

I fondamenti scientifici del test di Rorschach: le caratteristiche psicometriche

The scientific basis of the Rorschach: a review of the psychometric properties

Salvatore Zizolfi

Abstract

The Rorschach test is one of the most widely used instrument for the assessment of personality and psychopathology. But controversies of the past several years (1999-2013), concerning psychometric characteristics of the test, have placed clinical and forensic psychologists and psychiatrists in a conflicted position, where they have questioned whether they can continue to use the Rorschach in practice. Aiming to elucidate the scientific basis of the Rorschach, the paper reviews the Rorschach controversy up to date.

Key words: Rorschach test, Rorschach Comprehensive System, Scuola Romana Rorschach, Rorschach norms, Rorschach reliability, Rorschach validity, Rorschach meta-analysis.

Riassunto

Il test di Rorschach, dal suo apparire a tutt'oggi, è sempre stato ed è tuttora fra gli strumenti psicodiagnostici più usati, nei più diversi campi di applicazione. Tuttavia, soprattutto a cavallo della fine del secolo scorso, è stato oggetto di molte critiche, riguardanti le caratteristiche psicometriche del test, ritenute insufficienti a garantire requisiti minimi di scientificità. Il presente contributo ripercorre le evidenze accumulate negli ultimi quindici anni a riguardo delle caratteristiche psicometriche del test, a fondamento del suo uso responsabile in ambito clinico e forense.

Parole chiave: test di Rorschach, Rorschach Comprehensive System, Scuola Romana Rorschach, Caratteristiche psicometriche del Rorschach

Per corrispondenza: Salvatore ZIZOLFI, Via Auguadri, 22, 22100 Como – Cell. 338/3941548, tel.: 031/241375 • e-mail: zizolfi@iol.it

Salvatore ZIZOLFI, Psichiatra, Psicoanalista - Membro della SRR (Scuola Romana Rorschach)

I fondamenti scientifici del test di Rorschach: le caratteristiche psicometriche

1. La fortuna del test di Rorschach in ambito clinico e forense

Fra cinque anni, il test di Rorschach spegnerà le prime cento candeline dalla pubblicazione di *Psychodiagnostik* (1921, 1942, 1981), il capolavoro di Herman Rorschach (1884-1922). L'opera fu un fallimento, e rovinò il piccolo editore svizzero che si era assunto l'onere della pubblicazione: i diritti del libro e delle tavole furono messi all'asta, e acquistati da Hans Huber, che però si guardò bene dal ripubblicare il testo per ben tredici anni, fino al 1934, quando cedette alle pressioni di Morgenthaler (Maresca, 2010). Da allora, la diffusione del Rorschach non ha conosciuto soste in tutto il mondo, e il test è sempre rimasto in vetta alla classifica dei test di maggior interesse e di più largo impiego, nonostante i tempi e i costi di apprendimento e di applicazione di molto superiori a quelli di qualsiasi altra tecnica di indagine psicologica (Camara, Nathan & Puente, 2000).

Nel 1964, la bibliografia in lingua italiana sul Rorschach comprendeva già 304 voci (Chiari, 1964), mentre la bibliografia internazionale arrivava a oltre 4000 contributi (Lang, 1964).

Nei ventanni fra il 1974 e il 1994, per tutti gli anni singolarmente considerati e sul totale del periodo, i lavori scientifici pubblicati sul test di Rorschach (N = 1969) erano secondi solo a quelli pubblicati sull'MMPI (Minnesota Multiphasic Personality Inventory) (N = 4339); il TAT

(Thematic Apperception Test), al terzo posto, totalizza solo 998 articoli (Butcher & Rouse, 1996).

Se digitiamo la parola 'rorschach', otteniamo 3892, 3590, 1217 e 8892 voci bibliografiche rispettivamente con PubMed, EmBase, PsycNET e PsycINFO. Se come termine di paragone utilizziamo l'acronimo MMPI del Minnesota Multiphasic Personality Inventory, risultano 7722, 5678, 1444 e 10576 voci, che diventano 2142, 4015, 1156 e 17414, se digitiamo 'Wechsler Adult Intelligence Scale' (WAIS).

Stando alle statistiche statunitensi, le uniche disponibili, nei venticinque anni fra il 1947 e il 1971, il test di Rorschach, all'inizio il più usato di tutti i test psicologici, cede il primo podio alla WAIS, rimanendo comunque sempre al secondo posto (Klopfer & Taulbee, 1976; Louttit & Browne, 1947). Nei cinquantanni fra il 1935 e il 1982, è stabilmente al quarto posto fra i 30 test psicologici di più frequente impiego, subito dopo WAIS, MMPI e Bender-Gestalt test; è al primo posto nel 1959 e al secondo nel 1969 (Lubin, Larsen, & Matarazzo, 1984). Nel 2000, è al quarto posto, dopo WAIS-R, MMPI-2, e WISC-R (Camara, Nathan, & Puente, 2000).

Se dalla clinica si passa al più specifico ambito forense, i dati disponibili per gli Stati Uniti negli ultimi ventanni sono del tutto sovrapponibili: il test di Rorschach, subito dopo MMPI e WAIS, è al terzo posto fra i test psicologici di più largo impiego nei più diversi settori della pratica psicologica e psichiatrica forense (Archer, Buffington-Vollum, Vauter Stredny, & Handel, 2006; Archer & Wheeler, 2013; Hinselroth & Strycker, 2004) (cfr. Tab. 1).

Area di Intervento (*)	N°	Fonte	MMPI	WAIS	MCMI	R
Criminal Responsibility	96	Borum & Grisso, 1995	86.0 %	83.0 %	18.0 %	30.0%
Competence to Stand Trial	102	Borum & Grisso, 1995	94.0 %	78.0 %	32.0 %	32.0 %
Personal Injury	100	Lees-Haley, et al., 1996	68.0 %	76.0 %	9.0 %	14.0 %
Personal Injury	80	Boccaccini & Brodsky, 1999	94.0 %	54.0 %	50.0 %	41.0 %
Child Custody	101	Ackerman & Ackerman, 1997	92.0 %	43.0 %	-	48.0 %
Child Custody	100	Hagen & Castagna, 2001	84.0 %	-	-	31.0 %
Child Custody	89	Bow et al., 2006	90.0 %	5.0 %	58.0 %	-
Tutte le aree	284	Wygant & Austin, in press	29.0 %	-	26.0 %	22.0 %

(*) È stata mantenuta la dizione in lingua originale, stante la non perfetta sovrapponibilità dei sistemi giudiziari italiano e degli USA. In accordo con Pacente & Grattagliano, 2007, la "Criminal Responsibility" sta per l'eventuale esistenza di un vizio totale o parziale di mente (artt. 88 e 89 c.p.), e la "Competence to Stand Trial" sta per la capacità di stare in processo, ovvero per le condizioni mentali dell'autore di reato durante le indagini e il dibattimento (art. 70 e segg. c.p.p.).

Tab. 1: Percentuale di professionisti che usano un test in ambito forense: Indagini in USA

Tuttavia, nonostante l'enorme successo nella pratica clinica e forense, il test di Rorschach ha avuto in sorte il destino singolare e paradossale di essere nello stesso tempo osannato dai suoi sostenitori come uno strumento imparreggiabile per l'esplorazione della personalità, e di essere vituperato dai suoi detrattori come esempio per eccellenza di "unscientific psychological assessment" (Hunsley & Bailey, 1999), o addirittura un'espressione di pseudoscienza paragonabile all'oroscopo e alla lettura dei tarocchi (Wood, Nezworski, Lilienfeld, & Garb, 2003). Nel 1965, Jensen si spingerà ad affermare che: "The rate of scientific progress in clinical psychology might well be measured by the speed and thoroughness with which it gets over the Rorschach" (Jensen, 1965). A sua volta, Anastasi (1908-2001), indiscussa regina della testistica psicologica, così scriveva con il suo noto equilibrio: "Projective techniques present a curious discrepancy between research and practice. When evaluated as psychometric instruments, the large majority make a poor showing. Yet their popularity in clinical use continues unabated" (Anastasi, 1982).

Negli Stati Uniti, la polemica sulla 'scientificità' del test, ha sempre avuto come bersaglio le proprietà psicometriche dello strumento, ritenute insufficienti dai suoi detrattori.

In effetti, fino a tutti gli anni cinquanta e sessanta, le scuole Rorschach più affermate erano di fatto poco sensibili al richiamo e alle esigenze nomotetiche della psicomетria, e arroccate in una interpretazione idiografica del test come strumento proiettivo, basata sulla clinica e poco o punto attenta alle evidenze della ricerca empirica (Klopfer & Kelley, 1942). Le critiche hanno avuto perciò vita facile, e colpito nel segno, condizionando un declino nell'uso e nell'insegnamento universitario del test intorno alla fine degli anni sessanta.

A partire dai primi anni settanta, Exner (1928-2006) sviluppa il Comprehensive System (CS), il sistema di somministrazione, siglatura ed interpretazione del test di Rorschach nato con l'ambizione di sistematizzare la metodica Rorschach, integrando le cinque principali scuole Rorschach allora più diffuse negli USA (Beck, Hertz, Klopfer, Piotrowski, Rapaport-Schafer) (Exner, 1969). Il CS codifica e standardizza le modalità di somministrazione, siglatura e interpretazione, fornisce norme adeguate di riferimento e rifonda il test su basi empiriche e su un continuo lavoro di ricerca e di aggiornamento (Exner, 1969, 1974, 1978, 1986, 1991, 1993, 1997; Exner & Erdberg, 2005; Exner, Porcelli, & Appoggetti, 2001; Lis, Zennaro, Salcuni, Parolin, & Mazineschi, 2007). "Un sistema vivo e mobile che di anno in anno si fregiava di nuovi indicatori e ne cestinava di obsoleti" (Zennaro, 2015), sulla scorta dei risultati della ricerca, e che di fatto, almeno negli USA, ma anche in molti altri paesi, ha acquisito nel tempo una supremazia incontrastata.

Nel 1988, Anastasi saluta così la novità nella sesta edizione del suo classico *Psychological Testing*: "The availability of this system, together with the research completed thus far, has injected new life into the Rorschach as a potential psychometric instrument" (Anastasi, 1988).

Dieci anni dopo, nel 1998, il successo è tale che l'American Psychological Association conferisce ad Exner l'Award for Distinguished Professional Contributions to Knowledge con una motivazione che non si fa problema di riecheggiare un linguaggio da Nuovo Testamento: "Exner has almost single-handedly rescued the Rorschach

and brought it back to life. The result is the resurrection of perhaps the single most powerful psychometric instrument ever envisioned" (Board of Professional Affairs, 1998).

Ma le critiche non hanno risparmiato neppure il Comprehensive System (CS), che si troverà per così dire fra due fuochi, sotto il tiro incrociato dei nostalgici dell'impostazione clinica ideografica da un lato, e dei puristi della psicomетria dall'altro.

Nel 1999, l'attacco al CS raggiungerà il suo acme, e Howard N. Garb, professore associato di psichiatria a Pittsburg, da sempre fra i più acuti avversari del test, insieme ai colleghi James Wood e Teresa Nezworski, professori di psicologia alla Texas University, arriverà a chiedere una moratoria nell'uso clinico e forense del test (Garb, 1999; Wood, Nezworski, Lilienfeld, & Garb, 2003).

Con l'effetto di stimolare i sostenitori del test, e in primo luogo il direttivo della Society for Personality Assessment, il Rorschach Research Council (fondato da Exner nel 1997 e attivo fino al giorno della sua scomparsa) e il gruppo di ricerca della Toledo University a Toledo, OH, intorno alle figure di Gregory Meyer, Donald Viglione e Joni Mihura, ad approfondire gli aspetti teorici del problema, e a moltiplicare ricerca di buona qualità e studi di meta-analisi sullo strumento, proprio allo scopo di verificare la solidità delle sue basi psicometriche.

Il 2005 è l'anno che sancisce la riscossa dei sostenitori delle qualità psicometriche del Rorschach.

In quell'anno, il Board of Trustees della Society for Personality Assessment pubblica un 'Official Statement' su 'The status of the Rorschach in Clinical and Forensic Practice', che sostanzialmente afferma che il test di Rorschach possiede caratteristiche di attendibilità e di validità simili a quelle di altri strumenti di valutazione della personalità largamente accreditati, e che pertanto il suo uso responsabile nel campo della valutazione della personalità va ritenuto appropriato e giustificato.

Il 20 febbraio 2006, a 78 anni, Exner muore. I suoi eredi legali, estranei al mondo e alla mentalità scientifica, bloccano lo sviluppo del Comprehensive System (CS).

Il gruppo di studiosi che intorno alla figura di Exner aveva dato vita negli anni al Rorschach Research Council, non si perde d'animo, e oltre a continuare ad effettuare studi di meta-analisi sul CS, metterà a punto un nuovo sistema Rorschach, erede e perfezionamento del CS, che verrà presentato nel 2011, alla Annual Convention dell'APA (American Psychological Association) (Meyer & Viglione, 2011), con l'acronimo R-PAS: il Rorschach Performance Assessment System (Meyer, Viglione, Mihura, Erard, & Erdberg, 2013; traduzione italiana, 2015).

Si giunge infine allo studio seminale di meta-analisi di Mihura, Meyer, Dumitrascu, & Bombel (2013) sulla validità di 65 variabili del CS: un articolo di 38 pagine su *Psychological Bulletin*, forse la più antica e prestigiosa rivista di psicologia degli USA (con Impact Factor di 14.39 per il 2014), che, nei suoi 107 anni di storia, a partire dal 1904, ha pubblicato solo 48 lavori sul Rorschach, nessuno dopo il 1988.

Poche settimane fa, sulla stessa rivista, la dichiarazione di tregua, se non di cessazione definitiva delle ostilità, da parte degli avversari di sempre (l'articolo è stato però scritto subito dopo la lettura di quello di Mihura, Meyer, Dumitrascu & Bombel, 2013, e accettato per la pubblicazione già

il 2 dicembre 2013): “Nearly 15 years ago, one of the authors of this Comment published a recommendation that a moratorium be placed upon use of the Rorschach in clinical and forensic settings because of the test’s weak scientific foundation (Garb, 1999). He and the other authors of this Comment agree that, in light of the compelling evidence laid out by Mihura et al. (2013), the time has come to withdraw this recommendation so far as it applies to the Cognitive Quartet of Rorschach scores” (Wood, Garb, Nezworski, Lilienfeld, & Duke, 2015).

2. La storia delle critiche alle caratteristiche psicometriche del test di Rorschach

La storia della ‘Rorschach controversy’ coincide con la storia delle critiche al Comprehensive System (CS) di Exner, dal 1974 al 2011 la versione più diffusa e più accreditata del test di Rorschach dal punto di vista delle caratteristiche psicometriche. Nel 1999 e nel 2000, tre riviste peer-reviewed (Psychological Assessment, Assessment, Journal of Clinical Psychology) (rispettivamente, con Impact Factor per il 2014 di: 2.84, 3.29 e 2.11) hanno pubblicato una sezione speciale dedicata proprio al dibattito fra critici e sostenitori delle proprietà psicometriche del CS.

Le critiche al Comprehensive System (CS) hanno coinvolto: l’accuratezza, l’adeguatezza e la generalizzabilità interculturale delle norme di riferimento; l’attendibilità inter-rater della siglatura; l’attendibilità test-retest degli indici del CS; la validità delle variabili del CS; l’accessibilità della ricerca sul CS (Garb, Wood, Lilienfeld, & Nezworski, 2005; Lilienfeld, Wood, & Garb, 2000).

Il CS era stato apprezzato proprio per la disponibilità di dati normativi di riferimento, basati su un campione di 700 soggetti adulti normali (Exner, 1991, 1993). Diversamente, come si ricorderà, Klopfer era apertamente contrario all’uso di norme di riferimento per il Rorschach (Klopfer & Kelley, 1942) e si rifiutò sempre di includere valori normativi nelle diverse edizioni del suo manuale, mentre Beck (1944), Rapaport (Rapaport & Schafer, 1945) e Piotrowski (1957) si erano serviti di norme, ma sulla base di campioni troppo piccoli e non rappresentativi (cfr le critiche in proposito da parte di Hertz, 1959). Successivamente, però, lo stesso Exner era stato costretto ad una rettifica, chiarendo che in realtà il campione di riferimento era costituito da solo 479 soggetti, e che per un errore 221 protocolli erano stati conteggiati due volte, portando così il totale dei soggetti a 700 (Exner, 2001). Cercò subito di rimediare, aggiungendo altri 121 protocolli, per un numero totale di 600 (Exner, 2001). Ma ai suoi critici non sfuggì che i 600 protocolli del campione di riferimento erano stati selezionati secondo criteri di convenienza, e non casualmente, come sarebbe stato più rigoroso procedere, e che i test erano stati raccolti e siglati nel corso degli anni settanta e ottanta, ovvero 15-25 anni prima, secondo i criteri di somministrazione e di siglatura allora vigenti, e non erano stati risiglati con i criteri riformulati nel 2001 (Garb, Wood, Lilienfeld, & Nezworski, 2005). In pratica, i soggetti esaminati negli anni duemila finivano per essere confrontati con un campione di riferimento non casuale, raccolto 15-25 anni prima, secondo criteri di som-

ministrazione e di siglatura non ritenuti più validi. Non ultimo argomento a sfavore, era l’esiguità numerica del campione di riferimento, soprattutto se paragonato a quelli disponibili per la WAIS (intorno a 2400 soggetti) e l’MMPI-2 (Wood & Lilienfeld, 1999). Alcuni studi, inoltre, avevano nel frattempo evidenziato la scarsa generalizzabilità interculturale e la conseguente inadeguatezza delle norme del CS dell’epoca quando applicate a minoranze etniche – neri, ispanici e nativi americani –: soggetti normali finivano per riportare valori patologici in molti indici del CS (Garb, Wood, Nezworski, Grove, & Stejskal, 2001; Lilienfeld, Wood & Garb, 2000; Wood & Lilienfeld, 1999).

Quel che è più sconcertante, alcune ricerche, a fine secolo scorso, avevano messo in evidenza che le norme del CS allora disponibili non erano sufficientemente rappresentative neppure della popolazione USA normale. In altre parole, numerosi campioni di soggetti normali, e soprattutto di bambini, risultavano anormali, ovvero riportavano, per diverse e importanti variabili del CS, punteggi ricompresi nell’intervallo considerato patologico con riferimento ai valori normativi (Hamel, Shaffer, & Erdberg, 2000; Shaffer, Erdberg, & Haroian, 1999) (cfr la rassegna di 32 studi, a cura di Wood, Nezworski, Garb, & Lilienfeld, 2001). Poco convincenti, risultavano gli argomenti invocati a spiegazione di queste evidenze. In particolare, non reggevano ad un’analisi serrata le ipotesi che: 1) il campione di riferimento per il CS fosse costituito da soggetti con funzionamento psicologico superiore alla media; 2) i campioni di soggetti considerati normali negli studi effettuati fossero in realtà contaminati dalla presenza di pazienti psichiatrici; 3) le discrepanze rispetto ai dati normativi fossero riconducibili a differenze culturali e di provenienza etnica; 4) nel corso degli anni intercorsi fra la raccolta dei dati normativi e gli studi effettuati su campioni di soggetti normali, fosse aumentata la prevalenza di tratti psicopatologici nella popolazione cosiddetta normale; 5) le discrepanze rispetto ai valori normativi fossero dovute alla inadeguata somministrazione del test da parte di esaminatori poco esperti (Garb, Wood, Lilienfeld, & Nezworski, 2005).

A sua volta, l’attendibilità inter-rater della maggior parte degli indici del CS è sempre risultata particolarmente elevata, con una mediana di 0.96, un minimo di 0.78 e un massimo di 0.99 (Exner, 1993). I critici del CS fanno però notare che questi valori si basano su una misura di attendibilità, la percentuale di accordo, che tende a sovrastimare il grado di attendibilità inter-rater (McDowell & Acklin, 1996; Wood, Nezworski, & Stejskal, 1996a, 1996b, 1997). Se si utilizzano coefficienti che tengono conto della quota di accordo dovuta al caso, come il kappa di Cohen, o come il coefficiente di correlazione intraclassa (ICC), si otterrebbero valori soddisfacenti (pari o superiori a 0.85) solo nel 50% delle variabili considerate, con una mediana di 0.80, un massimo di 1,0 e un minimo di 0.20 (Acklin, McDowell, Verschell, & Chan, 2000; Gronnerod, 1999; Nakata, 1999; Garb, Wood, Nezworski, Grove, & Stejskal, 2001).

Per quanto riguarda poi l’attendibilità test-retest delle diverse variabili del CS, fino al 2000 era nota – o almeno pubblicata – solo per 40 su 125 indici del CS, con valori compresi fra 0.30 e 0.90, e una mediana di 0.80. Non era stata pubblicata, per molti indici importanti (Garb, Wood, Nezworski, Grove, & Stejskal, 2001; Lilienfeld, Wood, & Garb, 2000; Wood & Lilienfeld, 1999).

Si rinvia ai lavori originali per quanto riguarda le critiche alla validità degli indici del CS e del Rorschach in generale (Garb, Wood, Lilienfeld, & Nezworski, 2005; Lilienfeld, Wood, & Garb, 2000). In estrema sintesi, anche i più severi critici del Rorschach riconoscono che “Advocates and critics of the test agree that at least a small subset of Rorschach scores is valid when used for particular tasks. Specifically, even psychologist who are critical of the test generally agree that some scores from various Rorschach systems can be helpful for detecting thought disorder, diagnosing mental disorders characterized by thought disorder, measuring dependency, and predicting treatment outcome” (Garb, Wood, Lilienfeld, & Nezworski, 2005). Sul versante opposto, sostenitori e critici del test concordavano circa il fatto che “...many variables given fairly substantial interpretive emphasis have received little or no attention... (i.e.: non sono state studiate e validate adeguatamente)... These include the Coping Deficit Index, Obsessive Style Index, Hypervigilance Index, active-to-passive movement ratio, D-score, food content, anatomy and X-ray content, Intellectualization Index, and Isolation Index” (Meyer & Archer, 2001). Critici e sostenitori ugualmente concordano sulla necessità che gli studi di validità si basino su ricerche ben condotte metodologicamente, anche se non pubblicate, e su meta-analisi globali, ma anche su ‘focused meta-analyses’, centrate su specifiche variabili Rorschach.

Infine, non ha giovato al CS la ridotta accessibilità di tutto il materiale concernente i Rorschach Workshop Studies. Come si sa, gli studi non pubblicati costituiscono il fondamento empirico principale del CS. A questo riguardo, Wood e Lilienfeld (1999) si sono presi la briga di calcolare che Exner cita 157 suoi lavori nelle diverse edizioni del suo manuale fino al 1993: 95 di queste citazioni (il 63%) riguardano studi non pubblicati effettuati nell’ambito dei Rorschach Workshops. C’è di più: nel 1993 e nel 1994, Wood e tre Suoi Colleghi hanno indipendentemente richiesto ai Rorschach Workshops di poter avere accesso ai dati e/o ai manoscritti non pubblicati riguardanti 23 studi recenti, allo scopo di effettuare una rivisitazione critica di tutto il materiale. Non sono riusciti a ottenere nient’altro che cortesi lettere di risposta in cui si declinava l’invito con le motivazioni più diverse (Garb, Wood, Nezworski, Grove, & Stejskal, 2001; Wood & Lilienfeld, 1999).

In definitiva, nel loro articolo del 2001, i più competenti critici del test di Rorschach e del CS, (Garb, Wood, Nezworski, Grove, & Stejskal, 2001), sottolineavano che la richiesta di moratoria avanzata da Garb (1999) si riferiva all’uso clinico e forense del test, non certo all’attività di ricerca. E soprattutto era stata formulata “...not because all Rorschach scores are invalid, but because we do not know which Rorschach studies are valid and which ones are invalid” (Garb, Wood, Nezworski, Grove, & Stejskal, 2001, pag. 443). Auspicavano: il fiorire di ricerche metodologicamente solide, che evitassero i quattro bias più comuni in questo campo (il mancato uso di appropriati gruppi di controllo; la contaminazione del criterio - che si verifica quando la stessa fonte di informazione viene utilizzata per effettuare assessment ratings e criterion ratings -; l’indebita selezione dei risultati, con esclusione per lo più degli studi con evidenze contrarie rispetto a quelle desiderate; il mancato controllo del livello di alfa quando si effettuano più test statistici); la conduzione di apposite indagini di conva-

lida (replication studies); la piena accessibilità ai manoscritti non pubblicati e ai relativi data files; e soprattutto il ricorso a meta-analisi ben condotte, con due esplicite raccomandazioni: inserire anche i risultati derivati da studi non pubblicati (con particolare riguardo alle Doctoral Dissertations), includere tutti i tipi di possibili predittori di una particolare variabile in esame (Garb, Wood, Nezworski, Grove, & Stejskal, 2001).

3. Gli studi recenti a sostegno delle qualità psicometriche del test di Rorschach: le norme di riferimento e l’attendibilità inter-rater

A partire dall’estate del 1997, vari ricercatori di diversi paesi del mondo, tutti accomunati dall’uso del Comprehensive System (CS), hanno dato vita ad un ambizioso progetto internazionale di raccolta e condivisione di protocolli raccolti in campioni normativi di soggetti non clinici. A questo grande sforzo collettivo, hanno partecipato, fra l’altro, anche studiosi che avevano espresso critiche sulla validità dei precedenti campioni normativi per il CS (fra gli altri, gli stessi Erdberg e Shaffer, in prima linea).

La finalità principale era quella di fornire un insieme di norme valide a livello locale o nazionale, in accordo con il suggerimento dalla psicomетria classica, secondo il quale norme locali sono sempre preferibili a norme nazionali o internazionali, perché meglio possono tener conto di specifiche variabili culturali (Anastasi, 1982). Un ulteriore obiettivo prevedeva la messa a punto di un campione normativo internazionale di riferimento, da cui derivare le International Composite Reference Sample Norms, dette anche Composite International Reference Values (CIRVs) (Meyer, Erdberg, & Shaffer, 2007).

I risultati sono stati pubblicati dieci anni dopo, nel 2007, in un numero speciale del Journal of Personality Assessment, che riporta i risultati di 28 gruppi diversi di ricerca in altrettanti campioni di soggetti, provenienti da 16 paesi (USA; Israele; Argentina, Brasile e Perù; Australia; Giappone; Belgio di lingua francese, Danimarca, Finlandia, Grecia, Italia, Olanda, Portogallo, Romania, Spagna): sei campioni di bambini, due di adolescenti, uno di anziani e diciannove di adulti per complessivi 5815 soggetti, di cui 4506 adulti, 1257 fra bambini e adolescenti, e 52 anziani. In ordine alfabetico per nazionalità: Argentina, Adulti, N = 506; Argentina, Adulti, N = 90; Australia, Adulti, N = 128; Belgio di lingua francese, Adulti, N = 100; Brasile, Adulti, N = 409; Danimarca, Adulti, N = 141; Danimarca, Bambini di 9 anni, N = 75; Finlandia, Adulti, N = 143; Grecia, Adulti, N = 98; Israele, Adulti, N = 41; Italia, Adulti, N = 249; Italia, bambini di 5-7 anni e di 8-11 anni, N = 75+148; Italia, Preadolescenti ed Adolescenti, N = 116+117; Giappone, Adulti, N = 240; Giappone, Bambini, N = 190; Olanda, Adulti, N = 108; Perù, Adulti, N = 233; Portogallo, Adulti, N = 309; Portogallo, Bambini, N = 357; Romania, Adulti, N = 111; Spagna, Adulti, N = 517; USA, Adulti, N = 450 (nuovo campione normativo raccolto da Exner); USA, Adulti, N = 283; USA, Adulti Anziani, N = 52; USA, Bambini, N = 100; USA, Bambini di origine messicana, N = 42; USA, Adolescenti, N = 37.

Per ragioni di spazio, in questa sede, non è possibile riferire i dettagli di questa impresa, che ha posto particolare

attenzione alla distinzione fra campioni normativi e campioni di soggetti normali, ai criteri di inclusione e di esclusione, all'impiego o meno di altri strumenti psicodiagnostici da associare al Rorschach, alla rappresentatività dei campioni ed alla conseguente generalizzabilità dei risultati, all'importanza della qualità della somministrazione, della siglatura e dell'interpretazione, e di conseguenza del livello di esperienza e preparazione dei somministratori, all'analisi delle differenze fra gli esaminatori, al monitoraggio della qualità della somministrazione del test, alla selezione dei protocolli e degli esaminatori in rapporto all'attendibilità della siglatura, e al monitoraggio della qualità della siglatura dei test (Shaffer, Erdberg, & Meyer, 2007). All'interno di questo colossale sforzo di messa a punto delle procedure più adeguate, è stata effettuata anche un'apposita indagine volta a delucidare l'influenza del livello di training nell'impiego del CS, sulle modalità di somministrazione, inchiesta e siglatura dei protocolli. È stato evidenziato un impatto drammatico di un maggior livello di training sulla siglatura dei protocolli, e di conseguenza l'importanza di un livello standard di training minimo comune a tutti i ricercatori coinvolti (Lis, Parolin, Calvo, Zennaro, & Meyer, 2007).

L'adozione di metodologie sostanzialmente sovrapponibili nella selezione dei casi, nella raccolta dei protocolli, nella siglatura delle risposte, nei controlli adottati, nella individuazione delle 143 variabili da considerare, nonché nella scelta delle statistiche descrittive (media e SD; punteggi T), le stesse per tutti gli studi condotti, ha grandemente facilitato la comparazione fra i risultati ottenuti. Evidenziando una serie di elementi di grande rilievo ed interesse.

Innanzitutto, in tutti i campioni considerati salvo due eccezioni, e quindi su un totale di 997 protocolli, è stata valutata l'attendibilità inter-rater fra i diversi somministratori, calcolando la percentuale di accordo e il coefficiente iota (Janson & Olsson, 2004), individuati come misure di attendibilità più adeguate per le variabili Rorschach rispetto ai coefficienti di correlazione intraclasse ICC e il coefficiente kappa di Cohen (Lis, Parolin, Calvo, Zennaro, & Meyer, 2007; Meyer, 1997). Tenendo conto della differente numerosità dei campioni, la percentuale media di accordo era superiore a 0.90 per tutte le variabili, mentre il valore medio del coefficiente iota era intorno a 0.84: valori più che soddisfacenti (Meyer, Erdberg, & Shaffer, 2007).

I protocolli raccolti in adulti sono stati poi accorpati in un unico grande campione di 4704 soggetti non clinici (4506 provenienti dagli studi pubblicati su questo Supplemento, gli altri raccolti da Sultan et al, 2004 e 2006). Sono stati calcolati, per ciascun campione, e nel campione generale, i punteggi T per ciascuna delle 136 variabili del CS prese in considerazione. Dal campione generale sono state ricavate le International Composite Reference Sample Norms, dette anche Composite International Reference Values (CIRVs). Quando si confrontano i 20 campioni presi in esame con il campione generale, si scopre che il punteggio T medio per tutte le 136 variabili, con poche trascurabili eccezioni relative a singole variabili in singoli campioni, cade all'interno dell'intervallo fra 40 e 60 punti T del campione globale, con metà delle variabili ricomprese nel punteggio T fra 47 e 53. Si tratta cioè di differenze inferiori a 7.5 punti T, e soprattutto, di punteggi non superiori al valore soglia di 65, né inferiori al valore soglia di 35, notoriamente considerati non significativi a livello clinico (Greene,

2000). In altre parole, le differenze fra un campione nazionale e l'altro, e fra ogni campione nazionale e quello generale da cui erano state ricavate le International Composite Reference Sample Norms, risultavano minime e sostanzialmente trascurabili a livello clinico, a riprova della generalizzabilità interculturale delle norme, relativamente indipendenti dall'influenza di lingua, cultura e nazionalità (Meyer, Erdberg, & Shaffer, 2007). Al contrario, le norme così ottenute differivano in misura significativa da quelle del campione normativo di Exner del 2001: queste ultime presentano valori più elevati (in senso positivo), per più variabili, a riprova della loro insufficiente rappresentatività, e del rischio di far apparire patologici soggetti invece normali (in effetti, anche il campione di Exner del 2007 presenta valori più elevati per l'X+%, con T = 62, rispetto a tutti gli altri campioni, ma si mantiene al di sotto del valore soglia di 65 punti T). In approfondimenti successivi, è stato possibile dimostrare che, nell'ambito del campione CIRVs, non emergono differenze statisticamente significative fra tre sottogruppi individuati in base alla qualità nella raccolta e selezione dei dati, e che è possibile osservare differenze maggiori fra campioni raccolti da diverse zone nella stessa nazione piuttosto che fra questi campioni e le CIRVs (Meyer, Shaffer, Erdberg, & Horn, in press).

In definitiva, come negli obiettivi del progetto, le norme CS del 2007 possono essere utilmente impiegate come valide norme di riferimento, con la possibilità di scegliere fra campioni nazionali, quando presenti, ovvero uno o più campioni di culture vicine (ad esempio, i campioni argentino o brasiliano o peruviano per altri paesi dell'america del sud), ovvero il campione internazionale.

Successivamente, Gregory Meyer, anima dello studio normativo internazionale e dello sviluppo dell'R-PAS, ha chiesto agli Autori che avevano collaborato allo studio del 2007, se volessero mettere a disposizione i loro protocolli per ulteriori ricerche. Quasi tutti hanno fornito un esplicito consenso scritto (Exner, morto nel frattempo, aveva lasciato una liberatoria scritta per soli 67 dei Suoi 450 protocolli), così da poter attingere a 15 campioni nazionali. Per ognuno di questi, venivano estratti casualmente solo 100 protocolli, in modo da bilanciare la diversa numerosità (Il numero totale (N = 1396) è inferiore a 1500 di 104 unità, perché il campione greco è di soli 98 (-2) soggetti, quello israeliano di 93 (-7), quello di Exner di 67 (-33) e quello di Viglione di 38 (-62)). I test disponibili sono stati utilizzati per mettere a punto il c.d. "Campione Normativo di Riferimento per Adulti - 1396" o, più brevemente, il "Legacy sample of the CS", che pertanto si qualifica come un ulteriore campione internazionale di riferimento normativo per il metodo CS (Meyer, Viglione, Mihura, Erard, & Erdberg, 2015). Da questo campione è stato quindi estratto, attraverso complesse procedure, il "Campione modellato statisticamente", comprensivo di 640 protocolli. Si riporta di seguito il contributo di ciascuno dei campioni nazionali alla composizione rispettivamente del "Campione Normativo di Riferimento per Adulti - 1396" (N = 1396) e del "Campione Modellato Statisticamente" (N = 640): per ognuno dei campioni nazionali, vengono riportati due numeri, che si riferiscono rispettivamente il primo, al campione N = 1396, il secondo al campione N = 640: Argentina (N = 100, 25); Belgio (N = 100, 55); Brasile (N = 100, 36); Danimarca (N = 100, 54); Finlandia (N = 100, 44); Francia (N = 100, 47); Grecia (N

= 98, 34); Israele (N = 93, 26); Italia (N = 100, 38); Portogallo (N = 100, 43); Romania (N = 100, 47); Spagna (N = 100, 54); USA (N = 205, 137: a partire dai tre gruppi originari di Exner, Shaffer e collaboratori e Viglione).

Il “Campione modellato statisticamente” costituisce al presente il campione di riferimento provvisorio per l’R-PAS, nelle more della raccolta di campioni internazionali e nazionali di soggetti non clinici secondo la procedura di somministrazione dell’R-PAS. Si ricorda a questo riguardo che la formula di consegna e la modalità di somministrazione dell’R-PAS differiscono e non poco da quelle del CS, della Scuola Romana Rorschach e di tutte le altre scuole Rorschach. Allo scopo di limitare infatti la variabilità del numero di risposte, alla base di molte difficoltà nel trattamento statistico dei dati Rorschach, il probando viene invitato a non superare il numero di quattro risposte per tavola (cosiddetta “Somministrazione R-Ottimizzata”). Si rinvia ai manuali originali dell’R-PAS per maggiori dettagli (Meyer, Viglione, Mihura, Erard, & Erdberg, 2011 e 2015). Va detto che gli studi finora condotti con l’R-PAS su campioni di soggetti non clinici, hanno riportato risultati in linea con i valori normativi del “Campione modellato statisticamente”. “È importante, tuttavia, riconoscere che questa soluzione (ovvero, utilizzare delle stime statistiche basate su protocolli raccolti con il CS, invece di utilizzare dati effettivamente raccolti con il metodo R-PAS) non è comunque ideale. Queste stime devono quindi essere considerate come provvisorie, e hanno unicamente l’obiettivo di fornire delle norme quanto più accurate possibili in attesa di disporre di norme più definitive, messe a punto utilizzando le linee guida della somministrazione R-PAS” (Meyer, Viglione, Mihura, Erard, & Erdberg, 2015). Va comunque segnalato che, nell’ambito del “Campione modellato statisticamente”, così come in altri due campioni di pazienti psichiatrici adulti (N = 249) o in età infantile e giovanile (N = 241), non si osservano associazioni statisticamente significative fra le variabili Rorschach, il sesso, l’appartenenza etnica e l’età adulta, mentre, al contrario, l’età giovanile, e il livello culturale influenzano decisamente alcuni indici Rorschach (Meyer, Gironi, Viglione, Reese, & Mihura, 2015).

Dal canto suo, la Scuola Romana Rorschach (SRR), la più antica istituzione Rorschach al mondo, fondata nel 1938 da Carlo Rizzo (1895-1983), è stata sempre particolarmente attenta alla standardizzazione delle modalità di somministrazione, inchiesta e siglatura del test di Rorschach, e alle problematiche di inter-rater reliability, e già nel 1980, si dotava di un apposito *Manuale per la raccolta, localizzazione e siglatura delle interpretazioni Rorschach* (Rizzo, Parisi, & Pes, 1980), comprensivo della siglatura codificata di circa 6000 interpretazioni raccolte da protocolli di soggetti non clinici. Al presente, la SRR può contare su un database di risposte siglate di oltre 23000 interpretazioni su piattaforma informatica aggiornata di continuo e su una Commissione di 21 esperti nazionali e internazionali, consultabile su richiesta via web entro 24 ore, così da garantire la massima omogeneità inter-rater nella siglatura delle risposte. Ad oggi, il campione normativo della Scuola Romana Rorschach (SRR) è costituito da 792 protocolli (a titolo di paragone, il campione normativo italiano di riferimento consta di 1424 soggetti per la WAIS-IV e di 1375 soggetti per l’MMPI-2), raccolti secondo una campionatura casuale, da più testisti con documentata esperienza nelle

procedure SRR, fra il 1987 e il 1993, secondo i criteri codificati nel Manuale SRR del 1980 già ricordato (Rizzo, Parisi, & Pes, 1980), in altrettanti soggetti adulti normali di 18-54 anni, 396 maschi e 396 femmine, suddivisi in quattro classi di età di uguale numerosità (198 soggetti per ogni classe: 18-24, 25-34, 35-44 e 45-54), e in tre gruppi equipollenti per livello di istruzione (264 con scolarità elementare o media inferiore, e altrettanti con scolarità media superiore e con laurea). La siglatura è stata effettuata per tutti i protocolli, su base concorde, da tre dirigenti della SRR (Salvatore Parisi, Patrizia Pes e Francesco Sartoretti). Venivano esclusi i protocolli con più di due rifiuti, e i soggetti con anamnesi positiva per patologie psichiatriche (consultazioni, trattamenti e ricoveri psichiatrici), tossicodipendenza, alcoolismo e gravi malattie organiche, psicoterapia in atto. Sono attualmente in corso di elaborazione i risultati relativi al nuovo campione aggiornato di 800 protocolli, raccolti a distanza di dieci anni, fra il 2003 e il 2013, secondo la stessa metodologia, somministrati e siglati secondo i nuovi criteri SIGLAROR.

4. Gli studi recenti a sostegno delle proprietà psicometriche del test di Rorschach: la validità del test

A Viglione (1999) si deve un’importante rassegna sistematica della letteratura sulla validità del test di Rorschach, in era anteriore all’affermarsi delle tecniche di meta-analisi. In estrema sintesi, va sottolineato che l’Autore non esamina la validità di singole variabili del CS, ma piuttosto quella di costrutti psicologici derivati dal CS, giungendo alle seguenti quattro conclusioni principali: a) lo Schizophrenia Index (altrimenti detto Perceptual Thinking Index) possiede una validità adeguata quando si tratta di individuare la schizofrenia o altri disturbi psicotici negli adulti, ma non quando vengono esaminati bambini; b) il Depression Index non dovrebbe essere utilizzato per diagnosticare la depressione, benchè abbia qualche capacità predittiva limitatamente alla reazione depressiva in seguito ad eventi di vita negativi; c) la Suicide Constellation è un predittore valido di comportamenti autodistruttivi; d) la validità di due indici di stress situazionale (Movimento Inanimato e Chiaroscuro Diffuso) è sorretta dai dati di letteratura.

Nel 2005, il Board of Trustees of the Society for Personality Assessment ha riesaminato la validità del test di Rorschach alla luce di una serie di meta-analisi condotte sulla letteratura più recente e qualificata allora disponibile.

In primo luogo, ci si è preoccupati di situare comparativamente la validità del test di Rorschach nel contesto più generale dei test psicologici e delle varie metodiche di valutazione diagnostica in medicina. Sono stati riportati in dettaglio i risultati di un articolo di Meyer, Finn, Eyde, Kay, Moreland, Dies, Eisman, Kubiszyn, & Reed, 2001, che illustra i risultati di 125 meta-analisi e di 800 ricerche, dimostrando che gli strumenti di valutazione psicologica sono paragonabili come strumenti di misura, con valori di r compresi fra 0.30 e 0.50, ad indagini di largo uso in altre aree della medicina, come per esempio l’elettrocardiogramma, la mammografia, la risonanza magnetica nucleare, le radiografie dei

denti, il PAP test, la tomografia ad emissione di positroni e il dosaggio del colesterolo. Se si fa riferimento ai 144 coefficienti di correlazione elencati in un'apposita tabella, e se si considerano solo i valori più significativi, pari o superiori a 0.35, si può facilmente evidenziare, in ordine crescente di correlazione, un rapporto stretto fra variabili Rorschach e criteri esterni (0.35), fra Pap test e capacità predittiva del cancro della cervice (0.36), fra test di gravidanza fatti in casa ed esattezza della predizione (0.38), fra livelli di PSA e presenza di un cancro alla prostata (0.40), fra radiografia dei denti e diagnosi della carie interdentale (0.43), fra la PRS Prognostic Rating Scale al Rorschach e la risposta alla psicoterapia (0.44), fra livelli di colesterolo nel sangue e variazioni del colesterolo nella dieta (0.50), e così via.

In secondo luogo, sono stati richiamati i risultati dello studio di Meyer & Archer, 2001, che ha affrontato il tema della validità del test di Rorschach comparata a quella del QI (WAIS), dei questionari MMPI ed MMPI-2, e di altri strumenti di valutazione, lungo tre direttrici di indagine:

- a) attraverso un'indagine comparativa della validità globale di Rorschach, MMPI-2 e WAIS sulla scorta di meta-analisi della letteratura riguardanti distintamente i tre strumenti: i valori di r erano ampiamente sovrapponibili per i tre test, variando da 0.27 a 0.36 per il Rorschach, da 0.22 a 0.55 per l'MMPI, e da 0.33 a 0.57 per la WAIS;
- b) mediante una sintesi degli studi di meta-analisi che mettevano a confronto la capacità predittiva rispetto a un criterio di alcuni indici Rorschach a confronto di quella di altri indicatori; in particolare è stato evidenziato che:
 - a. il comportamento dipendente, così come valutato da un'osservatore esterno, correla con la Rorschach Oral Dependence Scale (0.37) più che con la TAT Dependency Scale (0.34), la MMPI Dependency Scale (0.20) e la MCMI Dependency Scale (0.17) (Bornstein, 1999);
 - b. l'esito della psicoterapia, correla con la PRS Rorschach Prognostic Rating Scale (0.45) più che con la Ego Strength Scale all'MMPI (0.02) e con il QI (0.15) (Meyer, 2000; Meyer & Handler, 1997, 2000);
 - c. l'insorgenza di una malattia fisica, in studi retrospettivi, correla con la Rorschach Oral Dependence Scale (0.72) molto più che con le valutazioni di dipendenza ottenute al TAT (0.29), con appositi questionari (0.18) e all'intervista secondo il DSM (0.09) (Bornstein, 1998);
- c) riportando l'effect size degli studi di meta-analisi riguardanti il valore predittivo specifico, rispetto a un determinato criterio, di diverse variabili Rorschach ovvero del QI e di differenti scale dell'MMPI e dell'MMPI-2. Va sottolineato che in queste indagini viene coinvolto un numero molto elevato di soggetti, e che l'effect size varia da 0.14 a 0.46 per il Rorschach, da 0.02 a 0.45 per l'MMPI e da 0.12 a 0.52 per il QI. In sei delle nove meta-analisi riguardanti il test di Rorschach, l'effect size supera il valore soglia di 0.30, rispettivamente in quelle che indagano la validità generale delle ipotesi Rorschach in assenza di fattori di confondimento (0.32), la validità incrementale del Rorschach rispetto al solo QI nel predire la risposta alla psicoterapia (0.36), la validità concorrente della Rorschach Oral Dependence Scale

rispetto al comportamento dipendente (0.37), l'indice SCZI al Rorschach e la diagnosi di disturbo psicotico (0.44), il punteggio PRS e l'esito della psicoterapia (0.45), il punteggio X+% e la capacità di distinguere fra soggetti normali e patologici (0.46).

In terzo luogo, vengono esposte analiticamente le evidenze a favore della validità incrementale del Rorschach, ovvero della capacità esibita da alcuni indici Rorschach di migliorare la validità predittiva e/o concorrente di altre misure di valutazione, in settori specifici di indagine, così come emergono dalla rassegna della letteratura 1977-1997 condotta da Viglione & Hinselroth (2001): si rinvia, per ragioni di spazio, alla stessa rassegna e ai tredici lavori originali.

Per finire, per garantire la massima imparzialità nella valutazione della validità comparativa fra Rorschach ed MMPI/MMPI-2, è stata designata una apposita commissione (un 'blue-ribbon panel': un panel 'bollino blu'), presieduta da Robert Rosenthal, un esperto statistico metodologo, particolarmente versato in studi di meta-analisi, senza esperienze in campo Rorschach ed MMPI e senza nessun interesse professionale e personale nella psicodiagnostica, e composta da due esperti di meta-analisi applicata rispettivamente al Rorschach (Robert F. Bornstein) e all'MMPI/MMPI-2 (David T.R. Berry). La loro meta-analisi comparativa ha esaminato 30 studi sull'MMPI (N = 4920) e 30 studi sul Rorschach (N = 1713), secondo criteri metodologici più moderni e rigorosi di quanto non si fosse verificato in passato (cfr le meta-analisi di Atkinson, 1986 e di Parker, Hanson, & Hunsley, 1988, e le relative critiche mosse da Garb, Florio, & Growe, 1998). I risultati sono confortanti: in quanto a validità, Rorschach ed MMPI/MMPI-2 sono sostanzialmente sovrapponibili e del tutto soddisfacenti, con valori di r stabilmente intorno a 0.30, in linea con le attese per i test di valutazione della personalità (Hiller, Rosenthal, Bornstein, Berry, & Brunell-Neuleib, 1999; Rosenthal, Hiller, Bornstein, Berry, & Brunell-Neuleib, 2001).

Fin qui, il lavoro svolto dal Board of Trustees of the Society for Personality Assessment.

Il 27 marzo 2008, al Meeting Annuale della Society of Personality Assessment, a New Orleans, in Louisiana, e l'anno dopo, il 27 agosto 2009, al IX Congresso dell'Associazione Europea Rorschach, a Praga, Mihura, Meyer, Dumitrascu e Bombel presentano il lavoro che verrà pubblicato online il 27 agosto 2012 e quindi, come s'è già ricordato, nel 2013, su *Psychological Bulletin*.

Si tratta dello sforzo più grande mai condotto a verifica della validità delle singole variabili Rorschach secondo il CS. Fin da subito, gli Autori prendono posizione concordando con le proposte e gli appelli dei principali critici del CS: "Strictly speaking, it is imprecise to ask if the Comprehensive System of the Rorschach is valid... the validity of each variable must be established separately" (Wood, Nezworski, & Stejskal, 1996a). L'obiettivo riconosciuto e condiviso diventa quello di "... testing the scientific support for the individual components of the Rorschach, not the overall validity of the test" (Mihura, Meyer, Dumitrascu, & Bombel, 2013), attraverso meta-analisi condotte secondo le linee-guida più rigorose (American Psychological Association, 2008; Moher, Liberati, Tetzlaff, Altman, & the PRISMA Group, 2009). Sono state prese in considerazione 65 variabili principali del CS, per ognuna delle quali vengono forniti una precisa defi-

nizione, un esempio di risposta e l'interpretazione corrente nel contesto del CS. Le 65 variabili sono state suddivise in otto gruppi, a seconda che afferiscano alla valutazione degli indici di controllo e di stress situazionale ($N = 17$), degli aspetti emotivi ($N = 8$), della percezione interpersonale ($N = 11$), della percezione di se stessi ($N = 7$), della processazione delle informazioni ($N = 6$), della mediazione cognitiva ($N = 3$), dell'ideazione ($N = 7$) e degli indici derivati ($N = 6$). Sono stati valutati tutti gli studi che esaminavano la validità concorrente delle variabili considerate sia rispetto a criteri 'introspectively assessed' (questionari di autovalutazione, interviste strutturate), sia rispetto a criteri 'externally assessed' (diagnosi DSM, osservatori esterni, test di performance cognitiva), individuati attraverso PsycINFO e Medline, usando 'Rorschach' come parola chiave, e limitandosi agli anni dal 1974 (prima edizione del CS) al 2011. Sono stati così raccolti 3074 coefficienti di validità, ridotti a 1156 in seguito ad appropriata selezione, a fronte di 215 campioni indipendenti (152 di adulti), e un totale di 25795 soggetti partecipanti. 770 coefficienti riguardano criteri 'externally assessed', 386 si riferiscono a criteri 'internally assessed'. Per dieci variabili non era possibile recuperare alcuna ricerca; per ognuna delle 55 variabili rimanenti, vengono riportati, in apposita tabella, il numero di soggetti coinvolti negli studi (in media, 2073, con mediana di 1075), il numero di campioni esaminati (in media, 14.2, con mediana di 13), il numero di effect size ottenuti (in media, 21.0, con mediana di 18), il valore della correlazione r_{RE} con relativo livello di significatività statistica p , e tutta una serie di statistiche relative all'effect size. In accordo con le precedenti risultanze in letteratura, e con quanto era lecito attendersi, la validità concorrente media r è più elevata per gli studi ($N = 770$) che utilizzano un criterio 'externally assessed' (53 meta-analisi; $r = 0.27$), rispetto a quelli ($N = 386$) che ricorrono ad un criterio 'internally assessed' (43 meta-analisi, $r = 0.08$), in buon accordo con la natura stessa del test di Rorschach, che attinge a variabili di personalità non direttamente né facilmente accessibili tramite l'osservazione introspettiva. Restrizzando il campo ai soli studi con criterio 'externally assessed', e raccogliendo le indicazioni di Hemphill (2003) a guida della valutazione della grandezza e della significatività degli effect size, è possibile distinguere cinque gruppi di variabili (nella traduzione dall'inglese delle diverse variabili, ci si attiene alla versione italiana di Lis, Zennaro, Salcuni, Parolin, & Mazzeschi, 2007):

- a) 13 (tredici) variabili CS con validità fortemente supportata dai risultati della meta-analisi ($r \geq 33$; $p < 0,001$; $FSN > 50$): movimento umano, esperienza attuale, totale risposte di chiaroscuro, movimento inanimato, risposte a contenuto anatomico e radiografico, risposte sintetizzate, risposte di forma convenzionale, risposte di forma distorta, risposte di forma appropriata, critical special scores, critical special scores severe, Perceptual Thinking Index, Suicide Constellation;
- b) 17 (diciassette) variabili CS con validità ben supportata dai risultati della meta-analisi ($r \geq 21$; $p < 0,05$; $FSN \geq 10$): numero totale risposte, lambda, sommatoria risposte colore, risposte di chiaroscuro diffuso, difference score, colore acromatico, rapporto forma-colore, rapporto affettivo, indice di complessità, movimento cooperativo, risposte umane intere buone, risposte di testura, buone rappresentazioni umane, risposte mor-

bose, risposte riflesso, organizational frequency, popolari;

- c) 10 (dieci) variabili CS con validità abbastanza supportata dai risultati della meta-analisi ($r \geq 21$; $p < 0,05$; $FSN < 10$; oppure: $r = 0.15-0.20$; $p < 0.05$; $FSN \geq 10$): stimolazione sperimentata, risposte personalizzate, rappresentazioni umane povere, risposte vista, risposte vaghe, perseverazioni, forme inusuali, risposte di movimento umano con forma distorta, Depression Index, Coping Deficit Index;
- d) 3 (tre) variabili CS con validità poco supportata dai risultati della meta-analisi ($p < 0,05$; $r < 0.15$ o $FSN < 10$): movimento animale, movimenti aggressivi, risposte cibo;
- e) 10 (dieci) variabili CS con validità non supportata dai risultati della meta-analisi: Adjusted Difference Score, stile di coping estratensivo versus introversivo, risposte di spazio bianco, risposte di colore puro, indice di isolamento, rapporto attivo-passivo, form dimension, indice di egocentricità, indice di efficienza, indice di ipervigilanza.

Indipendentemente dai risultati ottenuti, comunque rilevanti, questo studio, qui impietosamente riassunto per ragioni di spazio, rappresenta una vera pietra miliare della letteratura sul test di Rorschach. Come hanno riconosciuto anche i più determinati critici del CS, subito all'indomani della lettura dell'articolo: "...the most important contribution of the Mihura et al. (2013) meta-analysis is as a guide to future research... The Mihura et al. meta-analysis should be seen as a valuable impetus for further investigation, rather than the final word regarding the scientific status of clinical psychology's most contested measure" (Wood, Garb, Nezworski, Lilienfeld, & Duke, 2015). È quanto si auguravano gli stessi Autori della meta-analisi: "We hope our findings provide evidence-based guidance for clinical work and stimulate new research to answer lingering questions about how best to understand people based on how they perform, as a complement to what they say based on introspection... we do not expect ours to be the final word on the topic. Indeed, we hope that it is not. As Karl Popper (1934/1959) stated: The wrong view of science betrays itself in the craving to be right; for it is not his possession of knowledge, of irrefutable truth, that makes the man of science, but his persistent and recklessly critical quest for truth" (Mihura, Meyer, Dumitrascu, & Bombel, 2013). Che sinceramente ringraziano gli avversari di sempre a chiusura del loro articolo di replica: "We thank the authors for the opportunity to continue this dialogue in the quest to make the Rorschach a more valid and useful clinical instrument" (Mihura, Meyer, Bombel e Dumitrascu, 2015).

5. Scopi e limiti del presente lavoro, note conclusive

In primo luogo, va sottolineato che la verifica empirica delle caratteristiche psicometriche di base del test di Rorschach non esaurisce il tema dei fondamenti scientifici dello strumento, né tanto meno l'argomento, assai più vasto, della sua utilità in ambito clinico e forense, e delle linee guida cui attenersi perché il test possa dare il meglio di sé in ogni campo di applicazione.

Certamente, però, una buona qualità delle caratteristiche psicometriche di base rappresenta la conditio sine qua non per ogni discorso ulteriore, il fondamento nel senso letterale oltre che metaforico, la premessa indispensabile di ogni altro possibile approfondimento e sviluppo.

Senza verificare preliminarmente attendibilità e validità delle singole variabili Rorschach, non è possibile proseguire oltre, e misurarsi con i problemi più complessi della utilità generale del test di Rorschach nell'ambito clinico e forense, variabili da settore a settore di impiego.

Di qui la scelta di cominciare da questa particolare prospettiva, forse un po' ostica e meno appetibile per i non addetti ai lavori, ma assolutamente imprescindibile: tanto è vero che è stata oggetto di un vero e proprio fuoco di sbarramento, come si è visto, da parte dei detrattori del test.

In secondo luogo, va rimarcato che il presente lavoro non è, e non voleva essere, un'indagine scientifica sulle caratteristiche psicometriche del test di Rorschach. Non riporta nuovi dati sperimentali e clinici al riguardo, non discute problemi teorici di psicomelia applicata al test di Rorschach, e non entra nel merito degli argomenti contrapposti invocati da critici e sostenitori del test a proposito di sottili questioni di natura psicometrica (a quest'ultimo riguardo, vedi ad esempio: Dawes, 1999; Hunsley & Bailey, 2001; Meyer, 1997; Meyer & Archer, 2001; Strycker & Gold, 1999; Weiner, 1995, 1996, 2001).

L'obiettivo del presente lavoro è solo quello di fornire una rassegna aggiornata sulle qualità psicometriche del test di Rorschach, a vantaggio di chi usa il test in ambito clinico e forense, perché in ogni circostanza possa giustificarne e legittimarne l'uso critico e responsabile.

A questo scopo, sono stati ripercorsi oltre vent'anni di controversie sulle qualità psicometriche del Rorschach, per quanto possibile in maniera esaustiva e in modo completo ed equilibrato, sulla scorta della letteratura più rilevante in proposito, variamente individuata e raccolta, con la massima cura nel seguire e illustrare da presso lo sviluppo storico e logico della polemica. Troppo spesso, invece, anche in contributi e dibattiti recenti, e di qualche eco scientifica e professionale, si citano uno o poco più articoli isolati delle due fazioni contrapposte, senza curarsi di collocarli all'interno dello sviluppo cronologico della controversia, come se si trattasse di punti fermi, e senza curarsi di verificare se e in

quale misura siano stati superati da contributi e approfondimenti successivi.

Le conclusioni sono presto dette: in pieno accordo con quanto affermato già nel 2005 dal Board of Trustees of the Society for Personality Assessment, è possibile ribadire che il test di Rorschach possiede caratteristiche di attendibilità e di validità simili a quelle di altri strumenti psicodiagnostici largamente accreditati, e che pertanto il suo uso responsabile nel campo della valutazione della personalità va ritenuto appropriato e giustificato.

Ciò è ancor più vero oggi, alla luce dei dati riportati nella meta-analisi fondamentale di Mihura, Meyer, Dumitrascu, & Bombel (2013), e della dichiarazione di tregua, all'indomani di quest'ultimo articolo, a firma dei critici di sempre (Wood, Garb, Nezworski, Lilienfeld, & Duke, 2015).

Come si è visto, molto è stato fatto, nel settore degli studi sulle caratteristiche psicometriche del test di Rorschach, soprattutto negli ultimi dieci-quindici anni, ma moltissimo rimane ancora da fare.

In primo luogo, bisognerà verificare (convalidare/diconfermare) i risultati ottenuti dalle ricerche americane con il CS e l'R-PAS, in soggetti di cultura europea, e nello specifico nella popolazione italiana, eventualmente anche con altri sistemi Rorschach.

Primo fra tutti, quello della Scuola Romana Rorschach (SRR), che possiede i requisiti preliminari imprescindibili di standardizzazione delle metodiche di somministrazione, di siglatura e di interpretazione, e che da sempre mira a contemperare le esigenze psicometriche delle scuole americane, con quelle cliniche e psicodinamiche delle scuole europee.

S'è già detto del campione normativo italiano di soggetti non clinici messo a punto dalla SRR – il più numeroso e rappresentativo per la popolazione italiana, raccolto secondo criteri metodologici rigorosi e del tutto sovrapponibili a quelli del campione normativo del CS del 2007, e per di più costantemente aggiornato ogni dieci-quindici anni.

La tabella 2 sintetizza i punti di contatto e di divergenza fra CS, R-PAS ed SRR, premessa e base di ogni futura indagine di comparazione e di confronto fra i tre diversi metodi che più di tutti gli altri si sono premurati di standardizzare e codificare le procedure di somministrazione, di siglatura e di interpretazione del test di Rorschach, a premessa e garanzia del livello di qualità delle proprietà psicometriche dello strumento.

	CS	R-PAS	SRR
Indicazioni dettagliate e manualizzate per la somministrazione del test	X	X	X
Registrazione del tempo di latenza per le singole tavole			X
Procedure per ottimizzare il numero delle risposte		X	
Regole di base e indicazioni specifiche per la siglatura delle risposte	X	X	X
Indicazioni dettagliate per risolvere i dubbi di siglatura	X	X	X
Tabelle sinottiche per la valutazione della qualità formale	X	X	X
Supporto alla siglatura con software dedicato (SIGLAROR)			X
Supporto alla siglatura con piattaforma informatica su web		X	X
Norme di riferimento da campioni nazionali (Italia)	X	X	X
Norme di riferimento da campioni internazionali	X	X	X
Valorizzazione della specificità di stimolo delle singole tavole			X
Analisi sequenziale dei contenuti dinamicamente orientata			X

Tab. 2: Caratteristiche comparative di CS, R-PAS ed SRR

Occorre poi proseguire sulla strada già tracciata, sottoponendo ad un esame rigoroso tutte le variabili dei vari sistemi Rorschach, da quelle elementari e più semplici, a quelle più complesse, fino agli indici derivati, privilegiando indagini di validità concorrente rispetto a criteri esterni.

Nulla vieta, per concludere, di applicare analoghe metodologie scientifiche, in grado di soddisfare le esigenze e i requisiti richiesti dalla ricerca più avanzata, allo studio e alla verifica della miriade di indicazioni psicodiagnostiche che ci sono state trasmesse in eredità da ormai quasi un secolo di esperienze cliniche e di ricerca con il test. Le ipotesi psicodiagnostiche avanzate dalle più diverse scuole Rorschach, infatti, non possono più sottrarsi al vaglio della ricerca e della verifica empirica, se non vogliono essere considerate obsolete e prive di fondamento scientifico.

Il test di Rorschach ha attraversato quasi indenne un intero secolo di ricerca scientifica e di pratica clinica: un caso eccezionale di longevità in psicologia e in psichiatria. Come tutti gli strumenti diagnostici e gli interventi terapeutici in psicologia, psichiatria e medicina, nati all'inizio del secolo scorso e in massima parte obsoleti e scomparsi, si è affermato in un'epoca in cui né la comunità scientifica, né le autorità regolatorie avevano posto in essere criteri di verifica e di validazione neppure lontanamente paragonabili a quelli attualmente richiesti.

Se i clinici di più diversa formazione ed estrazione sono sempre stati intimamente convinti della validità del test, l'attacco alle caratteristiche psicometriche del reattivo ci ha fatto veramente correre il rischio concreto di (dover) "buttare il bambino insieme all'acqua sporca".

Scampato il pericolo, occorre evitare che il rischio si ripeta, e che altri aspetti dell'uso e dell'interpretazione del test (ad esempio: il differente valore di stimolo delle diverse tavole) divengano il bersaglio di una critica pregiudizialmente malevola.

Mi si consenta di concludere con una nota autobiografica, e un sentito ringraziamento ai Maestri. Antonio Giambelluca, un allievo diretto di Carlo Rizzo, e mio personale Maestro, generosissimo di insegnamenti ma poco propenso alla produzione scientifica, a metà degli anni settanta del secolo scorso, all'inizio del mio apprendistato personale in psicodiagnostica Rorschach, solleva ripetermi che non c'era più nulla da indagare e da pubblicare, sul test di Rorschach, perché chi ci aveva preceduto aveva già scritto tutto quello che c'era da scrivere. Era evidentemente del tutto ignaro dei fermenti che si agitavano oltre oceano, della nascita del Comprehensive System (CS), e delle nuove esigenze metodologiche, in base alle quali lo stesso Exner (1986), avrebbe considerato scientificamente non validi i due terzi della ricerca Rorschach pubblicata prima del 1970.

Al contrario, oggi, ancor più dopo la recentissima pubblicazione della traduzione italiana del Manuale dell'R-PAS, si apre uno scenario completamente diverso, in cui la ricerca sul test di Rorschach va in gran parte riscritta di nuovo, secondo le indicazioni metodologiche più rigorose e aggiornate, con una serie di replication studies a verifica (convalida, disconferma, precisazione) delle evidenze che ci arrivano dagli Stati Uniti, e con una serie di indagini di buona qualità su quegli aspetti tipicamente europei nell'impiego del test, che non sono ancora stati sottoposti al vaglio della ricerca empirica.

"The absence of evidence is not the evidence of ab-

sence". Chi è convinto, come lo scrivente, della bontà di questo aforisma, sa bene che l'unica via, per non correre il rischio di essere smentiti, in una corsia di ospedale o in un'aula di tribunale, è quella di moltiplicare e pubblicare ricerca di buona qualità e rigorosi studi di meta-analisi.

6. Le prospettive in ambito criminologico e psichiatrico forense

La verifica della solidità dei fondamenti psicometrici del test di Rorschach, e il riconoscimento della validità degli stessi fondamenti da parte dei critici e detrattori di sempre, non possono avere altra conseguenza che rilanciare su nuove e più robuste basi l'impiego del test non solo in ambito clinico, ma anche nel settore criminologico e forense.

Non è certo questa la sede più adatta per una rassegna delle opportunità e delle problematiche di utilizzo del test di Rorschach in criminologia e in psichiatria forense: tematiche il cui sviluppo, e la cui illustrazione, anche solo sommaria, meriterebbe senz'altro un distinto contributo, ben più ampio e diversamente articolato.

Preme però sottolineare, alla fine di questa carrellata sul lungo percorso della ricerca psicometrica sul test di Rorschach, i punti fermi che sono stati definitivamente messi a fuoco, e che costituiscono d'ora in avanti una sorta di condizioni preliminari da rispettare rigorosamente per qualsiasi applicazione del test in ambito clinico e ancor più in criminologia e in psichiatria forense.

In primo luogo, da un punto di vista psicometrico, l'impiego del test di Rorschach può essere riconosciuto attendibile e valido solo se il test viene somministrato e valutato secondo i tre sistemi – il CS (Comprehensive System), l'R-PAS (Rorschach-Performance Assessment System) e il metodo della SRR (Scuola Romana Rorschach) – che soli fra tutti possono contare, come già ricordato, su indicazioni dettagliate e manualizzate per la somministrazione la siglatura e l'interpretazione, su adeguati campioni normativi di riferimento nazionali e internazionali, e sul continuo sviluppo della ricerca clinica e psicometrica.

In secondo luogo, va ricordato che, allo stato attuale della ricerca psicometrica, non tutte le variabili Rorschach hanno dimostrato di possedere in ugual misura, e comunque in grado sufficiente e soddisfacente, caratteristiche psicometriche di attendibilità e di validità. Fortunatamente, le variabili che hanno dimostrato l'attendibilità e la validità più elevate, sono proprio quelle più spesso e più strettamente legate a caratteristiche psicologiche, cliniche e di funzionamento di primaria importanza ai fini clinici, criminologici e psichiatrico forensi. Con riferimento allo studio seminale più volte citato di Mihura, Meyer, Dumitrascu, & Bombel, 2013, le variabili Rorschach che hanno dimostrato di possedere la più elevata validità rispetto ad un criterio esterno sono quelle di movimento umano, esperienza attuale, totale risposte di chiaroscuro, movimento inanimato, risposte a contenuto anatomico e radiografico, risposte sintetizzate, risposte di forma convenzionale, risposte di forma distorta, risposte di forma appropriata, critical special scores, critical special scores severe, Perceptual Thinking Index, Suicide Constellation. Ovvero, rispettivamente, i descrittori Rorschach che misurano le capacità mentali di immagina-

zione, programmazione ed empatia; le risorse cognitive ed emotive; l'ansia e la sensazione di mancanza di aiuto, spesso come conseguenza di condizioni moderate o gravi di stress; la tensione psichica, spesso in risposta a stress moderati o gravi; il grado di preoccupazione per la salute fisica e per il funzionamento corporeo; la capacità di sintesi e di formazione dei concetti; la capacità di percepire la realtà in modo usuale, adeguato e comunemente condiviso; le distorsioni percettive; l'adeguatezza del funzionamento percettivo; le anomalie del pensiero; le anomalie gravi del pensiero; le distorsioni del pensiero e della percezione; il rischio di comportamenti suicidari. In altre parole, i descrittori Rorschach più spesso e maggiormente compromessi nelle gravi patologie psichiatriche, e più direttamente espressivi delle capacità di critica e di giudizio, e infine della capacità di intendere e di volere ai fini psichiatrico forensi.

In terzo luogo, bisogna prendere atto che le singole variabili Rorschach si riferiscono a descrittori di personalità e di funzionamento ovvero a specifici costrutti psicodiagnostici che non corrispondono quasi mai direttamente alle categorie e alle dimensioni della nosografia psicologica e psichiatrica classica. In altri termini, non è più possibile pensare di perseguire l'obiettivo di una corrispondenza precisa fra diagnosi Rorschach e diagnosi psichiatrica, e a maggior ragione fra diagnosi Rorschach e diagnosi criminologica e psichiatrica forense. Come si esprimeva senza mezzi termini lo stesso Exner, lo scopo del Rorschach non è quello di porre una diagnosi psichiatrica, per il semplice motivo che una diagnosi psichiatrica a tipo DSM non è nelle possibilità dello strumento. Il massimo che si può chiedere al test di Rorschach, è quello di valutare la compatibilità maggiore o minore di una particolare costellazione psicodiagnostica Rorschach con uno specifico quadro psicopatologico psichiatrico. Può sembrare poco, soprattutto in ambito forense, dove vige la regola della certezza della prova e della massima rilevanza possibile degli indizi. Ma è invece tantissimo, se si tiene presente che nella pratica clinica e forense è estremamente frequente il riscontro di un grado molto elevato di compatibilità fra peculiari costellazioni psicodiagnostiche Rorschach, e sindromi psichiatriche specifiche, per molti differenti quadri psichiatrici. Si tratta di evidenze accumulate in una pratica clinica pressochè centenaria, che devono ancora passare il vaglio severo della metodologia scientifica moderna più raffinata, ma che è facile prevedere che supereranno i controlli più rigorosi in un prossimo futuro. Verrà così confermato, per questa via, l'enorme vantaggio di poter confortare o falsificare una diagnosi clinica psichiatrica sulla base di elementi ed evidenze psicodiagnostiche Rorschach che per natura, provenienza e raccolta, sono completamente eterogenei e indipendenti rispetto alle categorie e alle dimensioni della semeiotica psichiatrica classica. Elementi ed evidenze psicodiagnostiche che, nonostante la deprecabile diffusione sul web delle informazioni più disparate sul test di Rorschach, è per di più e tuttora molto più difficile simulare e dissimulare di quanto non sia possibile per molti sintomi clinici.

Ancora una volta, è facile prevedere che il test di Rorschach festeggerà il suo primo centenario, fra cinque anni, lasciandosi alle spalle un glorioso passato, e guardando con fiducia ad un radioso avvenire davanti a sé.

Bibliografia

- Ackerman, M.J., & Ackerman, M.C. (1997). Custody evaluations practices: a survey of experienced professionals (revisited). *Professional Psychology: Research and Practice*, 28, 137-145.
- 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 Psychological Association. (2008). Meta-Analysis Reporting Standards.
- Anastasi, A. (1982). *Psychological testing*. (5th Edition). NY: Mac-Millan.
- Anastasi, A. (1983). *I test psicologici. Nuova edizione riveduta e ampliata*. Milano, Franco Angeli.
- Anastasi, A. (1988). *Psychological testing*. (6th Edition). NY: Mac-Millan.
- Archer, R.P., Buffington-Vollum, J.K., Vauter Stredny, R. & Handel, R.V. (2006). A Survey of Psychological Test Use Patterns Among Forensic Psychologists. *Journal of Personality Assessment*, 87(1), 84-94.
- Archer, R.P., & Weeler, E.M.A. (Eds) (2013). *Forensic uses of clinical assessment instruments*. (2nd ed.). New York, NY: Routledge.
- Atkinson, L. (1986). The comparative validities of the Rorschach and MMPI: a meta-analysis. *Canadian Psychology*, 27, 238-247.
- Beck, S.J. (1944). Rorschach's test. Basic processes. New York: NY: Grune & Stratton.
- Board of Professional Affairs. (1998). Awards for distinguished professional contributions: John Exner. *American Psychologist*, 53, 391-392.
- Boccaccini, M.T., & Brodsky, S.L. (1999). Diagnostic test use by forensic psychologists in emotional injury cases. *Professional Psychology: Research and Practice*, 30, 253-259.
- Bornstein, R.F. (1998). Interpersonal dependency and physical illness: a meta-analytic review of retrospective and prospective studies. *Journal of Research in Personality*, 32, 480-497.
- Bornstein, R.F. (1999). Criterion validity of objective and projective dependency test: a meta-analytic assessment of behavioral prediction. *Psychological Assessment*, 11, 48-57.
- Borum, M., & Grisso, T. (1995). Psychological test use in criminal forensic evaluations. *Professional Psychology: Research and Practice*, 26, 465-473.
- Bow, J.N., Gould, J.W., Flens, J.R., & Greenhut, D. (2006). Testing in child custody evaluations: selections, usage, and Daubert admissibility: a survey of psychologists. *Journal of Forensic Psychology Practice*, 6, 17-38.
- Butcher, J.N., & Rouse, S.V. (1996). Personality: individual differences and clinical assessment. *Annual Review of Psychology*, 47, 87-111.
- Camara, W.J., Nathan, J.S., & Puente, A.E. (2000). Psychological test usage: implications in professional psychology. *Professional Psychology: Research and Practice*, 31, 141-154.
- Chiari, S. (1964). Bibliografia italiana sul reattivo di Rorschach. *Supplemento al Bollettino di Psicologia Applicata*, 61-62, 1-81.
- Dawes, R.M. (1999). Two methods for studying the incremental validity of a Rorschach variable. *Psychological Assessment*, 11, 297-302.
- Exner, J.E. (1969). *The Rorschach systems*. New York, NY: Grune & Stratton.
- Exner, J.E. (1974). *The Rorschach: a Comprehensive System. Volume 1*. New York, NY: Wiley.
- Exner, J.E. (1978). *The Rorschach: a Comprehensive System. Volume 2: Current Research and Advanced Interpretation*. New York, NY: Wiley.
- Exner, J.E. (1986). *The Rorschach: a Comprehensive System. Volume 1: Basic foundations* (2nd edition). New York, NY: Wiley.
- Exner, J.E. (1991). *The Rorschach: a Comprehensive System. Volume*

- 2: *Interpretation* (2nd Edition). New York, NY: Wiley.
- Exner, J.E. (1993). *The Rorschach: a Comprehensive System. Volume 1: Basic foundations* (3rd edition). New York, NY: Wiley.
- Exner, J.E. (1997). *The Rorschach: a Comprehensive System. Volume 1: Basic foundations and Principles of Interpretation* (4th edition). New York, NY: Wiley.
- Exner, J.E. (2001). *Rorschach workbook for the Comprehensive System*. (5th Edition). Asheville, NC: Rorschach Workshops.
- Exner, J.E., & Erdberg, P. (2005). *The Rorschach: a Comprehensive System. Volume 2: Advanced Interpretation* (3rd Edition). New York, NY: Wiley.
- Exner, J.E., Porceppi, P., Appoggetti, P. (2001). *Il test di Rorschach secondo il sistema di Exner. Trento: Erikson*.
- Garb, H.N. (1999). Call for a moratorium on the use of the Rorschach Inkblot in clinical and forensic settings. *Assessment*, 6, 313-317.
- 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., Wood, J.M., Lilienfeld, S.O. & Nezworski, M.T. (2005). Roots of 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.
- Greene, R. L. (2000). *The MMPI-2. An interpretive Manual*. (2nd Edition). Boston, MA: Allyn & Bacon.
- Gronnerod, C. (1999). Rorschach interrater agreement estimates: an empirical evaluation. *Scandinavian Journal of Psychology*, 40, 115-120.
- Hagen, M.A. & Castagna, N. (2001). The real numbers: psychological testing in custody evaluations. *Professional Psychology: Research and Practice*, 32, 269-271.
- Hamel, M., Shaffer, T.W., & Erdberg, P. (2000). A study of nonpatient preadolescent Rorschach protocols. *Journal of Personality Assessment*, 75, 280-294.
- Hemphill, J.F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, 58, 78-79.
- Hertz, M.R. (1959). The use and misuse of the Rorschach method. I. Variations in the Rorschach procedure. *Journal of Projective Techniques*, 23, 33-48.
- 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.
- Hinselroth, M.J., & Strycker, G. (2004). A Consideration of Challenges to Psychological Assessment Instruments Used in Forensic Settings: Rorschach as Exemplar. *Journal of Personality Assessment*, 83(2), 141-152.
- 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.
- Janson, H., & Olsson, U. (2004). A measure of agreement for interval or nominal multivariate observations by different sets of judges. *Educational and Psychological Measurement*, 64, 62-70.
- Jensen, A.R. (1965). A review of the Rorschach. In O.K. Buros (Ed.), *Sixth mental measurement yearbook* (pp. 501-509). Highland Park, NH: Gryphon.
- Klopper, W.G. & Kelley, D.M. (1942). *The Rorschach Technique*. Yonkers on Hudson, NY: World Book.
- Klopper, W.G. & Taulbee, E.S. (1976). Projective tests. *Annual Review of Psychology*, 27, 543-567.
- Lang, A. (1964). *Rorschach-Bibliographie. Bibliographie Rorschach. Rorschach Bibliography. 1921-1964*. Bern: Hans Huber.
- Lees-Haley, P.R., Smith, H.H., Williams, C.W. & Dunn, J.T. (1996). Forensic neuropsychological test usage: an empirical survey. *Archives of Clinical Neuropsychology*, 11, 45-51.
- 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.
- Lis, A., Zennaro, A., Salcuni, S., Parolin, L. & Mazzeschi, C. (2007). *Il Rorschach secondo il Sistema Comprensivo di Exner. Manuale per l'utilizzo dello strumento*. Milano: Raffaello Cortina Editore.
- Lis, A., Parolin, L., Calvo, V., Zennaro, A., & Meyer, G. (2007). The impact of administration and inquiry on Rorschach Comprehensive System protocols in a national reference sample. *Journal of Personality Assessment*, 89 (S1), 193-200.
- Louitt, C.M., & Brown, C.G. (1947). The use of psychometric instruments in psychological clinics. *Journal of Consulting Psychology*, 11, 49-54.
- Lubin, B., Larsen, R.M. & Matarazzo, J.D. (1984). Patterns of psychological test usage in the United States: 1935-1982. *American Psychologist*, 39, 451-454.
- Maresca, L. (2010). *Storia del Rorschach. I contributi scientifici che hanno dato forma ai diversi metodi*. Roma: LAS.
- McDowell, C., & Acklin, M.W. (1996). Standardizing procedures for calculating Rorschach interater reliability: conceptual and empirical foundations. *Journal of Personality Assessment*, 66, 308-320.
- Meyer, G.J. (1997). Assessing reliability: critical corrections for a critical examination of the Rorschach Comprehensive System. *Psychological Assessment*, 9, 480-489.
- Meyer, G.J. (2000). The incremental validity of the Rorschach Prognostic Rating Scale over the MMPI Ego Strength Scale and IQ. *Journal of Personality Assessment*, 74, 356-370.
- 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., Erdberg, P., & Shaffer, T.W. (2007). Toward International Normative Reference Data for the Comprehensive System. *Journal of Personality Assessment*, 89, S1, S201-S216.
- Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.D., Eisman, E.J., Kubiszyn, T.W., & Reed, G.M. (2001). Psychological testing and psychological assessment: a review of evidences and issues. *American Psychologist*, 56, 128-165.
- Meyer, G.J., Giromini, L., Viglione, D.J., Reese, J.B., & Mihura, J.L. (2015). The association of gender, ethnicity, age and education with Rorschach scores. *Assessment*, 22, 46-64.
- Meyer, G.J., & Handler, L. (1997). The ability of the Rorschach to predict subsequent outcome: a meta-analysis of the Rorschach Prognostic Rating Scale. *Journal of Personality Assessment*, 69, 1-38.
- Meyer, G.J., & Handler, L. (2000). The ability of the Rorschach to predict subsequent outcome: a meta-analysis of the Rorschach Prognostic Rating Scale: Correction. *Journal of Personality Assessment*, 74, 504-506.
- Meyer, G.J., Shaffer, T.W., Erdberg, P., & Horn, S.L. (in press). Addressing issues in the development and use of the Composite International Reference Value sas Rorschach norms for Adults. *Journal of Personality Assessment*, in press.
- Meyer, G.J., & Viglione, D.J. (2011). New developments in Rorschach-based behavioral assessment. Annual Convention of the American Psychological Association, August, 4-7, Washington, DC.
- Meyer, G.J., Viglione, D.J., Mihura, J.L., Erard, R.E. & Erdberg, P. (2011). *Rorschach Performance Assessment System. TM. Administration, Coding, Interpretation and Technical Manual*. Toledo, OH: Rorschach Performance Assessment System, LLC.
- Meyer, G.J., Viglione, D.J., Mihura, J.L., Erard, R.E. & Erdberg, P. (2015). *Rorschach Performance Assessment System. TM. Somministrazione, siglatura, interpretazione e manuale tecnico*. Edizione Italiana a cura di Lis, A. & Zennaro, A. Milano: Raffaello Cortina.
- Mihura, J.L., Meyer, G.J., Bombel, G. & Dumitrascu, N. (2015).

- Standards, accuracy and question 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.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS medicine*, 6 (7), 1-9.
- Nakata, L.M. (1999). Interrater reliability of the Comprehensive System for the Rorschach: clinical and non-clinical protocols (Doctoral Dissertation, Pacific Graduate School of Psychology). Dissertation abstracts International.
- Pacente, F., & Grattagliano, F. (2007). Il reattivo di Rorschach secondo l'Exner Comprehensive System in ambito medico-legale e psichiatrico-forense. *Rassegna Italiana di Criminologia*, 1, 3, 170-198.
- Parker, K.C.H., Hanson, R.K., & Hunsley, J. (1988). MMPI, Rorschach and WAIS: a meta-analytic comparison of reliability, stability and validity. *Psychological Bulletin*, 103, 367-373.
- Piotrowski, Z.A. (1957). *Perceptanalysis*. New York, NY: Macmillan.
- Rapaport, D. & Schafer, R. (1945). The Rorschach test: a clinical evaluation. *Bulletin of the Menninger Clinic*, 9, 73-77.
- Rizzo, C., Parisi, S., & Pes, P. (1980). *Manuale per la raccolta, localizzazione e siglatura delle interpretazioni Rorschach*. Roma: Kappa.
- 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.
- Rorschach, H. (1921). *Psychodiagnostik. Methodik und Ergebnisse eines Wahrnehmungsdiagnostischen Experiments (Deutenlassen von Zufallsformen)*. Bern: Hans Huber.
- Rorschach, H. (1942). *Psychodiagnostics*. Bern: Hans Huber.
- Rorschach, H. (1981). *Psychodiagnostica. Metodologia e risultati di un esperimento diagnostico basato sulla percezione (Interpretazione di forme casuali)*. Roma: Kappa.
- Shaffer, T.W., Erdberg, P., & Haroian, J. (1999). Current nonpatient data for the Rorschach, WAIS-R and MMPI-2. *Journal of Personality Assessment*, 73, 305-316.
- Shaffer, T.W., Erdberg, P., & Meyer, G.Y. (2007). Introduction to the JPA Special Supplement on International Reference Samples for the Rorschach Comprehensive System. *Journal of Personality Assessment*, 89, S1, S2-S6.
- Society for Personality Assessment. (2005). The Status of the Rorschach in Clinical and Forensic Practice: An Official Statement by the Board of Trustees of the Society for Personality Assessment. *Journal of Personality Assessment*, 85, 219-237.
- Stricker, G., & Gold, J.R. (1999). The Rorschach. Toward a non-mothetically based, idiographically applicable configurational model. *Psychological Assessment*, 11, 240-250.
- Sultan, S., Andronikof, A., Fouques, D., Lemmel, G., Mormont, C., Reveillere, C., & Saïas, T. (2004). Vers des normes francophones pour le Rorschach en Système intégré: premiers résultats sur un échantillon de 146 adultes. *Psychologie Française*, 49, 7-24.
- Sultan, S., Andronikof, A., Reveillere, C., & Lemmel, G. (2006). A Rorschach stability study in a nonpatient adult sample. *Journal of Personality Assessment*, 87, 330-348.
- Viglione, D. (1999). A review of recent research addressing the utility of the Rorschach. *Psychological Assessment*, 11, 251-265.
- Viglione, D.J., & Hinselroth, M.J. (2001). The Rorschach: facts, fictions, and future. *Psychological Assessment*, 13, 452-471.
- Weiner, I.B. (1995). Methodological considerations in Rorschach research. *Psychological Assessment*, 7, 330-337.
- Weiner, I.B. (1996). Some observations on the validity of the Rorschach Inkblot Method. *Psychological Assessment*, 8, 206-213.
- Weiner, I.B. (2001). Advancing the science of psychological assessment: the Rorschach Inkblot Method as exemplar. *Psychological Assessment*, 13, 423-432.
- 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. (1999). The Rorschach Inkblot Test: a case of overstatement? *Assessment*, 6, 341-349.
- Wood, J.M., Nezworski, M.T., Garb, H.N., & Lilienfeld, S.O. (2001). 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., Nezworski, M.T., Lilienfeld, S.O., & Garb, H.N. (2003). *What's wrong with the Rorschach?* San Francisco, CA: John Wiley & sons.
- Wood, J.M., Nezworski, M.T., & Stejskal, W.J. (1996a). The Comprehensive System for the Rorschach: A critical examination. *Psychological Science*, 7, 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.
- Wygant, D.B. & Austin, K.D. (in preparazione). *A survey of psychological test use patterns: adult personality and psychopathological measures*.
- Zennaro, A. (2015). Presentazione dell'edizione italiana. In: Meyer, G.J., Viglione, D.J., Mihura, J.L., Erard, R.E. & Erdberg, P. (2015), .i.c.