 151–163. DOI 10.1002 /Jhbs.20080.

John, O. P., & Soto, C. J. (2007). The Importance of Being Valid. Handbook of Research Methods in Personality Psychology, 461-494.

Jolliffe, D., & Farrington, D. P. (2004). Empathy and Offending: A Systematic Review and Meta-Analysis. Aggression and Violent Behavior, 9(5), 441-476.

Jonason, P. K., & Krause, L. (2013). The Emotional Deficits Associated With The Dark Triad Traits: Cognitive Empathy, Affective Empathy, And Alexithymia. Personality and Individual Differences, 55(5), 532-537.

Jöreskog, K. G., & Sörbom, D. (2005). LISREL 8.72: Interactive LISREL for MS Windows. Lincolnwood: Scientific Software International, Inc.

Kanske, P., Böckler, A., Trautwein, F. M., & Singer, T. (2015). Dissecting the Social Brain: Introducing the Empatom to Reveal Distinct Neural Networks and

Brain-Behavior Relations for Empathy and Theory of Mind. Neuroimage, 122, 6–19. Https://Doi.Org/10.1016/J.Neuroimage.2015.07.082

Katko, N. J., Meyer, G. J., Mihura, J. L. & Bombel, G. (2010). A Principal Components Analysis of Rorschach Aggression and Hostility Variables. Journal of Personality Assessment, 92, 594–598. Doi:10.1080/00223891.2010.513309.

Kenny D.A., Mohr C.D., Levesque M.J. (2001). A Social relations variance partitioning of dyadic behavior. Psychological Bulletin. Vol. 127(1), 128-141.

Khorashad, B. S., Baron-Cohen, S., Roshan, G. M., Kazemian, M., Khazai, L., Aghili, Z., and others, & Afkhamizadeh, M. (2015). The "Reading the Mind in the Eyes" Test: Investigation of Psychometric Properties and Test–Retest Reliability of the Persian Version. Journal of Autism and Developmental Disorders, 45(9), 2651-2666.

Lang, P. J. (2005). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical report.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International Affective Picture System (IAPS): Technical Manual and Affective Ratings. NIMH Center for the Study of Emotion and Attention, 39-58.

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring Empathy: Reliability and Validity of the Empathy Quotient. Psychological Medicine, 34(05), 911-920.

Leiberg, S., & Anders, S. (2006). The Multiple Facets of Empathy: A Survey of Theory and Evidence. Progress In Brain Research, 156, 419-440.

Lerner, P. M. (2000). A nonreviewer's comment: On the Rorschach and baseball. Journal of Clinical Psychology, 56, 439.

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. Psychological Science in the Public Interest, 1, 27-66

Lipps, T. (1903). Einfühlung, Innere Nachahmung Und Organempfindungen.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To Parcel Or Not To Parcel: Exploring The Question, Weighing The Merits. Structural Equation Modeling, 9(2), 151-173.

Lyons C., Hazler R.J. (2002). The Influence Of Student Development Level On Improving Counselor Student Empathy. Counselor Education And Supervision. Vol 42.

Majorek, K., Wolfkühler, W., Küper, C., Saimeh, N., Juckel, G., & Brüne, M. (2009). Theory of Mind and Executive Functioning In Forensic Patients with Schizophrenia. Journal of Forensic Sciences, 54(2), 469–473.

Marcoux, L. A., Michon, P. E., Lemelin, S., Voisin, J. A., Vachon-Presseau, E., & Jackson, P. L. (2014). Feeling but Not Caring: Empathic Alteration in Narcissistic Men with High Psychopathic Traits. Psychiatry Research: Neuroimaging, 224(3), 341-348.

Markon, K. E., Chmielewski, M. & Miller, C. J. (2011). The Reliability and Validity of Discrete and Continuous Measures of Psychopathology: A Quantitative Review. Psychological Bulletin, 137, 856–879. Doi: 10.1037/A0023678

Mccrae, R. R., & Costa, P. T. (2004). A Contemplated Revision of the NEO Five-Factor Inventory. Personality and Individual Differences, 36(3), 587–596. Doi: 10.1016/S0191-8869(03)00118-1.

Mehrabian A., Epstein N. (1972). A Measure Of Emotional Empathy. Journal Of Personality. Vol. 40(4), 525-543.

Meyer, G. J. (1997a). Assessing reliability: Critical corrections for a critical examination of the Rorschach Comprehensive System. Psychological Assessment, 9, 480-489.

Meyer, G. J. (1997b). Thinking clearly about reliability: More critical corrections regarding the Rorschach Comprehensive System. Psychological Assessment, 9, 495-498.

Meyer, G. J. (2000). On the science of Rorschach research. Journal of Personality Assessment, 75, 46 - 81.

Meyer, G. J. (2001a). Erratum: Correction to "On the science of Rorschach research". Journal of Personality Assessment, 77, 580-580.

Meyer, G. J. (2001b). Evidence to correct misperceptions about Rorschach norms. Clinical Psychology: Science and Practice, 8, 389-396.

Meyer, G. J., & Archer, R. P. (2001). The hard science of Rorschach research: What do we know and where do we go? Psychological Assessment, 13, 486-502.

Meyer, G. J., & Eblin, J. J. (2012). An Overview of the Rorschach Performance Assessment System (R-PAS). Psychological Injury and Law, 5(2), 107-121.

Meyer, G. J., Erard, R. E., Erdberg, P., Mihura, J. L., & Viglione, D. J. (2011). Rorschach Performance Assessment System: Administration, Coding, Interpretation, And Technical Manual. LLC.

Miguel, F. K., Giromini, L., Colombarolli, M. S., Zuanazzi, A. C., & Zennaro, A. (2016). A Brazilian Investigation of the 36-And 16-Item Difficulties in Emotion Regulation Scales. Journal of Clinical Psychology, Advance Online Publication.

Mihura, J. L., Meyer, G. J., Bombel, G., & Dumitrascu, N. (2015). Standards, accuracy, and questions of bias in Rorschach meta-analyses: Reply to Wood, Garb, Nezworski, Lilienfeld, and Duke (2015). Psychological Bulletin. 141, 250-260.

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the comprehensive system. Psychological Bulletin, 139, 548-605.

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2016). On conducting construct validity meta-analyses for the Rorschach: A Reply to Tibon-Czopp and Zeligman (2016). Journal of Personality Assessment, 98, 343-350.

Mitchell, L. J., Gumley, A., Reilly, E. S., Macbeth, A., Lysaker, P., Carcione, A., Et Al. (2012). Metacognition In Forensic Patients With Schizophrenia and a History of Inter- Personal Violence: an Exploratory Study. Psychosis, 4(1), 42–51.

Mitsopoulou, E., Kafetsios, K., Karademas, E., Papastefanakis, E., & Simos, P. G. (2013). The Greek Version Of The Difficulties In Emotion Regulation Scale: Testing the Factor Structure, Reliability and Validity In an Adult Community Sample. Journal of Psychopathology and Behavioral Assessment, 35(1), 123-131.

Morey, L. C. (2007). Personality Assessment Inventory (PAI). Professional Manual. 2nd Edition. John Wiley & Sons, Inc..

Munro, G. E. S., Dywan, J., Harris, G. T., Mckee, S., Unsal, A., & Segalowitz, S. J. (2007). ERN Varies With Degree Of Psychopathy In An Emotion Discrimination Task. Biological Psychology, 76(1–2), 31–42.

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo Study Investigating The Impact Of Item Parceling On Measures Of Fit In Confirmatory Factor Analysis. Educational and Psychological Measurement, 63(5), 729-757.

Naughton, M. J., & Wiklund, I. (1993). A Critical Review of Dimension-Specific Measures of Health-Related Quality of Life in Cross-Cultural Research. Quality Of Life Research, 2(6), 397-432.

Nelson, W. D. (1954). An Evaluation of the White Space Response on the Rorschach as Figure-Ground Reversal and Intellectual Opposition. (Unpublished Doctoral Dissertation). Michigan State University, East Lansing, MI.

Nezworski, M. T., & Wood, J. M. (1995). Narcissism in the Comprehensive System for the Rorschach. Clinical Psychology: Science and Practice, 2, 179-199.

Ofir-Eyal, S., Hasson-Ohayon, I., & Kravetz, S. (2014). Affective and Cognitive Empathy and Social Quality of Life in Schizophrenia: A Comparison between a Parallel Process Model and an Integrative Meditation Model. Psychiatry Research, 220(1), 51-57.

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R. D. (2015). A Psychometric Analysis Of The Reading The Mind In The Eyes Test: Toward A Brief Form For Research And Applied Settings. Frontiers in Psychology, 6, 1503.

Parker, K. C. H., Hunsley, J., & Hanson, R. K. (1999). Old wine from old skins sometimes tastes like vinegar: A response to Garb, Florio, and Grove. Psychological Science, 10, 291-292.

Pignolo, C., Giromini, L., Ando', A., Ghirardello, D., Di Girolamo, M., Ales, F., & Zennaro, A. (2017). An Interrater Reliability Study of Rorschach Performance Assessment System (R–PAS) Raw and Complexity-Adjusted Scores. Journal of Personality Assessment, 1-7.v DOI 10.1080/00223891.2017.1296844.

Premack, D., & Woodruff, G. (1978). Does The Chimpanzee Have a Theory Of Mind? Behavioral and Brain Sciences, 1(4), 515-526.

Preston, S.D., & De Waal, F.B. (2002). Empathy: Its Ultimate and Proximate Bases. Behav. Brain Sci., 25, 1–20.

Preti, A., Vellante, M., Baron-Cohen, S., Zucca, G., Petretto, D. R., & Masala, C. (2011). The Empathy Quotient: A Cross-Cultural Comparison of the Italian Version. Cognitive Neuropsychiatry, 16(1), 50-70.

Ragsdale, G., & Foley, R. A. (2011). A Maternal Influence on Reading the Mind in the Eyes Mediated By Executive Function: Differential Parental Influences on Full and Half-Siblings. Plos One, 6(8), E23236.

Rapaport, D., Gill, M., & Schafer, R. (1945). Diagnostic Psychological Testing (Vol. 1).

Reese, J. B., Viglione, D. J., & Giromini, L. (2014). A Comparison between Comprehensive System and an Early Version of the Rorschach Performance Assessment System Administration with Outpatient Children and Adolescents. Journal of Personality Assessment, 96(5), 515-522.

Reniers R., Corcoran R., Drake J., Shryane N., Vollm B. (2011). The QCAE: A Questionnaire Of Cognitive And Affective Empathy. Journal Of Personality Assessment. 93(1):84-95.

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

Richell, R. A., Mitchell, D. G., Newman, C., Leonard, A., Baron-Cohen, S., & Blair, R. J. (2003). 'Theory Of Mind And Psychopathy: Can Psychopathic Individuals Read The 'Language Of The Eyes'? Neuropsychologia, 41, 523–526.

Ritter, K., Dziobek, I., Preißler, S., Rüter, A., Vater, A., Fydrich, T., et al. & Roepke, S. (2011). Lack of empathy in patients with narcissistic personality disorder. Psychiatry research, 187(1), 241-247.

Ritter, P., Lorig, K., Laurent, D., & Matthews, K. (2004). Internet versus Mailed Questionnaires: A Randomized Comparison. Journal of Medical Internet Research, 6(3), E29.

Riva, G., Teruzzi, T., & Anolli, L. (2003). The Use of the Internet in Psychological Research: Comparison of Online and Offline Questionnaires. Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 6(1), 73–80.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996) Premotor Cortex and the Recognition of Motor Actions. Brain Res. Cogn. Brain Res., 3, 131–141.

Rogers, C. (1959) a Theory of Therapy, Personality, and Interpersonal Relationships as Developed In the Client-Centered Framework. In: Koch, J.S. (Ed.) Psychology: A Study Of a Science: Vol. 3. Formulations of the Person in the Social Context. Mcgraw-Hill, New York, Pp. 184–256.

Rorschach, H. (1922). Psychodiagnostik [Psychodiagnostics]. Bern, Switzerland: Bircher.

Rosenthal, R., Hiller, J. B., Bornstein, R. F., Berry, D. T. R., & Brunell-Neuleib, S. (2001). Meta-analytic methods, the Rorschach, and the MMPI. Psychological Assessment, 13, 449-451.

Rugancı, R. N., & Gençöz, T. (2010). Psychometric Properties of a Turkish Version of the Difficulties in Emotion Regulation Scale. Journal of Clinical Psychology, 66(4), 442-455.

Schipper, M., & Petermann, F. (2013). Relating Empathy And Emotion Regulation: Do Deficits In Empathy Trigger Emotion Dysregulation? Social Neuroscience, 8(1), 101-107.

Schulz, M. S., & Waldinger, R. J. (2005). The value of pooling "naive" expertise: Comment. American Psychologist, 60, 656-657.

Shamay-Tsoory, S. G. (2011). The neural bases for empathy. The Neuroscientist, 17(1), 18-24.

Sijtsma, K. (2009). On The Use, The Misuse, And The Very Limited Usefulness Of Cronbach's Alpha. Psychometrika, 74(1), 107.

Silver, H., Goodman, C., Knoll, G., Isakov, V., & Modai, I. (2005). Schizophrenia Patients With A History Of Severe Violence Differ From Non-Violent Schizophrenia Patients in Perception of Emotions but not Cognitive Function. Journal of Clinical Psychiatry, 66(3), 300–308.

Skitka, L. J., & Sargis, E. G. (2006). The Internet as Psychological Laboratory. Annu. Rev. Psychol., 57, 529-555.

Sonnby-Borgstrom M., Jonsson P., Svensson O. (2003). Emotional Empathy as Related To Mimicry Reactions at Different Levels of Information Processing. Journal of Nonverbal Behavior. Vol.27 (1), 3-23.

Stein, M. L. (1973). An Empirical Validation of the Relation between Rorschach White-Space and Oppositionality. Perceptual and Motor Skills, 37, 375–381

Stellwagen, K. K., & Kerig, P. K. (2013). Dark triad personality traits and theory of mind among school-age children. Personality and Individual Differences, 54(1), 123-127.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal Lobe Contributions to Theory of Mind. Journal of Cognitive Neuroscience, 10(5), 640-656.

Stotland, E. (1969) Exploratory Investigations of Empathy. In: Berkowitz, L. (Ed.), Advances In Experimental Social Psychology. Academic Press, New York, Pp. 271–314.

Stotland, E., Sherman, S., & Shaver, K. (1971). Empathy and Birth Order: Some Experimental Explorations. Lincoln: University Of Nebraska Press.

Stricker, G., & Gold, J. R. (1999). The Rorschach: Toward a nomothetically based, idiographically applicable configurational model. Psychological Assessment, 11, 240-250.

Swart, M., Kortekaas, R., & Aleman, A. (2009). Dealing with Feelings: Characterization of Trait Alexithymia on Emotion Regulation Strategies and Cognitive-Emotional Processing. Plos One, 4(6), E5751.

Tibon Czopp, S., & Zeligman, R. (2016). The Rorschach Comprehensive System (CS) psychometric validity of individual variables. Journal of Personality Assessment, 98, 335-342.

Tomkins S. (1991). Affect, Imagery And Consciousness. New York: Springer Publishing Company.

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. The Quarterly Review of Biology, 46(1), 35-57.

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The "Reading the Mind in the Eyes" Test: Systematic Review

of Psychometric Properties and a Validation Study in Italy. Cognitive Neuropsychiatry, 18(4), 326-354.

Viglione, D. J. (1999). A review of recent research addressing the utility of the Rorschach. Psychological Assessment, 11, 251-265.

Viglione, D. J., & Hilsenroth, M. J. (2001). The Rorschach: Facts, fictions, and future. Psychological Assessment, 13, 452-471

Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., Mckie, S., ... & Elliott, R. (2006). Neuronal Correlates of Theory of Mind and Empathy: A Functional Magnetic Resonance Imaging Study in a Nonverbal Task. Neuroimage, 29(1), 90-98. 10.1016/J.Neuroimage.2005.07.022.

Voracek, M., & Dressler, S. G. (2006). Lack of Correlation between Digit Ratio (2D: 4D) And Baron-Cohen's "Reading the Mind in the Eyes" Test, Empathy, Systemising, And Autism-Spectrum Quotients in a General Population Sample. Personality and Individual Differences, 41(8), 1481-1491.

Weiner, I. B. (1996). Some observations on the validity of the Rorschach Inkblot Method. Psychological Assessment, 8, 206-213.

Weiner, I. B. (1999). What the Rorschach can do for you: Incremental validity in clinical applications. Assessment, 6, 327-339.

Weiner, I. B. (2000). Using the Rorschach properly in practice and research. Journal of Clinical Psychology, 56, 435-438.

Weiner, I. B. (2001). Advancing the science of psychological assessment: The Rorschach Inkblot Method as exemplar. Psychological Assessment, 13, 423-432.

Weisbuch, M., Ambady, N., Slepian, M. L., & Jimerson, D. C. (2011). Emotion Contagion Moderates The Relationship Between Emotionally-Negative Families And Abnormal Eating Behavior. International Journal of Eating Disorders, 44(8), 716-720.

Weiss, E. M., Kohler, C. G., Nolan, K. A., Czobor, P., Volavka, J., Platt, M. M., and others, & Gur, R. C. (2006). The Relationship Between History Of Violent And Criminal Behavior And Recognition Of Facial Expression Of Emotions In Men With Schizophrenia And Schizoaffective Disorder. Aggressive Behavior, 32(3), 187-194.

Westen, D., & Weinberger, J. (2004). When clinical description becomes statistical prediction. American Psychologist, 59, 595-613.

Westen, D., & Weinberger, J. (2005). Clinical judgment in science: Reply. American Psychologist, 60 659-661.

Widiger, T. A. (2001). The best and the worst of us? Clinical Psychology: Science and Practice, 8, 374-377.

Wolfkühler, W., Majorek, K., Tas, C., Küper, C., Saimed, N., Juckel, G., & Brüne, M. (2012). Emotion Recognition in Pictures Of Facial Affect: Is There A Difference Between Forensic and Non-Forensic Patients With Schizophrenia? European Journal of Psychiatry, 26(2).

Wood, J. M. (1996). Thinking Critically About the Comprehensive System for the Rorschach: A Reply to Exner. Psychological Science, 7, 11-17.

Wood, J. M., & Lilienfeld, S. O. (1999). The Rorschach inkblot test: A case of overstatement? Assessment, 6, 341-351.

Wood, J. M., & Nezworski, M. T. (2005). Science as a history of corrected mistakes: Comment. American Psychologist, 60, 657-658.

Wood, J. M., Garb, H. N., Lilienfeld, S. O., & Nezworski, M. T. (2002). Clinical assessment*. Annual Review of Psychology, 53, 519-543.

Wood, J. M., Garb, H. N., Nezworski, M. T., Lilienfeld, S. O., & Duke, M. C. (2015). A second look at the validity of widely used Rorschach indices: Comment on Mihura, Meyer, Dumitrascu, and Bombel (2013). Psychological Bulletin, 141, 236-249.

Wood, J. M., Lilienfeld, S. O., Garb, H. N., & Nezworski, M. T. (2000a). The Rorschach test in clinical diagnosis: A critical review, with a backward look at Garfield (1947). Journal of Clinical Psychology, 56, 395-430.

Wood, J. M., Lilienfeld, S. O., Garb, H. N., & Nezworski, M. T. (2000b). Limitations of the Rorschach as a diagnostic tool: A reply to Garfield (2000), Lerner (2000), and Weiner (2000). Journal of Clinical Psychology, 56, 441-448.

Wood, J. M., Lilienfeld, S. O., Nezworski, M. T., & Garb, H. N. (2001). Coming to grips with negative evidence for the Comprehensive System for the Rorschach: A comment on Gacono, Loving, and Bodholdt; Ganellen; and Bornstein. Journal of Personality Assessment, 77, 48 - 70.

Wood, J. M., Nezworski, M. T., & Garb, H. N. (2003). What's right with the Rorschach? The Scientific Review of Mental Health Practice, 2(7), 142-146.

Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996a). The Comprehensive System for the Rorschach: A critical examination. Psychological Science, 7 (1), 3-10.

Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996b). Thinking critically about the Comprehensive System for the Rorschach: A reply to Exner. Psychological Science, 7, 14-17.

Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1997). The reliability of the Comprehensive System for the Rorschach: A comment on Meyer (1997). Psychological Assessment, 9, 490-494.

Wood, J. M., Nezworski, M. T., Lilienfeld, S. O., & Garb, H. N. (2003). What's wrong with the Rorschach?: Science confronts the controversial inkblot test. San Francisco, CA, US: Jossey-Bass.

Wood, J. M., Nezworski, M. T., Stejskal, W. J., & Garven, S. (2001). Advancing scientific discourse in the controversy surrounding the Comprehensive System for the Rorschach: A rejoinder to Meyer (2000). Journal of Personality Assessment, 76, 369 - 378.

Wood, J. M., Nezworski, M. T., Stejskal, W. J., Garven, S., & West, S. G. (1999). Methodological issues in evaluating Rorschach validity: A comment on Burns and Viglione (1996), Weiner (1996), and Ganellen (1996). Assessment, 6, 115-129.

Wood, J. M., Nezworski, M. T., Stejskal, W. l. J., & McKinzey, R. K. (2001). Problems of the Comprehensive System for the Rorschach in forensic settings. Journal of Forensic Psychology Practice, 1, 93-108.

Wood, J. M., Teresa, M., Garb, H. N., & Lilienfeld, S. O. (2001a). Problems with the norms of the Comprehensive System for the Rorschach: Methodological and conceptual considerations. Clinical Psychology: Science and Practice, 8, 397-402.

Wood, J. M., Teresa, P. M., Garb, H. N., & Lilienfeld, S. O. (2001b). The misperception of psychopathology: Problems with the norms of the

Comprehensive System for the Rorschach. Clinical Psychology: Science and Practice, 8, 350-373.