

Testes Psicológicos mais Usados em Portugal: Resultados de um Inquérito a Psicólogos Portugueses

Most Commonly Used Psychological Tests in Portugal: Findings from a Survey of Portuguese Psychologists

Mário R. Simões¹, Maria João Seabra-Santos², Pedro Armelim Almiro³ e Leandro S. Almeida⁴

Resumo

A avaliação psicológica constituiu uma atividade relevante e exclusiva da Psicologia, justificando a análise da evolução da sua prática. Este artigo identifica os testes e outros instrumentos de avaliação psicológica mais utilizados em Portugal. Participaram 1383 psicólogos que responderam a um inquérito *online*. Os resultados permitem salientar que os testes de avaliação da inteligência (escalas de inteligência de Wechsler e, particularmente, a WISC) e de raciocínio não-verbal (Matrizes de Raven) estão entre os mais utilizados. Verifica-se um incremento no recurso a testes neuropsicológicos, testes de rastreio cognitivo, escalas de avaliação do comportamento e testes de rendimento ou realização escolar, e uma diminuição na utilização de testes projetivos. Assinala-se, ainda, uma maior utilização de inventários de sintomas psicopatológicos do que inventários mais clássicos de personalidade. A discussão dos resultados remete para a necessidade de mais investigação relativa aos instrumentos utilizados, explicitando os dados referentes à precisão, validade e normalização dos seus resultados, favorecendo a qualidade dos testes em uso e assegurando assim uma utilização mais criteriosa dos instrumentos disponíveis pelos psicólogos portugueses nas suas práticas de avaliação psicológica.

Palavras-chave: avaliação psicológica, testes psicológicos, psicométrica, prática psicológica, psicólogos portugueses

Abstract

Psychological assessment has been a relevant and exclusive activity of Psychology, justifying an analysis of the evolution of its practice. This article identifies the tests and other psychological assessment instruments most commonly used in Portugal. A total of 1383 psychologists took part and responded to an online survey. The results show that intelligence assessment tests (Wechsler Intelligence Scales, particularly the WISC) and non-verbal reasoning tests (Raven's Matrices) are among the most widely used. There has been an increase in the use of neuropsychological tests, cognitive screening tests, behavioral assessment scales and school performance or achievement tests, while there has been a decrease in the use of projective tests. On the other hand, there was a greater use of psychopathological symptom inventories than more classic personality inventories. The discussion of the results points to the need for more research into the instruments used, focusing on data relating to the reliability, validity and standardization of their results, promoting the quality of the tests in use and ensuring a more rigorous use of the instruments available to Portuguese psychologists in their psychological assessment practices.

Keywords: psychological assessment, psychological tests, psychometrics, psychological practice, portuguese psychologists

Agradecimentos: Os autores estão gratos à Ordem dos Psicólogos Portugueses (OPP) e aos psicólogos portugueses que responderam ao convite para participar neste estudo.

¹Doutor em Psicologia (Avaliação Psicológica). Professor Catedrático na Universidade de Coimbra. Diretor do Laboratório de Avaliação Psicológica e Psicométrica (PsyAssessmentLab). Universidade de Coimbra, Faculdade de Psicologia e de Ciências da Educação; CINEICC – Centro de Investigação em Neuropsicologia e Intervenção Cognitivo-Comportamental; PsyAssessmentLab. Rua do Colégio Novo, 3000-115 Coimbra, Portugal. Tel.: +351 239851450. E-mail: simoesmr@fpce.uc.pt

²Doutora em Psicologia (Avaliação Psicológica). Professora Associada com Agregação na Universidade de Coimbra. Universidade de Coimbra, Faculdade de Psicologia e de Ciências da Educação; CINEICC – Centro de Investigação em Neuropsicologia e Intervenção Cognitivo-Comportamental; PsyAssessmentLab. Rua do Colégio Novo, 3000-115 Coimbra, Portugal. Tel.: +351 239851450. E-mail: seabramj@fpce.uc.pt

³Doutor em Psicologia (Avaliação Psicológica). Professor Auxiliar na Universidade Autónoma de Lisboa. Universidade Autónoma de Lisboa, Departamento de Psicologia; CIP – Centro de Investigação em Psicologia; CINEICC – Centro de Investigação em Neuropsicologia e Intervenção Cognitivo-Comportamental; PsyAssessmentLab. Rua de Santa Marta, 47, 1150-293 Lisboa, Portugal. Tel.: +351 213177600. E-mail: pedroarmelimalmiro@gmail.com

⁴Doutor em Psicologia (Psicologia da Educação). Professor Catedrático na Universidade do Minho. Universidade do Minho, Escola de Psicologia; CIPsi – Centro de Investigação em Psicologia. Campus de Gualtar, 4710-057 Braga, Portugal. Tel.: +351 253604255. E-mail: leandro@psi.uminho.pt

Introdução

Os dados obtidos a partir da administração de testes psicológicos fazem geralmente parte do processo, mais vasto e complexo, de avaliação psicológica que inclui e integra igualmente informações provenientes de outras metodologias, técnicas e fontes de dados, em particular os dados provenientes das entrevistas e da observação direta (Fernandez-Ballesteros et al., 2001; Meyer et al., 2001). Neste contexto, os testes e outros instrumentos estandardizados assumem um papel insubstituível neste processo, contribuindo para a qualidade da resposta aos vários objetivos da Avaliação Psicológica: (i) descrição do funcionamento atual da pessoa, sua definição e caracterização psicológica (e.g., nas suas dimensões ou modos de resposta adaptativa, cognitiva, emocional, personalidade); (ii) estimativa do nível de funcionamento pré-mórbido do sujeito; (iii) objetivação quantificada de áreas de funcionamento problemático, incluindo o padrão e severidade dos défices cognitivos ou sintomas psicopatológicos, traços e outras características da personalidade; (iv) identificação de domínios de funcionamento positivo; (v) rastreio; (vi) processo de diagnóstico, incluindo diagnóstico diferencial, situações de comorbilidade; (vii) orientação vocacional; (viii) seleção profissional; (ix) planificação, monitorização e avaliação da intervenção; (x) avaliação de programas; (xi) resposta a questões de natureza forense; (xii) investigação (Haynes et al., 2019; Wright et al., 2017).

Por outro lado, o desenvolvimento, investigação e uso de instrumentos são elementos distintivos do ensino e da prática profissional da avaliação psicológica baseada na evidência científica. Neste plano, é tida em conta a investigação relevante sobre testes psicológicos relativa a diferentes dados de natureza psicométrica que permitem analisar a qualidade dos instrumentos: adaptação, construção e funcionamento dos itens, diversos índices de fiabilidade/precisão, validade de constructo e eficácia preditiva (Bornstein, 2017; Haynes et al., 2019), normas representativas (Hunsley & Meyer, 2003), e uma avaliação crítica, incluindo quer a identificação de limitações relativas à validade e generalização dos resultados dos instrumentos,

quer a indicação de sugestões acerca do seu aperfeiçoamento (Haynes et al., 2019). Importa igualmente admitir que a “evidência psicométrica é condicional, não é uma característica unidimensional ou estável de uma medida” e que a “validade não é uma característica invariante” do teste, dependendo de variáveis como as “populações/amostras, os objetivos da avaliação, aplicabilidade e sensibilidade aos indivíduos examinados e às diferenças individuais, reconhecendo-se ainda que os contextos específicos estudados podem não ser amplamente generalizáveis” (Haynes et al., 2019, p. 7).

As investigações acerca dos “testes e outros instrumentos de avaliação mais utilizados” constituem uma prática frequente, bem documentada em mais de uma centena de publicações. A primeira investigação conhecida deste tipo foi publicada em 1947 (Louttit & Browne, 1947). De uma forma esquemática, e não exaustiva, estas investigações consideram o uso de testes: (i) simultaneamente em diferentes países (e.g., 17 países europeus: Evers et al., 2012; Estados- Unidos e Canadá: Wright et al., 2017); (ii) em países específicos como os Estados- Unidos (e.g., Lubin et al., 1984), Canadá (Hill & Demetriooff, 2019), Austrália (e.g., Martin et al., 2001), Espanha (e.g., Hernández et al., 2020; Muñoz et al., 2020), Alemanha (Wertz et al., 2022), Portugal (e.g., Diniz et al., 2007), Brasil (e.g., Reppold et al., 2020), China (e.g., Ryan et al., 1994), ou Filipinas (Tarroja et al., 2020); (iii) em diferentes áreas de exercício profissional da Psicologia, como é o caso dos contextos escolar (e.g., Benson et al., 2019), clínico (e.g., Wade & Baker, 1977), clínico com base em diferentes orientações/modelos (e.g., Watkins et al., 1995), aconselhamento (e.g., Watkins et al., 1988), forense (e.g., McLaughlin & Kan, 2014), neuropsicológico (e.g., Rabin et al., 2016), ou organizacional (Ryan & Sackett, 1987), incluindo serviços de emprego (Turcotte, 1994); (iv) tipos específicos de testes abrangendo diferentes domínios de avaliação, como a personalidade (Smith et al., 2010), ou os estilos de resposta como simulação, exagero de sintomas, esforço insuficiente, através de testes de validade de desempenho e de testes de validade de sintomas (e.g., Martin et al., 2015); (v) grupos etários distintos como crianças (Cook et al., 2017),

adolescentes (Archer & Newsom, 2000), adultos (Harrison et al., 1988), idosos (Simões et al., 2018); (vi) condições neurodegenerativas como a demência (Di Pucchio et al., 2018); (vii) reabilitação vocacional (Donoso et al., 2010); ou, ainda, (viii) custódia de crianças (e.g., Ackerman et al., 2021).

A investigação centrada na identificação dos testes mais usados serve diversos objetivos. Desde logo, tais estudos permitem identificar os testes disponíveis e aqueles que servem melhor as necessidades práticas dos psicólogos, podendo esta informação orientar a formação inicial (graduação) e contínua dos psicólogos. Nessa formação inclui-se a apreciação da fundamentação teórica dos testes e, em particular, as evidências e atualização das suas propriedades psicométricas (precisão, validade, normas). Por outro lado, tomando um país ou conjunto de países, ou uma área ou várias áreas da intervenção psicológica, estes estudos permitem detetar a evolução e os avanços na área dos testes, considerando dados de natureza psicométrica ou a sua atualização com a introdução de novas revisões de instrumentos existentes, assim como a emergência e utilização de instrumentos com novos referenciais teóricos ou resposta a contextos e a necessidades mais recentes da avaliação psicológica. Por último, este tipo de estudos permite uma análise interativa da investigação-formação-prática envolvendo as condições e recursos da avaliação psicológica num determinado país ou domínio do exercício da Psicologia.

A relevância destes estudos é reconhecida no apoio explícito à sua realização por parte de organizações científicas e profissionais de psicólogos. A título exemplificativo é de mencionar a implementação deste tipo de estudos por diferentes divisões e sociedades da *American Psychological Association* (APA) (e.g., *Quantitative and Qualitative Methods*, *Society of Clinical Psychology*, *School Psychology*, *Society of Counseling Psychology*, *Society for Clinical Neuropsychology*, *Psychologists in Independent Practice*); *National Academy of Neuropsychology* (NAN), *International Neuropsychological Society* (INS); *National Association of School Psychologists* (NASP); *Society of Personality Assessment* (SPA); *International Test Commission* (ITC); *European Federation of Psychologists' Associations* (EFPA);

Colegio Oficial de Psicólogos (Espanha) e, no caso da presente investigação, a *Ordem dos Psicólogos Portugueses* (OPP). Todas estas investigações referem o uso de um número elevado de testes psicológicos, evidenciando que esta utilização constitui uma prática comum e, simultaneamente, uma tarefa fundamental dos psicólogos, sinalizando também que os testes psicológicos são incontornáveis no processo de avaliação psicológica (e.g., Groth-Marnat & Wright, 2016) e neuropsicológica (e.g., Sherman et al., 2023).

Neste contexto, a escolha criteriosa dos testes, bem como a sua correta administração, cotação e interpretação dos resultados, desempenha um papel essencial na avaliação psicológica. A investigação empírica aponta vários fatores que influenciam a escolha ou seleção dos testes: experiência obtida com testes específicos no decurso da formação e da prática, possibilidade de resposta a necessidades particulares da avaliação, fiabilidade e validade dos resultados, existência de normas (e de normas atuais) para interpretar os resultados, condições impostas pela instituição de trabalho ou por exigências legais (e.g., defensabilidade dos testes no tribunal), disponibilidade dos testes na instituição empregadora, popularidade/aceitação geral de certos testes entre os profissionais, facilidade e celeridade na administração e cotação, e familiaridade ou comodidade na experiência com o teste (e.g., Lees-Haley et al., 1996; Watkins et al., 1995).

A questão da escolha (e uso) dos testes tem sido por vezes colocada de forma distinta em diferentes domínios de exercício da avaliação psicológica. Comparativamente a áreas como a educação ou clínica, a escolha, a utilização e os resultados nos testes são mais escrutinados, nas suas potencialidades e limites, na área da avaliação (neuro)psicológica forense. Neste âmbito, a investigação de Lally (2003) identificou opiniões de especialistas em psicologia forense sobre aceitabilidade do uso de testes específicos e em função de áreas particulares de avaliação psicológica forense, de acordo com a seguinte tipologia: “Recomendado” ou “Aceitável” (e.g., WAIS-III, MMPI-2, na avaliação do estado mental aquando da ofensa, simulação e capacidade para testemunhar em tribunal e, juntamente com PCL-R, na avaliação do risco de violência), “Equivocoinaceitável” (e.g., Rorschach, na avaliação do

estado mental aquando da ofensa), “Inaceitável” (e.g., Rorschach, TAT, desenhos projetivos, na avaliação do risco de violência, capacidade para testemunhar em tribunal, simulação); “Sem opinião”. Mais recentemente, e com base nas normas de Daubert, Archer e colaboradores (2016) definiram critérios que devem ser considerados na escolha e admissibilidade de testes psicológicos no âmbito de uma “avaliação empiricamente fundamentada”. Esses critérios incluem vários tipos de evidências: (i) standardização adequada e qualificações exigidas ao utilizador do teste; (ii) níveis aceitáveis de fiabilidade e de validade das pontuações no teste, quando este é usado para objetivos forenses específicos (este critério inclui a margem de erro conhecida para o teste, que pode ser expressa pelo erro padrão de medida e/ou através dos valores de sensibilidade e especificidade); (iii) “aceitação pela comunidade científica”, incluindo aqui um conjunto suficiente de investigação examinada por pares e publicada em revistas de referência, bem como o consenso por parte de psicólogos especialistas na área forense; (iv) adaptações internacionais do teste, com o objetivo de identificar possíveis diferenças de natureza cultural nos resultados; (v) dados de natureza comparativa adequados (e.g., dados normativos específicos para ajudar a interpretar os resultados no teste nos diferentes grupos de interesse nos quais o teste é utilizado e que permitam responder às questões de avaliação (e.g., legais) (cf. igualmente Bow et al., 2006).

Com o presente estudo pretende-se: (i) identificar os testes psicológicos mais utilizados em Portugal; (ii) comparar os dados obtidos com estudos congéneres nacionais e internacionais; (iii) analisar em que medida estes testes mais utilizados acompanham linhas orientadoras (e.g., evidência de fiabilidade/precisão e validade, normas representativas) ou as atualizações resultantes de novas revisões e edições de testes considerados como relevantes; (iv) identificar se são mencionados instrumentos recentemente disponíveis em Portugal e que procuram atender a novas necessidades de avaliação. É importante acrescentar que, no âmbito deste projeto de investigação centrado na utilização de testes em Portugal, foram já publicados três trabalhos (Almeida et al., 2021; Seabra-Santos et al., 2019, 2021).

Método

Amostra

Nas investigações recenseadas é muito variável o número efetivo de respondentes (e de questionários completos e utilizáveis para análise). Constituídas por psicólogos profissionais e/ou doutorados em Psicologia, as amostras incluem os seguintes números: 12.606, de 17 países europeus (Evers et al., 2012); 3.126, que responderam à pergunta enviada aos 51.545 membros do Colégio Oficial de Psicólogos (Muñiz & Fernández-Hermida, 2010); 1.946 (Louttit & Browne, 1947); 1.383 (presente investigação); 1.072 (Hernández et al., 2020); 747 neuropsicólogos clínicos doutorados membros da Divisão 40 da *American Psychological Association* (APA), da *National Academy of Neuropsychology* (NAN) ou da *International Neuropsychological Society* (INS) (Rabin et al., 2005); 648 psicólogos escolares membros da NASP (Shapiro & Heick, 2004); 512 membros doutorados da NAN e do INS (Rabin et al., 2016); 488 psicólogos escolares (Benson et al., 2019); 471 (Wade et al., 1978); 449 neuropsicólogos (Guilmette et al., 1990); 440 psicólogos inscritos no Conselho Federal de Psicologia, do Instituto Brasileiro de Avaliação Psicológica e da Associação Nacional de Pesquisa e Pós-Graduação em Psicologia (Reppold et al., 2020); 412 (Watkins et al., 1995). Outros estudos conhecidos, e frequentemente citados, remetem para amostras mais reduzidas: 211, numa investigação realizada junto de psicólogos doutorados americanos e canadianos (Wright et al., 2017); 64 psicólogos de 64 países, um por país (Oakland et al., 2016); 24 neuropsicólogos especialistas na área da simulação (Slick, 2004); 19 (de 41 peritos contactados, de 21 países), incluindo Espanha, Portugal e países da América Latina (Muñiz et al., 1999).

No caso do presente estudo, um total de 1.921 psicólogos portugueses acederam, após convite, ao *site* da Ordem dos Psicólogos Portugueses (OPP) para participarem no estudo. Porém, alguns não iniciaram o preenchimento do inquérito ($n=211$; 11.0%) e outros ficaram apenas pela caracterização inicial do respondente ($n=187$; 9.7%). De entre os 1.523 respondentes que completaram o questionário de forma válida, 140 (9.2%)

afirmaram não utilizar testes e outros instrumentos de avaliação. As análises reportadas neste artigo incidem sobre os restantes 1.383 (90.8%) participantes, que afirmaram utilizar testes e/ou questionários psicológicos, de forma habitual, no exercício da sua atividade profissional, ou seja, responderam positivamente à questão: “*Na sua atividade profissional utiliza habitualmente testes e/ou questionários psicológicos?*”. Destes, 1215 eram do sexo feminino (87.85%) e 168 do sexo masculino (12.15%). Relativamente ao grau de formação dos psicólogos respondentes, 336 (24.30%) concluiu o grau de Mestre pós-Bolonha, 320 (23.14%) completou uma Licenciatura de cinco anos pré-Bolonha, 207 (15.0%) obteve uma especialização pós-licenciatura e 127 (9.2%) eram doutorados. As idades dos participantes oscilavam entre os 24 e os 76 anos, com uma média de 39.77 anos ($DP=9.24$). O número de anos de atividade profissional variava entre 1 e 45, com a moda situada nos 10 anos ($n=112$) ($M=13.50$, $DP=8.35$) e uma maior concentração de valores entre os 10 e os 15 anos de prática. Relativamente às áreas de especialização dos psicólogos participantes que mencionam a utilização de testes e questionários psicológicos, importa referir que uma parte muito substantiva destes trabalha na área da Psicologia Clínica e da Saúde ($n=972$; 70.3%). Seguem-se as áreas da Psicologia Escolar e da Educação ($n=493$; 35.6%), Neuropsicologia ($n=228$, 16.5%), Psicologia do Trabalho ($n=139$; 10.1%), Psicologia Social e Comunitária ($n=124$; 9.0%) e Psicologia Forense ($n=119$; 8.6%). Na leitura destes dados é necessário reconhecer que alguns psicólogos trabalham em mais do que uma área de especialização.

Dos 140 respondentes que referiram “não” utilizar instrumentos de avaliação na sua prática profissional, 51 participantes justificaram essa opção de resposta. Algumas razões apontadas para o não uso de testes e/ou questionários psicológicos aparentam ser compreensíveis e são as mais frequentemente citadas: a sua área de intervenção (ação social, investigação, docência, direção/coordenação de serviços de psicologia) não torna necessária a realização da avaliação psicológica e, em particular, o uso de testes e outros instrumentos; a avaliação psicológica não se adequa e não é aplicável (sem a apresentação de outras razões); a avaliação não está prevista no

planeamento das suas atividades; não reconhecem benefício no recurso à avaliação psicológica nas situações habituais da sua prática profissional; estar desempregado. Outras justificações, ainda que raramente mencionadas, poderão ser consideradas problemáticas: a instituição onde trabalham não adquiriu testes; os testes não são indispensáveis para a prática clínica ou a psicoterapia; falta de instrumentos para as populações específicas de trabalho profissional; escassez de tempo e condições; ou a preferência pelo recurso exclusivo à entrevista.

Instrumento

O questionário aplicado incluía várias secções: (i) Breve caracterização do(a) psicólogo(a) respondente, nomeadamente Idade, Sexo, País, Região, Nível de formação, Campo de atuação atual, Experiência profissional (anos), Área(s) de formação na pós-graduação, Área(s) de prática na Psicologia; (ii) Descrição da utilização de testes psicológicos, nomeadamente Recurso habitual a testes psicológicos; Razões, se não recorre; Objetivo (Diagnóstico, Recrutamento, Orientação ou Aconselhamento, Avaliação de programas); População (idades) e Testes usados na prática profissional; (iii) Atitudes face à formação e utilização de provas psicológicas, à necessidade de regulamentação no seu uso e ao uso exclusivo destes instrumentos por parte de psicólogos; e (iv) Problemas identificados na utilização de testes. Em particular, neste estudo analisam-se os testes mencionados pelos psicólogos participantes na resposta à questão colocada relativa aos 10 testes ou questionários mais utilizados na sua prática profissional.

Procedimentos

É diverso o procedimento empregue no processo de recolha das respostas acerca dos testes mais utilizados. A maior parte das investigações coloca uma pergunta aberta acerca dos testes mais utilizados na prática profissional, frequentemente sem qualquer restrição quanto ao seu número (e.g., Reppold et al., 2020) ou com algum grau de limitação: três instrumentos mais utilizados na avaliação da personalidade e três instrumentos mais utilizados na avaliação da inteligência, para cada uma de sete faixas etárias (Brown & McGuire, 1976), 10 testes (como acontece no presente

estudo). Em outras investigações é previamente enviado aos respondentes uma lista muito variável de testes (Ryan et al., 1994), que pode ser previamente identificada com base em “resumos da literatura” (Butler et al., 1991), a partir da consulta de relatórios de avaliação neuropsicológica forense elaborados por peritos (Lees-Haley et al., 1996), com base numa lista de 145 instrumentos de avaliação comercializados por nove editoras de testes (Noronha et al., 2005) ou em 154 instrumentos identificados a partir de quatro manuais de testes (Peterson et al., 2014).

No presente estudo, o questionário foi aplicado *online*. No início apresentavam-se os objetivos do estudo e a relevância do maior número de respostas para assegurar maior robustez nas análises e conclusões. Para além de não ser requerida a identificação pessoal, assegurou-se o tratamento anónimo das respostas. Como atrás mencionado, os participantes listaram os testes mais usados na sua prática sem que qualquer lista de testes lhes tenha sido previamente facultada.

As análises dos dados foram conduzidas com o programa IBM SPSS Statistics, versão 26.0 para Windows. Considerando os objetivos do presente artigo, as análises circunscreveram-se a estatísticas descritivas (frequências, percentagens) para caracterizar a amostra do ponto de vista demográfico, área de atividade profissional e de especialização, testes usados e respetivas menções, em geral e por grandes áreas ou domínios de avaliação.

Resultados

Para esta análise considerámos os 50 “testes e/ou questionários psicológicos” mais referenciados pelos respondentes. Neste particular, importa referir que nas investigações congêneres é muito variável o número de testes valorizados nas análises: 300 instrumentos sinalizados pelos participantes (Goth et al., 1981), 150 (Benson et al., 2019), 116 (Butler et al., 1991), 113 (Louttit & Browne, 1947), 73 (Lubin et al., 1971), 44 (Lees-Haley et al., 1996), 40 (Ryan et al., 1994), 38 (Watkins et al., 1995), 31 (Sundberg, 1961), 30 (Cashel et al., 2002), 29 (Diniz et al., 2007), 25 (Muñiz & Fernández-Hermida, 2010), 20 (Camara, 2000), 14 (Giromini et al. 2022), 13 (Wright et al., 2017), ou 10 (Oakland et al., 2016). De acrescentar também alguma diversidade nos critérios fixados

para identificar os testes mais utilizados. Por exemplo, no estudo de Louttit e Browne (1947) incluíram-se os testes com quatro ou mais menções, enquanto no estudo de Sundberg (1961) se consideraram os testes usados por 10% ou mais da amostra.

A Tabela 1 apresenta os 50 instrumentos de avaliação mais utilizados pelos psicólogos portugueses na sua prática profissional. Estes instrumentos são de natureza diversificada (testes, escalas, questionários, inventários) e abrangem a maior parte das áreas ou domínios comumente contemplados nos processos de avaliação psicológica (inteligência, funções cognitivas específicas, personalidade, psicopatologia, comportamento, interesses, desenvolvimento, realização/rendimento escolar).

O maior destaque está nas escalas de inteligência de Wechsler, de modo particular a WISC-III (o 1º instrumento da lista), a WAIS-III (4º) e, num registo menos expressivo, a WPPSI-R (20º). Dentro da área da cognição merecem relevo três grupos de testes. Por um lado, os testes de raciocínio, em particular duas das versões dos testes das Matrizes de Raven – SPM (3º) e CPM/CPM-P (6º) – as BPRD/BPR (5º) e ainda as PMA (30º), uma bateria de aptidões que inclui também a avaliação do raciocínio lógico. Ainda no domínio da cognição, especificamente a cognição geral, é indispensável mencionar os testes de rastreio cognitivo: MMSE (12º), MoCA (26º) e ACE-R (43º). Por outro lado, surgem com saliência os testes neuropsicológicos de avaliação de funções cognitivas como a memória, atenção ou funções executivas: FCR (2º), d2 (8º), TP (22º), Bender (39º), WMS-III (39º), TMT (42º) ou uma bateria de 15 testes neuropsicológicos, recentemente editada, a BANC (38º).

Bem presentes nas práticas de avaliação psicológica dos psicólogos respondentes estão, também, quer os inventários de avaliação da psicopatologia – BDI/BDI-II (8º), SCL-90-R (15º), BSI (16º), YSR (17º), CDI/CDI-2 (36º), GDS-30 (45º), PHSCS-2 (47º), quer as escalas de avaliação do comportamento de crianças e adolescentes completadas por pais e professores – TRF (13º), CBCL (14º) e Conners (24º).

Relativamente aos instrumentos de avaliação da personalidade é possível considerar dois grandes grupos: os testes projetivos e os inventários de

Tabela 1. Testes e instrumentos de avaliação mais frequentemente utilizados

Posição	Teste / Instrumento	n	%
1	Escala de Inteligência de Wechsler para Crianças (WISC, WISC-R, WISC- III, WISC-IV, WISC-V)	676	49.10
2	Figura Complexa de Rey (FCR)	296	21.40
3	Matrizes Progressivas de Raven (SPM/PM38)	192	13.88
4	Escala de Inteligência de Wechsler para Adultos (WAIS-R, WAIS-III, WAIS-IV)	190	13.74
5	Bateria de Testes de Raciocínio (Diferencial) (BPRD/BPR)	165	11.93
6	Matrizes Progressivas Coloridas de Raven (CPM/CPM-P)	148	10.70
7	Interesses e Preferências Vocacionais (IPP/IPP-R)	146	10.55
8	<i>Beck Depression Inventory</i> (BDI/BDI-II)	137	9.91
8	Teste de Atenção (d2)	137	9.91
10	Rorschach	122	8.75
11	Teste de Apercepção para Crianças (CAT-A/ CAT-H)	116	8.39
12	<i>Mini-Mental State Examination</i> (MMSE)	112	8.10
13	<i>Teacher Report Form</i> (TRF)	109	7.88
14	<i>Child Behavior Checklist</i> (CBCL)	107	7.74
15	<i>Symptom Checklist-90-Revised</i> (SCL-90-R)	103	7.45
16	<i>Brief Symptom Inventory</i> (BSI)	102	7.38
17	<i>Youth Self-Report</i> (YSR)	98	7.09
18	<i>Family Drawing</i> (Corman)	96	6.94
19	Escalas de Desenvolvimento Mental de Griffiths (0-2 e 2-8 anos)	91	6.58
20	Escala de Inteligência de Wechsler para a Idade Pré-escolar e Primária (WPPSI-R, WPPSI-III, WPPSI-IV)	83	5.93
21	Teste de Apercepção Temática (TAT)	81	5.86
22	Bateria de Aptidões para a Aprendizagem Escolar (BAPAE)	80	5.78
23	Teste de Barragem de Toulouse e Piéron (TP)	73	5.28
24	Escala Revista de Conners (Pais e Professores)	67	4.84
24	<i>Minnesota Multiphasic Personality Inventory</i> (MMPI/MMPI-2)	67	4.84
26	Avaliação Cognitiva de Montreal (MoCA)	63	4.56
27	<i>California Occupational Preference System</i> (COPS/COPS-R)	62	4.48
28	<i>Roberts Apperception Test</i> (RATC/RATC-2)	58	4.19
29	Inventário de Personalidade NEO- Revisto (NEO-PI-R)	54	3.90
30	Aptidões Mentais Primárias (PMA)	49	3.54
31	<i>Semistructured Clinical Interview for Children and Adolescents</i> (SCICA)	46	3.33
32	<i>Millon Clinical Multiaxial Inventory</i> (MCMI/MCMI-II/MCMI-III)	45	3.35
33	Era Uma Vez ... Prova Projectiva para Crianças	44	3.18
34	Bar-Ilan: Teste de Figuras para Crianças	42	3.03
35	Questionário Fatorial de Personalidade (16PF/16PF-5)	41	2.96
36	Inventário de Depressão para Crianças (CDI/CDI-2)	37	2.68
36	<i>Mini-Mult</i> (versão reduzida do MMPI)	37	2.68
38	Bateria de Avaliação Neuropsicológica de Coimbra (BANC)	33	2.39
39	Teste Guestáltico Visomotor de Bender (Bender)	32	2.31
39	Escala de Memória de Wechsler (WMS-R, WMS-III)	32	2.31
39	O Rei – Teste de Fluência e Precisão da Leitura	32	2.31
42	<i>Trail Making Test</i> (TMT)	31	2.24
43	<i>Addenbrooke's Cognitive Examination</i> (ACE-R)	30	2.17
43	<i>Draw-a-Person</i> (D-A-P, Goodenough)	30	2.17
45	Provas de Diagnóstico Pré-Escolar (Pré-Escolar)	29	2.10
45	<i>Geriatric Depression Scale</i> (GDS-30)	29	2.10
47	<i>Piers-Harris Self-Concept Scale</i> (PHSCS-2)	28	2.02
47	House-Tree-Person (H-T-P)	28	2.02
49	Escala de Avaliação das Competências de Desenvolvimento Infantil (SGS, SGS-II)	27	1.95
50	Teste Exploratório de Dislexia Específica (PEDE)	27	1.95

Nota. A denominação dos instrumentos é referida de acordo com a designação da edição comercial portuguesa, quando existe e está traduzida para português, ou seguindo a designação original.

personalidade. Os testes projetivos mais mencionados são: Rorschach (10°), CAT-A/ CAT-H (11°), Desenho da Família (18°), TAT (21°), RATC/RATC-2 (28°), Era uma vez (33°). Os inventários de personalidade referidos incluem: MMPI/MMPI-2 (24°), NEO-PI-R (29°), MCMI/MCMI-II/MCMI-III (32°), 16PF/16PF-5 (35°), *Mini-Mult* (versão reduzida do MMPI) (36°) e PHSCS-2 (47°).

Com menor relevo encontramos instrumentos orientados para outras áreas de avaliação: os

inventários de interesses vocacionais, IPP/IPP-R (7°) e o COPS/COPS-R (27°); os testes de realização/rendimento escolar, BAPAE (22°); O Rei (45°); Pré-Escolar (45°) e PEDE (50°) e, finalmente, os testes de desenvolvimento, Griffiths (0-2 e 2-8 anos) (19°) e SGS/SGS-II (49°).

Uma leitura mais agregada e resumida dos dados pode ser consultada na Tabela 2, onde se menciona o número de testes/instrumentos usados e o número de citações por grandes áreas ou

Tabela 2. Número de testes/instrumentos usados e número de citações por grandes áreas / domínios de avaliação

Área / Domínio	Número de testes usados	Número de citações
Inteligência / Cognição Geral	11	1738
Funções (Neuro)cognitivas	7	634
Personalidade / Testes Projetivos	8	587
Inventários de Sintomas Psicopatológicos	6	506
Comportamento	3	283
Personalidade / Inventários de autorrelato	6	272
Interesses	2	208
Realização/Rendimento escolar	4	168
Desenvolvimento	2	118
Total	50	4560

Tabela 3. Comparação entre estudos Portugueses relativos à utilização de testes / questionários psicológicos (presente estudo e Diniz et al., 2007), considerando instrumentos reportados em ambos os estudos

	Presente estudo <i>N</i> = 1383	Presente estudo Classificação	Diniz et al. (2007) <i>N</i> = 265	Diniz et al. (2007) Classificação
Escala de Inteligência de Wechsler para Crianças (WISC, WISC-R, WISC-III, WISC-IV, WISC-V)	676	1	111	1
Figura Complexa de Rey (FCR)	296	2	42	7
Matrizes Progressivas de Raven (SPM/PM38)	192	3	53	3
Escala de Inteligência de Wechsler para Adultos (WAIS, WAIS-R, WAIS-III, WAIS-IV)	190	4	51	4
Bateria de Testes de Raciocínio Diferencial (BPRD/BPR)	165	5	32	8
Matrizes Progressivas Coloridas de Raven (MPCR/MPCR-P)	148	6	11	19
<i>Beck Depression Inventory</i> (BDI/BDI-II)	137	8	13	15
Rorschach	122	10	71	2
Teste de Apercepção para Crianças (CAT-A / CAT-H)	116	11	45	5
<i>Family Drawing</i>	96	18	20	12
Escalas de Desenvolvimento Mental de Griffiths (0-2 e 2-8 anos)	91	19	23	10
Teste de Apercepção Temática (TAT)	81	21	43	6
Teste de Barragem de Toulouse e Piéron (TP)	73	23	21	11
<i>Minnesota Multiphasic Personality Inventory</i> (MMPI /MMPI-2)	67	24	30	9
Inventário de Personalidade NEO- Reviso (NEO-PI-R)	54	29	12	17
<i>Millon Clinical Multiaxial Inventory</i> (MCMI/MCMI-II/MCMI-III)	45	32	17	13
<i>Mini-Mult</i> (versão reduzida do MMPI)	37	36	12	17
Teste Gestáltico Visomotor de Bender (Bender)	32	39	16	14
<i>Draw-a-Person</i> (D-A-P, Goodenough)	30	43	13	15

domínios de avaliação. Totalizando 11 testes e 1738 citações, os instrumentos de avaliação da inteligência (sobretudo as escalas de inteligência de Wechsler) e da cognição em geral (testes de raciocínio não-verbal como as Matrizes de Raven, testes de rastreio cognitivo, como o MMSE) ocupam de forma destacada o 1º lugar relativo às grandes áreas de avaliação. A seguir, os testes neuropsicológicos (7 testes, 634 menções) suplantam os testes projetivos de avaliação da personalidade (8 testes, 587 menções). Com base num número igual de testes (6) identificamos os inventários de sintomas psicopatológicos com um número substancialmente mais elevado de menções (506) do que os inventários de personalidade (272 menções). Os três instrumentos/escalas de avaliação do comportamento apresentam mais menções (283) que os seis inventários de personalidade (272).

Destaque ainda para os dois inventários de interesses, com 208 menções.

Na Tabela 3 é apresentada a comparação entre estudos portugueses sobre utilização de testes / questionários psicológicos (presente estudo e Diniz et al., 2007), considerando instrumentos reportados em ambos os estudos. Importa mencionar que o número de testes indicado no estudo de Diniz e colaboradores (29 testes) é menor do que o número de testes listado no presente estudo (50 testes) e, ainda, que o número de testes em comum nos dois estudos é igual a 19.

Comparativamente ao estudo de 2007, o presente estudo evidencia três tendências relativamente aos testes mais utilizados. Por um lado, uma estabilidade relativa aos testes mais utilizados. Assim, e num registo muito destacado, a WISC-III ocupa o 1º lugar entre os testes mais utilizados nas duas investigações. Ainda em

comum e em posições cimeiras (8 primeiros lugares) encontram-se instrumentos de avaliação cognitiva: raciocínio (CPM, SPM; BPRD/BPR), inteligência (WAIS-III) e memória/funções executivas (FCR). A segunda tendência nesta classificação comparativa refere-se a um declínio sistemático quer no uso de testes projetivos (Rorschach; CAT-A/CAT-H; Desenho da Família; TAT), quer de inventários de avaliação da personalidade (MMPI /MMPI-2, NEO-PI-R, *Mini-Mult*). Esta descida de posições é extensiva a outros testes (Griffiths, TP, Bender, D-A-P). A terceira tendência comparativa remete para uma subida de posições de dois instrumentos: MPCR/MPCR-P e BDI/BDI-II.

Discussão

Cotejados com as outras investigações recenseadas, os dados do presente estudo sugerem que é possível concluir que em Portugal existem instrumentos de referência disponíveis que: viabilizam um protocolo de avaliação representativo; abrangendo a maior parte das dimensões de avaliação (nomeadamente, inteligência, funções neurocognitivas, personalidade, psicopatologia, comportamento, rendimento escolar) e remetendo para diferentes grupos etários. De um modo geral, os testes de inteligência geral/inteligência não-verbal/raciocínio (nomeadamente, WISC-III, WAIS-III, SPM, CPM) mantêm uma supremacia de utilização, à semelhança da investigação de Diniz et al. (2007) e da literatura internacional. Mais especificamente, as escalas de inteligência de Wechsler (sobretudo WISC e WAIS) ocupam um lugar de destaque na presente investigação corroborando dados dos estudos internacionais recenseados. À semelhança do presente estudo, a WISC surge frequentemente em primeiro lugar na lista dos testes mais utilizados, que inclui também a WAIS (e.g., Quinnell & Bow, 2001). Em outras investigações, que incluem habitualmente a WISC, a WAIS é listada em primeiro lugar (e.g., Lubin et al., 1971). A “popularidade” das escalas de inteligência de Wechsler é igualmente evidenciada na investigação de Evers et al. (2012), que coloca a WISC no grupo dos 10 testes mais utilizados em 16 dos 17 países inquiridos, e a WAIS em 14 desses

17 países. No mesmo sentido, uma revisão de 23 estudos realizados entre 1982 e 1992 conclui que as escalas de inteligência de Wechsler são os instrumentos mais utilizados (Piotrowski & Keller, 1992).

Esta estabilidade observada ao longo do tempo na utilização muito frequente das escalas de inteligência de Wechsler tem várias explicações. Estes instrumentos proporcionam informação importante, através da utilização da escala completa, de conjuntos de subtestes (permitindo calcular, por exemplo, índices fatoriais relativos à Velocidade de Processamento ou à Memória de Trabalho, frequentemente mais valorizados do que os QIs), ou de subtestes isolados. Por este motivo, as escalas de inteligência de Wechsler são bastante usadas (e mencionadas) em contextos muito diversos de exercício profissional da Psicologia: escolar (e.g., Benson et al., 2019), clínico e de saúde, educação (e.g., Hernández et al., 2020), neuropsicologia clínica (e.g., Rabin et al., 2016), neuropsicologia forense (e.g., Lees-Haley et al., 1996), e no exame de comportamentos de simulação, com base no teste da Memória de Dígitos (e.g., Martin et al., 2015). As Matrizes Progressivas de Raven (SPM, CPM) estão igualmente mencionadas no estudo de Diniz et al. (2007) e na investigação de Evers et al. (2012) surgindo na lista de 10 instrumentos mais usados em 10 países, e em primeiro lugar em três de 17 países (cf. no mesmo sentido, Hernández et al., 2020; Oakland et al., 2016). Estes testes, tradicionalmente considerados de inteligência/raciocínio não-verbais, são igualmente usados em contextos de avaliação neuropsicológica, nomeadamente no exame de funções viso-espaciais ou do raciocínio, uma das dimensões das funções executivas (e.g., Echemendia & Harris, 2004).

Ausentes do estudo português anterior de Diniz e colaboradores (2007), os instrumentos de rastreio cognitivo, mencionados no presente estudo (MMSE, MoCA, ACE-R), são atualmente objeto de utilização progressiva e sistemática, quer na avaliação de idosos, quer de adultos. Estes três instrumentos, que permitem igualmente o exame rápido de diferentes funções cognitivas, dispõem de estudos recentes de validação em Portugal abrangendo vários grupos clínicos e possuem normas representativas que cruzam variáveis como

a idade e a escolaridade, que tornam as suas pontuações mais diferenciadoras dos desempenhos cognitivos (e.g., casos de declínio cognitivo). Os testes de rastreio cognitivo (MMSE, sobretudo) são mencionados, em lugares de relevo, em estudos recentes (e.g., Giromini et al. 2022 (4º teste mais utilizado); Hernández et al., 2020 (12º); Wright et al., 2017 (4º)).

Por outro lado, a investigação reconhece a crescente expansão e especialização em diferentes áreas da psicologia (e da avaliação) como é particularmente o caso da neuropsicologia e da psicologia forense (e.g., Wright et al., 2017). O presente estudo testemunha bem esta especialização da neuropsicologia com a menção a sete testes neuropsicológicos. Com exceção do teste de TP (uma medida de atenção seletiva e sustentada e de velocidade de processamento) e da BANC (e.g., Simões et al., 2016), uma bateria de 15 testes desenvolvida e editada recentemente em Portugal, que avalia a atenção, orientação, memória, linguagem, funções executivas, motricidade e lateralidade, todos os outros testes são igualmente mencionados em investigações internacionais congéneres, no âmbito de estudos de neuropsicologia clínica ou forense, psicologia escolar ou psicologia clínica: FCR, com várias classificações possíveis, por exemplo, enquanto teste de funções viso-espaciais, memória/evocação, funções executivas (e.g., Rabin et al., 2005); d2, uma medida de atenção (e.g., Hernández et al., 2020); Bender, uma medida de funções viso-perceptivas e viso-motoras (e.g., Slick & Craig, 1991); Escalas de Memória de Wechsler (WMS-III), com recurso à escala completa ou a subtestes isolados (e.g., Sullivan & Bowden, 1997); TMT, uma medida de atenção e funções executivas (e.g., Archer et al., 2006). Na investigação de Diniz et al. (2007) constam três testes neuropsicológicos: FCR, TP e Bender. Deve ser mencionado que a WMS-III, o d2 e o TMT ainda não tinham nesta data estudos publicados em Portugal. Por conseguinte, quer a crescente disponibilização de testes neuropsicológicos (que estes “novos instrumentos” apenas em parte exemplificam), quer instrumentos clássicos como as escalas de inteligência de Wechsler (e.g., Sherman et al., 2023), alargam consideravelmente a possibilidade de concretizar avaliações (neuro)psicológicas mais rigorosas e completas.

No presente estudo, o incremento que acabamos de observar nos testes neuropsicológicos não é visível no caso dos instrumentos de avaliação mais específicos da psicologia forense. Apesar de mais de uma centena de psicólogos a trabalhar em contexto forense ter respondido ao inquérito, os dados parecem comprovar conclusões de outras investigações realizadas sobre os testes mais comumente usados por psicólogos forenses, que referem que os instrumentos de avaliação clínica tradicionais continuam a ser decisivos em muitas avaliações psicológicas forenses, apesar da multiplicação de testes considerados mais específicos, desenvolvidos para responder a questões particulares de natureza forense (Archer et al., 2016). No presente estudo, não surge qualquer menção à utilização de instrumentos mais específicos desta área, apesar de existirem vários com estudos de validação na população portuguesa (cf. Simões, Almeida & Gonçalves, 2017) em domínios relevantes como a psicopatologia (e.g., Escala de Psicopatologia de Hare – Versão Revista (PCL-R)); a avaliação do risco (e.g., Guião de Avaliação Histórica, Clínica e de Gestão do Risco: Versão 2 (HCR-20V2)), ou a simulação (e.g., Teste de Simulação de Problemas de Memória (TOMM)). Este último constitui o teste mais mencionado nas investigações centradas na avaliação da simulação, exagero de sintomas ou esforço reduzido (e.g., Brooks et al., 2016; Martin et al., 2015). Neste âmbito, importa reconhecer que testes tradicionais (e.g., SPM, FCR, WMS-III, MMPI-2) podem ser igualmente usados no exame de comportamentos de simulação (e.g., Dandachi-FitzGerald et al., 2013).

Os instrumentos de avaliação da personalidade ocupam um lugar importante nos protocolos de avaliação. Os testes projetivos nomeados no presente estudo (Rorschach, CAT-A/CAT-H, Desenho da Família, TAT, RATC/RATC-2, Bar-Ilan, H-T-P) surgem frequentemente nas primeiras investigações internacionais nas listas dos testes mais usados (e.g., Hutton et al., 1992). Nestas investigações, as listas dos 10 instrumentos mais usados incluía frequentemente vários testes projetivos e um menor número de testes de inteligência (e.g., Sundberg, 1961; Wade et al., 1978). Progressivamente observa-se neste tipo de investigação uma crescente utilização de escalas de inteligência de Wechsler, de outro tipo de

instrumentos de avaliação da personalidade (e.g., MMPI) e de escalas de avaliação do comportamento (e.g., CBCL), com uma concomitante diminuição do uso de testes projetivos (e.g., Muñiz & Fernández-Hermida, 2010).

À semelhança do presente estudo, o Rorschach é frequentemente o teste projetivo com mais menções (e.g., Sundeberg, 1961). Um predomínio evidente de “testes psicométricos” sobre os “testes projetivos” é reconhecido em países como Espanha e Portugal (Muñiz et al., 1999). Esta tendência contrasta fortemente com os resultados citados por Muñiz et al. (1999) em alguns países latino-americanos (Argentina, Brasil), onde os testes projetivos estão em maioria (cf. no mesmo sentido, Reppold et al., 2020). Em outras investigações, e exemplificando bem a diminuição de uso destes instrumentos, não surge referência a nenhum teste projetivo (e.g., Hernández et al., 2020, numa lista de 25 testes), ou apenas um único é indicado: o Rorschach (Bow et al., 2006; Evers et al., 2012); ou o CAT/TAT, na lista de 10 testes mais usados para avaliação de crianças e adolescentes, por psicólogos escolares, em 64 países (Oakland et al., 2016). Este declínio de utilização de testes projetivos é observado no presente estudo também comparativamente à anterior investigação portuguesa realizada por Diniz e colaboradores (2007). Nesta comparação constata-se uma descida sistemática de posições nos diferentes instrumentos projetivos comuns às duas investigações: Rorschach, CAT-A/CAT-H, Desenho da Família, TAT.

Estas constatações de declínio podem não ser alheias à sugestão de que estes testes não cumprem padrões psicométricos de utilização considerados indispensáveis (e.g., em procedimentos de natureza legal, não preencherem critérios de aceitabilidade/admissibilidade em tribunal) (e.g., Areh et al., 2022). Esta contestação a “técnicas de avaliação questionáveis e controversas” como o Rorschach, TAT, CAT, RATC, Desenho da Família, H-T-P, listados no presente estudo, é desenvolvida mais aprofundadamente em relação ao Rorschach, o teste projetivo simultaneamente mais estudado e mais comumente presente nas diferentes listas consultadas relativas aos testes mais utilizados. Mais especificamente, as críticas ao uso do Rorschach, em contexto forense, referem

a falta de artigos revistos por pares que examinem a testabilidade empírica da teoria subjacente, a falta de replicabilidade dos diagnósticos elaborados a partir do Sistema de Exner, a falta de um método de avaliação de estilos de resposta, a falta de validade apropriada para responder às questões legais, problemas estes que podem abrir a possibilidade a relatos e depoimentos que causem danos razoavelmente previsíveis e evitáveis (e.g., Areh et al., 2022). Acresce, ainda, a existência de diversos e muito distintos sistemas de cotação disponíveis considerados subjetivos, incluindo o Sistema Compreensivo de cotação de Exner (Hunsley et al., 2015). Uma perspetiva contrária foi defendida recentemente por De Ruiter et al. (2023), que consideram o Rorschach um instrumento útil em avaliações na área da saúde mental forense, com base no recurso ao novo sistema de cotação *Rorschach Performance Assessment System* (R-PAS). De modo convergente, na sua revisão sobre testes empiricamente fundamentados, Archer et al. (2016) observaram que “embora pouco utilizado, o Rorschach [administrado, cotado e interpretado de acordo com o Sistema Compreensivo de Exner] foi considerado a única medida projetiva de personalidade que atendia às linhas orientadoras de Daubert” (p. 7).

Esta controvérsia relativa à validade do uso de testes projetivos tem sido, por vezes, perspetivada num registo simplista como um conflito entre académicos e investigadores céticos que contestam a sua validade (e.g., Garb, 1999) e psicólogos clínicos que os utilizam e consideram úteis na conceptualização do caso e na planificação da intervenção (e.g., Weiner, 1999).

Outros testes projetivos (desenho H-T-P, desenho da família), mencionados no presente estudo, e em muitas outras investigações internacionais, apresentam uma fiabilidade e validade muito reduzidas ou mesmo inexistentes e são mesmo considerados “ultrapassados” (e.g., Heilbrun et al., 2014) e, por isso, é defendida a sua não inclusão nos protocolos de avaliação (e.g., Archer & Wygant, 2012). Por outro lado, o recurso ainda frequente a testes de desenhos pode ser compreensível, não pelas suas qualidades psicométricas, mas, sobretudo, quando se considera o seu potencial contributo para o estabelecimento de uma relação de avaliação, no caso das crianças, que habitualmente estão

familiarizadas com este tipo de tarefa (Cashel, 2002). Neste sentido, tem sido sugerido que muitos psicólogos tendem a preferir instrumentos de administração e cotação rápida e fácil de usar (testes de desenho) comparativamente a outros testes projetivos como o Rorschach e TAT (Goh et al., 1981).

Os outros instrumentos ditos objetivos de avaliação da personalidade, os questionários ou inventários de autorrelato, mencionados no presente estudo, constam também das listas de outras investigações internacionais, ocupando frequentemente posições antes preenchidas por testes projetivos. Neste plano, o MMPI/MMPI-2, é presença muito habitual nestas investigações, por vezes é mesmo o teste mais usado e, nas primeiras investigações, foi frequentemente o único instrumento objetivo de personalidade mencionado nos 10 instrumentos mais utilizados, quase exclusivamente ocupados por testes inteligência e projetivos (e.g., Harrison et al., 1988). Outros inventários surgem igualmente nas listas, com destaque para o MCMI (e.g., Wright et al., 2017), ainda que o seu uso seja considerado “inaceitável” em contextos forenses, de acordo com a pesquisa de Lally (2003), seguido pelo NEO-PI-R (e.g., Evers et al., 2012). Estes instrumentos (MMPI, MCMI, NEO-PI-R e, também, o *Mini-Mult*) são igualmente mencionados na investigação de Diniz et al. (2007) em posições classificativas mais elevadas por comparação com o presente estudo, sugerindo também aqui uma eventual diminuição na utilização de inventários de avaliação da personalidade.

Os Inventários de Sintomas Psicopatológicos identificados no presente estudo acompanham diferentes faixas etárias e são referidos em várias investigações: BDI/BDI-II (e.g., Hernández et al., 2020; único inventário de sintomas mencionado no anterior estudo de Diniz et al., 2007), SCL-90 (e.g., Watkins et al., 1995), YSR (e.g., Cashel, 2002), CDI/CDI-2 (e.g., Archer & Newsom, 2000) e GDS (e.g., Hernández et al., 2020).

Comparativamente ao estudo anterior de Diniz e colaboradores (2007), surgem pela primeira vez na presente investigação vários instrumentos desenvolvidos e/ou atualizados mais recentemente, com crescente utilização, como é o caso das escalas de avaliação do comportamento, completadas por pais (e.g., CBCL) ou professores (e.g., TRF) que

têm adquirido uma aceitação significativa também na investigação internacional (e.g., Benson et al., 2019; Oakland et al., 2016). Tal facto não é alheio a uma lógica de avaliação multi-informador e multi-método, bem exemplificada na área de avaliação das crianças e adolescentes através do modelo ASEBA, que integra outros instrumentos (inventário de psicopatologia YSR ou a entrevista SCICA respondidos pela criança ou adolescente), para além do CBCL e TRF.

Na pesquisa de Diniz et al. (2007) também não existe qualquer menção aos testes de realização/rendimento escolar. Acompanhando uma tendência observada em trabalhos internacionais (e.g., Benson et al., 2019), no presente estudo surgem mencionados quatro testes de rendimento escolar (e.g., BAPAE, O Rei).

Os resultados obtidos permitem-nos, ainda, alguns comentários sobre certos grupos de instrumentos. Por exemplo, a avaliação de interesses vocacionais e preferências profissionais está representada na presente investigação através de dois inventários: IPP/IPP-R e COPS/COPS-R. Este tipo de instrumentos não consta da pesquisa de Diniz et al. (2007), mas outros inventários de interesses são presença em outros estudos (e.g., Watkins et al., 1995). Por outro lado, na avaliação do desenvolvimento infantil a Griffiths é igualmente mencionada na lista de Diniz et al. (2007), mas não em nenhum dos estudos internacionais recenseados. Por último, em algumas investigações, as entrevistas e a observação são técnicas que correspondem a instrumentos específicos e formais integrantes e reconhecidos como indispensáveis ao processo e nos protocolos de avaliação psicológica (e.g., Benson et al., 2019). A não referência ao uso de outras entrevistas para além da SICA (e.g., entrevistas diagnósticas standardizadas) ou à observação direta poderá significar uma não utilização de instrumentos formais estruturados, a não alusão ao uso de medidas mais informais (e.g., entrevistas semiestruturadas) ou, mais provavelmente, ser consequência da pergunta colocada aos psicólogos remeter para apenas para o uso de testes e questionários.

Os dados das investigações relativas aos testes mais utilizados ajudam a definir a evolução ao longo dos anos das práticas e dos protocolos de avaliação psicológica. Assim, e com base na

identificação dos instrumentos mais utilizados, é possível identificar um protocolo progressivamente mais alargado de testes, que são mencionados no presente estudo. A investigação de Watkins (1991) sugere: WISC, WAIS, MMPI, Rorschach e TAT. Mais recentemente, e a partir de um estudo internacional que considerou respostas de psicólogos de 17 países europeus, constata-se que os testes mais comumente usados continuam a ser testes psicométricos clássicos de inteligência (WISC, WAIS, Raven), mas aos testes projetivos de personalidade atrás indicados acrescenta-se agora, sobretudo, uma maior ênfase nos instrumentos de autorresposta, considerados mais objetivos (e.g., MMPI, NEO-PI-R), a que se acrescentam os inventários de avaliação da psicopatologia (e.g., BDI, SCL-90) (Evers et al., 2012; Wright et al., 2017), as escalas de avaliação do comportamento (e.g., CBCL) (Benson et al., 2019; Wright et al., 2017) e os instrumentos de rastreio cognitivo (e.g., MMSE) (e.g., Hernández et al., 2020; Wright et al., 2017).

Encontramos no presente estudo outros cenários de utilização de testes que merecem ser comentados. Por um lado, existe menção relativamente isolada ao recurso a versões mais recentes de alguns instrumentos que não dispõem de estudos no nosso país (WISC-IV, WISC-V, WAIS-IV, WPPSI-III, WPPSI-IV, MMPI). Nestes casos, admitimos tratar-se de utilizações mais didáticas associadas ao ensino de instrumentos de avaliação psicológica. Por outro, constata-se a indicação de uso de versões anteriores com estudos em Portugal, mas com normas muito datadas (e.g., WISC). Consideramos aqui a possibilidade de incompletude na indicação genérica do nome do instrumento (mais provável) ou a formação insuficiente/falta de atualização de alguns profissionais. Existe ainda referência à utilização de versões sem estudos sistemáticos conhecidos em Portugal (WISC-R, WMS-R, MMPI, MCMI/MCM-II/MCMI-III/MCMI-IV). Nestes casos poder-se-á tratar de “versões de trabalho” utilizadas por alguns profissionais e que antecederam respetivamente as aferições portuguesas posteriores (WISC-III, WMS-III e MMPI-2/MMPI-2-RF), podendo também traduzir falta de atualização de conhecimentos e práticas. São igualmente mencionados e referidos como muito utilizados, na presente investigação, alguns

instrumentos (WISC-III, WAIS-III, WPPSI-R, MMPI-2, NEO-PI-R, ACE-R) que não correspondem às respetivas versões internacionais mais recentes (e.g., WISC-V, WAIS-IV, WPPSI-IV, MMPI-2-RF/MMPI-3, NEO-PI-3, ACE-III). Este último cenário remete para um potencial conflito ético entre usar uma versão mais antiga do teste, alegadamente ultrapassada, mas que dispõe de estudos psicométricos consistentes e normas representativas, e uma versão revista, atual, mais moderna, supostamente superior desse mesmo teste, mas sem estudos psicométricos/normas ou com estudos pouco representativos (Lees-Haley et al., 1996), como é, por exemplo, o caso do ACE-III em Portugal. Por outro lado, vários instrumentos mencionados não estão comercialmente editados, conhecendo-se a existência de mais do que uma única tradução dos itens (e.g., Conners, GDS).

Finalmente, os dados deste tipo de investigações levantam outras questões conhecidas. Por que motivo alguns testes menos representativos das práticas de avaliação mais atuais e sem estudos sistemáticos conhecidos são muito utilizados e outros instrumentos, mais relevantes e com estudos de validação publicados, são completamente ignorados? Neste sentido, algumas investigações têm concluído, de forma explícita, que alguns dos testes mais mencionados (“populares”) não são necessariamente os mais válidos (e.g., Butler et al., 1991). Ou seja, a utilização de certos instrumentos muito utilizados é questionável, ou por nunca terem provado a validade dos seus resultados, ou por não disporem de estudos de validação e normas representativas que assegurem uma utilização credível e útil, considerando os grupos nos quais esses instrumentos são especificamente administrados. Neste sentido, importa reconhecer problemas com alguns dos testes incluídos nesta lista quando se consideram critérios mais exigentes de escrutínio da sua validade. Especificamente, alguns instrumentos utilizados (e.g., Desenho da Família, *Mini-Mult*, Bender, H-T-P) não têm normas ou estudos psicométricos conhecidos ou aceitáveis (cf. no mesmo sentido Reppold et al., 2020). Para além dos problemas com estes testes é necessário admitir limitações associadas ao número reduzido de estudos de validação de muitos dos instrumentos, incluindo alguns testes considerados mais representativos e utilizados (e.g., WAIS-III,

WMS-III), alguns dos quais justificando provavelmente o desenvolvimento de novas normas ou a sua atualização com base em versões mais recentes (e.g., WISC-V relativamente à WISC-III). Muitos outros instrumentos carecem de estudos de validade e normas adicionais igualmente à luz dos critérios mais exigentes de utilização de testes, nomeadamente os relativos à investigação com amostras ou grupos mais específicos (e.g., perturbações do neurodesenvolvimento, grupos clínicos e forenses particulares). Ou seja, é urgente mais pesquisa relativa aos instrumentos já utilizados, incluindo alguns dos testes com mais citações, que recorram adicionalmente a outros modelos de investigação de evidência psicométrica (e.g., teoria de resposta ao item) e não apenas, como é mais comum, à teoria clássica dos testes.

Por outro lado, e com implicações para o ensino/formação/treino e divulgação de instrumentos de avaliação psicológica, deve ser sublinhado que, apesar de se encontrarem disponíveis em Portugal, são ignorados na presente investigação outros testes importantes. Abrangendo áreas relevantes de avaliação, alguns destes testes não estão editados comercialmente, mas têm estudos de validade realizados, sobretudo na última década, muitas vezes apenas com publicações em revistas internacionais e abrangendo domínios como: aprendizagem e memória (e.g., Teste de Recordação Livre e Seletiva Guiada de Palavras (FCSRT; e.g., Lemos et al., 2015), Teste de Aprendizagem e Memória Auditiva de Rey (RAVLT; Cavaco et al., 2015); funções executivas (Stroop; Fernandes, 2013); inteligência pré-mórbida (Teste de Leitura de Palavras Irregulares (TeLPI; e.g., Alves et al., 2013); atividades de vida diária/funcionalidade (Inventário de Avaliação Funcional de Adultos e Idosos (IAFAI; e.g., Sousa et al., 2017); e simulação (TOMM, e.g., Simões, Sousa et al., 2017)). Esta situação ilustra a necessidade de uma melhor divulgação de instrumentos relevantes disponíveis e com estudos realizados em Portugal.

Na lista dos testes não citados é de sublinhar igualmente a omissão a testes computadorizados. A sua menção nas investigações internacionais é aliás muito rara. Uma investigação realizada nos Estados-Unidos e no Canadá revelou que, de uma lista de 693 instrumentos, apenas 40 (6%) eram

instrumentos computadorizados, aliás raramente utilizados (Rabin et al., 2014). Existem explicações plausíveis para esta omissão, nomeadamente, o número reduzido de instrumentos computadorizados comercializados, o preço elevado destes instrumentos ou, no caso de alguns testes mais representativos (e.g., Cantab, NIH Toolbox-Cognitiva), a sua frequente revisão que torna as versões anteriores rapidamente não utilizáveis.

Por último, observa-se ainda uma forte dependência de instrumentos desenvolvidos noutros países. Dos 50 instrumentos listados, apenas quatro testes foram originalmente desenvolvidos em Portugal: BPRD/BPR, Era uma Vez, BANC e O REI.

Ainda com implicações para a prática da avaliação psicológica, e relativamente a outros instrumentos importantes mencionados em investigações internacionais, importa acrescentar que está prevista a edição e comercialização próximas em Portugal do *Behavior Rating Inventory of Executive Function* (BRIEF-P, BRIEF-A, BRIEF-2) (e.g., Rabin et al., 2016), BDI-II, MMPI-2-RF, MMPI-A, *Personality Assessment Inventory* (PAI) (Wright et al., 2017), e a WISC-V encontra-se em fase de aferição.

Finalmente, no plano da investigação importa sublinhar o papel decisivo que têm tido instituições como a Fundação para a Ciência e a Tecnologia e a Fundação Calouste Gulbenkian, ao financiarem, em concursos competitivos, projetos orientados para a adaptação e validação de testes e outros instrumentos muito utilizados a nível internacional (e.g., WISC-III, WPPSI-R, MPCR, MoCA, TeLPI, BDI-II) ou originalmente desenvolvidos em Portugal (e.g., BPRD, BANC, IAFAI).

Conclusões

Acompanhando tendências internacionais, a presente investigação permite concluir o seguinte: (i) os instrumentos mais tradicionais de avaliação da inteligência (escalas de inteligência de Wechsler e, particularmente, a WISC) e raciocínio (Matrizes de Raven) estão entre os testes mais utilizados pelos psicólogos portugueses; (ii) existe um incremento evidente do recurso a testes neuropsicológicos (e.g., FCR, d2) e a testes de rastreio cognitivo (e.g., MMSE) sugerindo que a avaliação psicológica alargou o seu âmbito e/ou se tornou mais neuropsicológica, bem como se

observa um aumento do uso das escalas de avaliação do comportamento (nomeadamente, CBCL e TRF, integrantes do modelo ASEBA de Achenbach); (iii) constata-se uma diminuição, tendencialmente progressiva, na utilização de testes projetivos entre os psicólogos portugueses; (iv) os inventários de sintomas psicopatológicos (e.g., BDI/BDI-II) apresentam um maior número de utilizações do que os inventários de personalidade (e.g., MMPI/MMPI-2); (v) os testes de rendimento/realização escolar surgem pela primeira vez nas investigações portuguesas relativas aos testes mais usados, podendo sugerir alterações nas práticas dos psicólogos e/ou nas necessidades da avaliação psicológica em contextos escolares (e.g., Maluf et al., 2022); e (vi) é indispensável a implementação de mais investigação relativa aos instrumentos já utilizados, uma melhor divulgação de instrumentos relevantes disponíveis com estudos realizados em Portugal e mais formação na área dos testes psicológicos. Por último, importa destacar que, num contexto de diminuição da durabilidade dos conhecimentos da psicologia profissional (Neimeyer et al., 2012), se torna urgente um maior investimento na área dos testes e da avaliação psicológica, quer por parte das instituições de ensino da Psicologia ao nível da formação atualizada, quer dos psicólogos no que diz à atualização de conhecimentos.

Referências

- Ackerman, M. J., Bow, J. N., & Mathy, N. (2021). Child custody evaluation practices: Where we were, where we are, and where we are going. *Professional Psychology: Research and Practice, 52*(4), 406-417.
<https://doi.org/10.1037/pro0000380>
- Almeida, L. S., Simões, M. R., Seabra-Santos, M. J., & Almiro, P. A. (2021, Julho). Utilização de testes psicológicos em Portugal. *PSIS21, 22*.
<https://dashboard.mazsystems.com/reader/74517?page=16>
- Alves, L., Simões, M. R., Martins, C., Freitas, S., & Santana, I. (2013). TeLPI performance in subjects with mild cognitive impairment and Alzheimer's disease: A validation study. *Alzheimer Disease and Associated Disorders, 27*(4), 324-329.
<http://doi.org/10.1097/WAD.0b013e31827bdc8c>
- Archer, R. P., & Newsom, C. R. (2000). Psychological test usage with adolescent clients: Survey update. *Assessment, 7*(3), 227-335.
<https://doi.org/10.1177/107319110000700303>
- Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment, 87*(1), 84-94.
https://doi.org/10.1207/s15327752jpa8701_07
- Archer, R. P., Wheeler, E. M. A., & Vauter, R. A. (2016). Empirically supported forensic assessment. *Clinical Psychology: Science and Practice, 23*, 348-364.
<https://doi.org/10.1111/cpsp.12171>
- Areh, I., Verkampt, F., & Allan, A. (2022). Critical review of the use of the Rorschach in European courts. *Psychiatry, Psychology and Law, 29*(2), 183-205.
<https://doi.org/10.1080/13218719.2021.1894260>
- Benson, N. F., Floyd, R. G., Kranzler, J. H., Eckert, T. L., Fefere, S. A., & Morgana, G. B. (2019). Test use and assessment practices of school psychologists in the United States: Findings from the 2017 National Survey. *Journal of School Psychology, 72*, 29-48.
<https://doi.org/10.1016/j.jsp.2018.12.004>
- Bornstein, R. F. (2017). Evidence-based psychological assessment. *Journal of Personality Assessment, 99*(4), 435-445.
<https://doi.org/10.1080/00223891.2016.1236343>
- Bow, J. N., Gould, J. W., Flens, J. R., & Greenhut, D. (2006). Testing in child custody evaluations-selection, usage, and Daubert admissibility: A survey of psychologists. *Journal of Forensic Psychology Practice, 6*(2), 17-38.
https://doi.org/10.1300/J158v06n02_02
- Brooks, B. L., Ploetz, D. M., & Kirkwood, M. W. (2016). A survey of neuropsychologists' use of validity tests with children and adolescents. *Child Neuropsychology, 22*(8), 1001-20.
<https://doi.org/10.1080/09297049.2015.1075491>

- Brown, W. R., & McGuire, J. M. (1976). Current psychological assessment practices. *Professional Psychology, 7*(4), 475-484. <https://doi.org/10.1037/0735-7028.7.4.475>
- Butler, M., Retzlaff, P., & Vanderploeg, R. (1991). Neuropsychological test usage. *Professional Psychology: Research and Practice, 22*(6), 510-512. <https://doi.org/10.1037/0735-7028.22.6.510>
- Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice, 31*(2), 141-154. <https://doi.org/10.1037/0735-7028.31.2.141>
- Cashel, M. L. (2002). Child and adolescent psychological assessment: Current clinical practices and the impact of managed care. *Professional Psychology: Research and Practice, 33*(5), 446-453. <https://doi.org/10.1037/0735-7028.33.5.446>
- Cavaco, S., Gonçalves, A., Pinto, C., Almeida, E., Gomes, F., Moreira, I., Fernandes, J., & Teixeira-Pinto, A. (2015). Auditory Verbal Learning Test in a large nonclinical Portuguese population. *Applied Neuropsychology: Adult, 22*(5), 321-331. <https://doi.org/10.1080/23279095.2014.927767>
- Cook, J. R., Hausman, E. M., Jensen-Doss, A., & Hawley, K. M. (2017). Assessment practices of child clinicians: Results from a national survey. *Assessment, 24*(2), 210-221. <https://doi.org/10.1177/1073191115604353>
- Dandachi-FitzGerald, B., Ponds, R. W. H. M., & Merten, T. (2013). Symptom validity and neuropsychological assessment: A survey of practices and beliefs of neuropsychologists in six European countries. *Archives of Clinical Neuropsychology, 28*, 771-783. <https://doi.org/10.1093/arclin/act073>
- de Ruiter, C., Giromini, L., Meyer, G. J., King, C. M., & Rubin, B. A. (2023). Clarifying sound and suspect use of the Rorschach in forensic mental health evaluations: A response to Areh et al. (2022). *Psychology Injury and Law, 16*, 158-176. <https://doi.org/10.1007/s12207-023-09472-6>
- Di Pucchio, A., Vanacore, N., Marzolini, F., Lacorte, E., Di Fiandra, T., I-DemObs Group, & Gasparini, M. (2018). Use of neuropsychological tests for the diagnosis of dementia: A survey of Italian memory clinics. *British Medical Journal Open, 8*(3), e017847. <https://doi.org/10.1136/bmjopen-2017-017847>
- Diniz, A. M., Almeida, L. S., & Pais, L. G. (2007). Contextos profissionais e práticas de avaliação psicológica: Inquérito aos psicólogos portugueses. *Psico-USF, 12*(1), 1-12. <http://dx.doi.org/10.1590/S1413-82712007000100002>
- Donoso, O. A., Hernandez, B., & Horin, E. V. (2010). Use of psychological tests within vocational rehabilitation. *Journal of Vocational Rehabilitation, 32*(3), 191-200. <https://doi.org/10.3233/JVR-2010-0509>
- Echemendia, R. J., & Harris, J. G. (2004). Neuropsychological test use with Hispanic/Latino populations in the U.S.: Part II of a national survey. *Applied Neuropsychology, 11*, 4-12. https://doi.org/10.1207/s15324826an1101_2
- Evers, A., Muñiz, J., Bartram, D., Boben, D., Egeland, J., Fernández-Hermida, J. R., Frans, Ö., Gintiliené, G., Hagemester, C., Halama, P., Iliescu, D., Jaworowska, A., Jiménez, P., Manthouli, M., Matesic, K., Schittekatte, M., Sümer, H. C., & Urbánek, T. (2012). Testing practices in the 21st century: Developments and European psychologists' opinions. *European Psychologist, 17*(4), 300-319. <https://doi.org/10.1027/1016-9040/a000102>
- Fernandes, S. (2013). *Stroop: Teste de Cores e de Palavras. Manual*. Hogrefe.
- Fernandez-Ballesteros, R., De Bruyn, E. E. J., Godoy, A., Horke, L. F., Laak, J. T., Vizcarro, C., Westhoof, K., Westmeyer, H., & Zaccagini, J. L. (2001). Guidelines for the Assessment Process (GAP): A proposal for discussion. *European Journal of Psychological Assessment, 17*(3), 187-200. <https://doi.org/10.1027/1015-5759.17.3.187>
- 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. <https://doi.org/10.1177/107319119900600402>
- Giromini, L., Pasqualini, S., Corgiat Loia, A., Pignolo, C., Di Girolamo, M., & Zennaro, A. (2022). A survey of practices and beliefs of Italian psychologists regarding malingering and

- Symptom Validity Assessment. *Psychological Injury and Law*, 15(2), 128-140.
<https://doi.org/10.1007/s12207-022-09452-2>
- Goh, D. S., Teslow, C. J., & Fuller, G. B. (1981). The practice of psychological assessment among school psychologists. *Professional Psychology*, 12, 696-706.
<https://doi.org/10.1037/0735-7028.12.6.696>
- Groth-Marnat, G., & Wright, A. J. (2016). *Handbook of psychological assessment* (6th ed.). Wiley.
- Guilmette, T. J., Faust, D., Hart, K., & Arkes, H. R. (1990). A national survey of psychologists who offer neuropsychological services. *Archives of Clinical Neuropsychology*, 5(4), 373-392.
[https://doi.org/10.1016/0887-6177\(90\)90016-I](https://doi.org/10.1016/0887-6177(90)90016-I)
- Harrison, P. L., Kaufman, A. S., Hickman, J. A., & Kaufman, N. L. (1988). A survey of tests used for adult assessment. *Journal of Psychoeducational Assessment*, 6(3), 188-198.
<https://doi.org/10.1177/073428298800600301402/313>
- Haynes, S. N., Smith, G. T., & Hunsley, J. D. (2019). *Scientific foundations of clinical assessment* (2nd ed.). Routledge.
- Heilbrun, K., DeMatteo, D., Holliday, S. B., & LaDuke, C. (2014). Teaching point: Whether and how to criticize material from the records. In K. Heilbrun, D. DeMatteo, H. S. Brooks, & C. LaDuke (Eds.), *Forensic mental health assessment* (pp. 411-412). Oxford University Press.
- Hernández, A., Elosua, P., Wechsler, S., & Muñiz, J. (2020). El uso de los tests en España y Latinoamérica: Perspectivas actuales y retos futuros. *Anuario Internacional de Revisiones en Psicología*, 1, 67-94.
<https://doi.org/10.14635/REVPSY.0.4>
- Hill, D., & Demetrio, S. (2019). Clinical-forensic psychology in Canada: A survey of practitioner characteristics, attitudes, and psychological assessment practices. *Canadian Psychology/Psychologie Canadienne*, 60(1), 55-63.
<https://doi.org/10.1037/cap0000152>
- Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment*, 15(4), 446-455.
<https://doi.org/10.1037/1040-3590.15.4.446>
- Hunsley, J., Lee, C. M., Wood, J. M., & Taylor, W. (2015). Controversial and questionable assessment techniques. In S. O. Lilienfeld, S. Y. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 42-82). Guilford.
- Hutton, J. B., Dubes, R., & Muir, S. (1992). Assessment practices of school psychologists: Ten years later. *School Psychology Review*, 21, 271-284.
<https://doi.org/10.1080/02796015.1992.12085614>
- Lally, S. J. (2003). What tests are acceptable for use in forensic evaluations? A survey of experts. *Professional Psychology: Research and Practice*, 34, 491-498.
<https://doi.org/10.1037/0735-7028.34.5.491>
- 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(1), 45-51.
[https://doi.org/10.1016/0887-6177\(95\)00011-9](https://doi.org/10.1016/0887-6177(95)00011-9)
- Lemos, R., Simões, M. R., Santiago, B., & Santana, I. (2015). The Free and Cued Selective Reminding Test: Validation for mild cognitive impairment and Alzheimer's disease. *Journal of Neuropsychology*, 9(2), 242-257.
<https://doi.org/10.1111/jnp.12048>
- Louitt, C. M., & Browne, C. G. (1947). The use of psychometric instruments in psychological clinics. *Journal of Consulting Psychology*, 11(1), 49-54.
<https://doi.org/10.1037/h0056136>
- Lubin, B., Larsen, R. M., & Matarazzo, J. D. (1984). Patterns of psychological test usage in the United States: 1935-1982. *American Psychologist*, 39(4), 451-454.
<https://doi.org/10.1037/0003-066X.39.4.451>
- Lubin, B., Wallis, R. R., & Paine, C. (1971). Patterns of psychological test usage in the United States: 1935-1969. *Professional Psychology*, 2, 70-74.
<https://doi.org/10.1037/h0031544>
- Maluf, R. G., Lopes, J. A. L., & Oliveira, C. R. G. (2022). Práticas de avaliação psicológica escolar: Uma revisão sistemática da literatura. *Revista Iberoamericana de Diagnóstico e Avaliação Psicológica*, 62(2),

- 185-215.
<https://doi.org/10.21865/RIDEP62.1.14>
- Martin, M., Allan, A., & Allan, M. M. (2001). The use of psychological tests by Australian psychologists who do assessments for the courts. *Australian Journal of Psychology*, 53(2), 77-82.
<https://doi.org/10.1080/00049530108255127>
- Martin, P. K., Schroeder, R. W., & Odland, A. P. (2015). Neuropsychologists' validity testing beliefs and practices: A survey of North American professionals. *The Clinical Neuropsychologist*, 29(6), 741-776.
<https://doi.org/10.1080/13854046.2015.1087597>
- McLaughlin, J. L., & Kan, L. Y. (2014). Test usage in four common types of forensic mental health assessment. *Professional Psychology: Research and Practice*, 45(2), 128-135.
<https://doi.org/10.1037/a0036318>
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, 56(2), 128-165.
<https://doi.org/10.1037/0003-066X.56.2.128>
- Muñiz, J., Hernández, A., & Fernández-Hermida, J. R. (2020). La utilización de los tests en España: El punto de vista de los psicólogos. *Papeles del Psicólogo*, 41, 1-15.
<https://doi.org/10.23923/pap.psicol2020.2921>
- Muñiz, J., & Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los tests. *Papeles del Psicólogo*, 31(1), 108-121.
<https://www.papelesdelpsicologo.es/pdf/1801.pdf>
- Muñiz, J., Bartram, D., Evers, A., Boben, D., Matesic, K., Glabeke, K., Fernández-Hermida, J. R., & Zaal, J. N. (2001). Testing practices in European countries. *European Journal of Psychological Assessment*, 17(3), 201-211.
<https://doi.org/10.1027/1015-5759.17.3.201>
- Muñiz, J., Prieto, G., Almeida, L., & Bartram, D. (1999). Test use in Spain, Portugal and Latin American Countries. *European Journal of Psychological Assessment*, 15(2), 151-157.
<https://doi.org/10.1027//1015-5759.15.2.151>
- Neimeyer, G. J., Taylor, J. M., & Rozensky, R. H. (2012). The diminishing durability of knowledge in professional psychology: A Delphi Poll of specialties and proficiencies. *Professional Psychology: Research and Practice*, 43(4), 364-371.
<https://doi.org/10.1037/a0028698>
- Noronha, A. P. P., Primi, R., & Alchieri, J. C. (2005). Instrumentos de avaliação mais conhecidos/utilizados por psicólogos e estudantes de psicologia. *Psicologia: Reflexão e Crítica*, 18(3), 390-401.
<https://dx.doi.org/10.1590/S0102-79722005000300013>
- Oakland, T. Douglas, S., & Kane, H. (2016). Top 10 standardized tests used internationally with children and youth by school psychologists in 64 countries: A 24-year follow-up study. *Journal of Psychoeducational Assessment*, 34(2), 166-176.
<https://doi.org/10.1177/0734282915595303>
- Peterson, C. H., Lomas, G. I., Neukrug, E. S., & Bonner, M. W. (2014). Assessment use by counselors in the United States: Implications for policy and practice. *Journal of Counseling & Development*, 92, 90-99.
<https://doi.org/10.1002/j.1556-6676.2014.00134.x>
- Piotrowski, C., & Keller, J. W. (1992). Psychological testing in applied settings: A literature review from 1982-1992. *The Journal of Training and Practice in Professional Psychology*, 6, 74-82.
<https://psycnet.apa.org/record/1993-24038-001>
- Quinnell, F. A., & Bow, J. N. (2001). Psychological tests used in child custody evaluations. *Behavioral Sciences and the Law*, 19, 491-501. <https://doi.org/10.1002/bsl.452>
- Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, 20, 33-65.
<https://doi.org/10.1016/j.acn.2004.02.005>
- Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in test-usage practices of clinical neuropsychologists in the United States and Canada over a 10-year period: A follow-up

- survey of INS and NAN members. *Archives of Clinical Neuropsychology*, 31, 206-230.
<https://doi.org/10.1093/arclin/acw007>
- Rabin, L. A., Spadaccini, A. T., Brodale, D. L., Grant, K. S., Elbulok-Charcape, M. M., & Barr, W. B. (2014). Utilization rates of computerized tests and test batteries among clinical neuropsychologists in the United States and Canada. *Professional Psychology: Research and Practice*, 45(5), 368-377.
<https://doi.org/10.1037/a0037987>
- Reppold, C. T., Wechsler, S. M., Almeida, L. S., Elosua, P., & Hutz, C. S. (2020). Perfil dos psicólogos brasileiros que utilizam testes psicológicos: Áreas e instrumentos utilizados. *Psicologia: Ciência e Profissão*, 40, e201348, 1-14.
<https://doi.org/10.1590/1982-3703003201348>
- Ryan, A. M., & Sackett, P. R. (1987). A survey of individual assessment practices by I/O psychologists. *Personnel Psychology*, 40(3), 455-488. <https://doi.org/10.1111/j.1744-6570.1987.tb00610.x>
- Ryan, J., Dai, X-D., & Zheng, L. (1994). Psychological test usage in the People's Republic of China. *Journal of Psychoeducational Assessment*, 12(4) 324-330.
<https://doi.org/10.1177/073428299401200402>
- Seabra-Santos, M. J., Simões, M. R., Almiro, P. A., & Almeida, L. S. (2021). Utilização de testes para avaliar crianças dos 0 aos 7 anos: Resultados de um inquérito a psicólogos portugueses. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 60(3), 81-94.
<https://doi.org/10.21865/RIDEP60.3.07>
- Seabra-Santos, M. J., Almiro, P. A., Simões, M. R., & Almeida, L. S. (2019). Testes psicológicos em Portugal: Atitudes, problemas e perfil dos utilizadores. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 53(4), 101-112.
<https://doi.org/10.21865/RIDEP53.4.08>
- Shapiro, E. S., & Heick, P. F. (2004). School psychologist assessment practices in the evaluation of students referred for social /behavioral/emotional problems. *Psychology in the Schools*, 41(5), 551-561.
<https://doi.org/10.1002/pits.10176>
- Sherman, E. M. S., Tan, J. E., & Hrabok, M. (Eds.) (2023). *A compendium of neuropsychological tests: Fundamentals of neuropsychological assessment and test reviews for clinical practice* (4th ed.). Oxford University Press.
- Simões, M. R., Albuquerque, C. P., Pinho, M. S., Vilar, M., Pereira, M., Lopes, A. F., Seabra-Santos, M. J., Alberto, I., Lopes, C., Martins, C., & Moura, O. (2016). *Bateria de Avaliação Neuropsicológica de Coimbra (BANC): Manual Técnico*. Cegoc.
- Simões, M. R., Almeida, L. S., & Gonçalves, M. M. (Coords.) (2017). *Psicologia Forense: Instrumentos de Avaliação*. Pactor.
- Simões, M. R., Almeida, L. S., Seabra-Santos, M., & Almiro, P. A. (2018, 12-13 de Outubro). Testes e outros instrumentos mais utilizados por psicólogos na avaliação de adultos idosos: Dados portugueses de um projeto internacional. Comunicação apresentada no 32º Congresso do Grupo de Estudos de Envelhecimento Cerebral e Demência. Lisboa.
- Simões, M. R., Sousa, L. B., Fonseca, M. S., Garcia, S. M., Pinho, J. I., Soares, D. M., Silva, D. C., Oliveira, C. S., Mota, M., Maior, F. S., Jesus, S. S., Fernandes, S. P., Faustino, J. P., Cruz, A. P., Barreto, M., Dias, I. T., Perdiz, C. P., Duro, D., Santana, I., & Pinho, M. S. (2017). Teste de Simulação de Problemas de Memória (TOMM). In M. R. Simões, L. S. Almeida, & M. M. Gonçalves (Coords.), *Psicologia Forense: Instrumentos de Avaliação* (pp. 353-375). Pactor.
- Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology*, 19, 465-473.
<https://doi.org/10.1016/j.acn.2003.04.001>
- Slick, D. J., & Craig, P. L. (1991). Neuropsychological assessment in public psychiatric hospitals: The changing state of the practice – 1979 to 1989. *Archives of Clinical Neuropsychology*, 6, 73-80.
[https://doi.org/10.1016/0887-6177\(91\)90023-3](https://doi.org/10.1016/0887-6177(91)90023-3)
- Smith, S. R., Gorske, T., Wiggins, C., & Little, J. A. (2010). Personality assessment use by clinical neuropsychologists. *International Journal of Testing*, 10, 6-20.
<https://doi.org/10.1080/15305050903534787>

- Sousa, L. B., Vilar, M., Prieto, G., & Simões, M. R. (2017). Inventário de Avaliação Funcional de Adultos e Idosos (IAFAI). In M. M. Gonçalves, M. R. Simões, & L. S. Almeida (Coords.), *Psicologia clínica e da saúde: instrumentos de avaliação* (pp. 235-249). Factor.
- Sullivan, K., & Bowden, S.C. (1997). Which tests do neuropsychologists use? *Journal of Clinical Psychology, 53*(7), 657-661. [https://doi.org/10.1002/\(SICI\)1097-4679\(199711\)53:7<657::AID-JCLP3>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-4679(199711)53:7<657::AID-JCLP3>3.0.CO;2-F)
- Sundberg, N. D. (1961). The practice of psychological testing in clinical services in the United States. *American Psychologist, 16*(2), 79-83. <https://doi.org/10.1037/h0040647>
- Tarroja, M. C., Alcalá, M. A., Simon, P. D., & Sanchez, J. D. (2020). A review of psychological assessment practice in the Philippines: What do some practitioners say?. *Philippine Journal of Psychology, 53*, 81-115. <https://doi.org/10.31710/pjp/0053.04>
- Turcotte, M. (1994). Use of tests in the Canadian employment services. *Journal of Employment Counseling, 31*(4), 188-199. <https://doi.org/10.1002/j.2161-1920.1994.tb00978.x>
- Wade, T. C., & Baker, T. B. (1977). Opinions and use of psychological tests: A survey of clinical psychologists. *American Psychologist, 32*(10), 874-882. <https://doi.org/10.1037/0003-066X.32.10.874>
- Wade, T. C., Baker, T. B., Morton, T. L., & Baker, L. J. (1978). The status of psychological testing in clinical psychology: Relationships between test use and professional activities and orientations. *Journal of Personality Assessment, 42*(1), 3-10. https://doi.org/10.1207/s15327752jpa4201_1
- Watkins, C. E., Campbell, V. L., & McGregor, P. (1988). Counseling psychologists' uses of and opinions about psychological tests: A contemporary perspective. *The Counseling Psychologist, 16*(3), 476-486. <https://doi.org/10.1177/0011000088163010>
- Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26*(1), 54-60. <https://doi.org/10.1037/0735-7028.26.1.54>
- Watkins, C. E., Jr. (1991). What have surveys taught us about the teaching and practice of psychological assessment?. *Journal of Personality Assessment, 56*(3), 426-437. http://dx.doi.org/10.1207/s15327752jpa5603_5
- Weiner, I. B. (1999). What the Rorschach can do for you: Incremental validity in clinical applications. *Assessment, 6*, 327-339. <https://doi.org/10.1177/107319119900600404>
- Wertz, M., Hank, L., Hausam, J., Konrad, N., Schiltz, K., Imhoff, R., & Rettenberger, M. (2022). The use and reporting practice of psychological tests in German risk and criminal responsibility expert reports. *Psychology, Crime & Law, 1*-18. <https://doi.org/10.1080/1068316X.2022.2063286>
- Wright, C. V., Beattie, S. G., Galper, D. I., Church, A. S., Bufka, L. F., Brabender, V. M., & Smith, B. L. (2017). Assessment practices of professional psychologists: Results of a national survey. *Professional Psychology: Research and Practice, 48*(2), 73-78. <https://doi.org/10.1037/pro0000086>