

O Zulliger no Sistema Compreensivo: Um Estudo de Fidedignidade

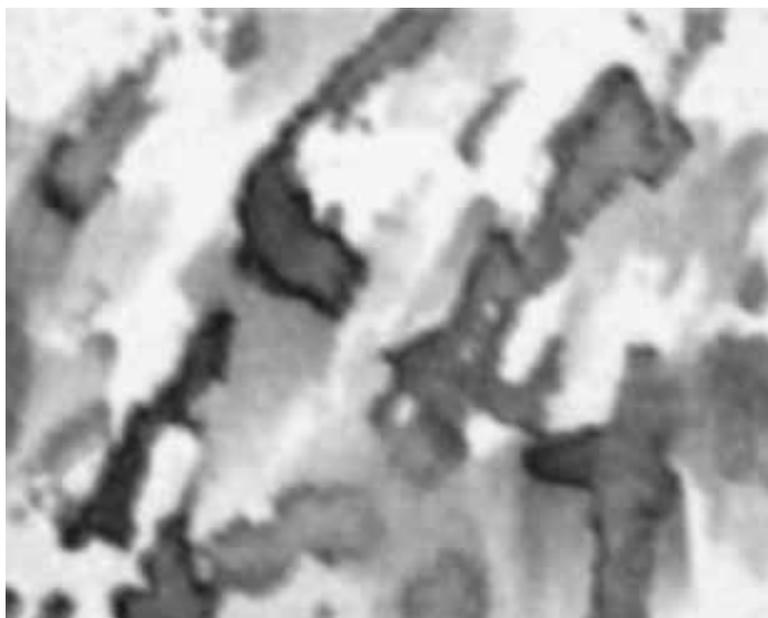
The Zulliger in the Comprehensive System:
A Reliability Study

El Zulliger en el Sistema Compreensivo:
Un Estudio de Fidedignidade

**Anna Elisa de Villemor-Amaral,
Maria Aparecida dos Santos
Machado & Ana Paula Porto
Noronha**

Universidade
São Francisco

Artigo



Resumo: Os testes psicológicos têm passado por verificações criteriosas de validade e precisão a fim de atender às exigências éticas e científicas. O objetivo desta pesquisa foi averiguar a precisão do Teste Zulliger no sistema compreensivo por meio do teste-reteste. Participaram do estudo 25 sujeitos, do sexo masculino, não pacientes, estudantes de teologia do interior de São Paulo. Utilizou-se um questionário para dados socioeconômico-familiar e o Zulliger. A aplicação dos instrumentos foi individual, durando cerca de 40 minutos, com intervalo de 5 meses entre teste-reteste. Dos protocolos, 25% foram recodificados por um juiz às cegas, obtendo-se índices satisfatórios de precisão. Foram selecionados 16 indicadores para se realizar as estatísticas descritivas e estudos de correlação com método de Pearson. Desses, dez alcançaram índices de precisão satisfatórios, entre 0,60-0,99; quatro tiveram uma precisão entre 0,40-0,60, todos com nível de significância entre 0,01 e 0,05. Os resultados foram satisfatórios e contribuíram para demonstrar a precisão do instrumento.

Palavras-chave: Técnica projetiva. Teste de Zulliger. Precisão. Avaliação psicológica.

Abstract: The psychological tests are the focus of several validity and reliability studies in order to attend ethical and scientific criteria. The objective of this study was to verify the reliability of the Zulliger in the comprehensive system through the testing-retesting method. The participants were 25 male, non clinical cases, theology students in the interior of São Paulo. A questionnaire was used for social-economic data, and the Zulliger. The application was performed individually during approximately 40 minutes, with a pause of 5 months between testing and retesting. 25% of the protocols were evaluated by an independent judge. Among the data, 16 indicators were selected for descriptive statistics and correlation with the Pearson method. From these indicators, 10 of them reached a satisfactory precision rate, between 0,60-0,99, four had a precision between 0,40-0,60, all of them with a significance level between 0,01 and 0,05. The results were satisfactory and contributed to demonstrate the precision of this instrument.

Keywords: Projective techniques. Zulliger Test. Reliability. Psychological assessment.

Resumen: Los testes psicológicos han pasado por verificaciones con criterios de validez y precisión a fin de atender a las exigencias éticas y científicas. El objetivo de esta pesquisa fue averiguar la precisión de la Prueba Zulliger en el sistema compreensivo por medio del test-retest. Participaron del estudio 25 sujetos, del sexo masculino, no pacientes, estudiantes de teología del interior de Sao Paulo. Se utilizó un cuestionario para datos socioeconómico-familiar y el Zulliger. La aplicación de los instrumentos fue individual, durando cerca de 40 minutos, con intervalo de 5 meses entre test-retest. De los protocolos, 25% fueron recodificados por un juez a ciegas, lográndose índices satisfactorios de precisión. Fueron seleccionados 16 indicadores para realizarse las estadísticas descriptivas y estudios de correlación con método de Pearson. De estos, diez alcanzaron índices de precisión satisfactorios, entre 0,60-0,99; cuatro tuvieron una precisión entre 0,40-0,60, todos con nivel de significación entre 0,01 y 0,05. Los resultados fueron satisfactorios y aportaron para demostrar la precisión del instrumento.

Palabras clave: Técnica Proyectiva. Test de Zulliger. Precisión. Evaluación psicológica.

Os testes psicológicos são técnicas utilizadas com o objetivo de observar, descrever e classificar o comportamento de indivíduos de forma sistemática. Para Pasquali (2001), esses instrumentos podem ser classificados em testes psicométricos e testes projetivos, e são utilizados na avaliação psicológica para os mais variados fins bem como na pesquisa para o desenvolvimento da ciência psicológica.

Muitos questionamentos são levantados em torno desses instrumentos, não só pela comunidade científica mas também pelos psicólogos e sociedade. Problemas sobre o desenvolvimento de pesquisas para validação, precisão e normatização, bem como sobre a elaboração de manuais, têm sido temas de várias investigações (Noronha, Sartori, Freitas, & Ottati, 2001).

A comunidade científica tem desenvolvido pesquisas para atualizar e melhorar os instrumentos de avaliação psicológica, principalmente quanto aos critérios psicométricos. A psicometria é um ramo da Psicologia que faz interface com a estatística, cujo objetivo é avaliar adequadamente as diferenças individuais sob o enfoque positivista, no qual a quantificação e a generalização dos fenômenos são fundamentais. Seus parâmetros de qualidade são padronização, validade e fidedignidade (Pasquali, 1999).

A comunidade científica tem desenvolvido pesquisas para atualizar e melhorar os instrumentos de avaliação psicológica, principalmente quanto aos critérios psicométricos. A psicometria é um ramo da Psicologia que faz interface com a estatística, cujo objetivo é avaliar adequadamente as diferenças individuais sob o enfoque positivista, no qual a quantificação e a generalização dos fenômenos são fundamentais. Seus parâmetros de qualidade são padronização, validade e fidedignidade (Pasquali, 1999).

Segundo Anastasi e Urbina (2000), a padronização diz respeito à uniformidade de procedimento na aplicação e na pontuação do teste, o que auxilia na objetividade e no controle das variáveis que poderão intervir no processo de testagem. Validade é a verificação daquilo que o teste mede e quão bem ele o faz. A fidedignidade, ou confiabilidade, refere-se à consistência dos escores obtidos pelas mesmas pessoas quando elas são reexaminadas com o mesmo teste em diferentes ocasiões, com diferentes conjuntos de itens equivalentes, ou sob outras condições variáveis de exame.

A consistência dos resultados no procedimento de teste-reteste mostra o quanto os resultados de uma pessoa podem ser generalizados para ocasiões diferentes. Quanto maior a estabilidade, menos suscetíveis serão os escores às mudanças aleatórias nas condições dos testandos ou do ambiente de testagem e, em conseqüência, maior será sua confiabilidade. Cada teste deve conter uma informação sobre o tipo de fidedignidade que foi utilizado, o intervalo das aplicações e a descrição das características dessa amostra. A medida de fidedignidade caracteriza o teste quando ele é aplicado em condições-padrão e com pessoas semelhantes àquelas que constituem a amostra normativa (Anastasi & Urbina, 2000).

Conseguir todos esses parâmetros psicométricos nos instrumentos psicológicos não é uma tarefa simples, porque os testes se diferenciam em seus construtos e na atividade proposta. Para as técnicas projetivas, a tarefa é ainda mais difícil. Güntert (2001) enfatiza que o modelo de ciência basicamente positivista, oriundo das ciências naturais, não é apropriado para grande parte dos fenômenos humanos, principalmente no que concerne aos critérios de objetividade máxima. Nesse sentido, não seria nem a questão do que é bom ou não, e sim, do que é adequado ou inadequado para um determinado objeto que se pretende investigar (Alves-Mazzotti, 2001).

Villemor-Amaral (2006) destaca quatro pontos que considera os principais na problemática dos procedimentos científicos utilizados, quando se busca a validação e a precisão dos métodos projetivos. O primeiro diz respeito à natureza dos fenômenos estudados, que nem sempre são observáveis diretamente pelo comportamento. O segundo ponto está ligado ao fato de determinados fenômenos ocorrerem uma única vez em um único indivíduo, talvez não se repetindo nem em outro e nem no próprio sujeito. O terceiro diz respeito à natureza da tarefa, que é proposta pela técnica projetiva ao se apresentarem estímulos pouco definidos, que podem gerar um número muito grande de respostas, e cada participante pode dar uma variedade de respostas diferentes. O último ponto aborda a questão de que, em pesquisa, quanto mais complexo o número de variáveis envolvidas na análise, maior deve ser o número de sujeitos estudados, o que é pouco compatível com investigações desse tipo, que despendem muito tempo para aplicação e análise dos dados, resultando geralmente em uma amostra mais reduzida do que a que seria ideal.

Chandler, citado por Alves (2006), preconiza que muitos autores afirmam que as técnicas projetivas não são testes, e, por isso, deveriam

estar isentas das restrições normalmente aplicadas aos instrumentos psicométricos. Entretanto, Alves (2006) sustenta que questões sobre a validade e a precisão precisam ser tratadas, considerando o problema sobre a adequação de se adotar os mesmos procedimentos para técnicas projetivas e para testes de aptidão.

Considerando que a validação de um teste abrange a avaliação de sua sensibilidade, fidedignidade e validade, Anzieu (1978) responde negativamente a essa questão, afirmando que

Os testes projetivos não exploram uma variável única, mas descrevem um indivíduo em termos de um esquema dinâmico de variáveis inter-relacionadas. Sua validação não consiste em verificar se os indivíduos testados se ordenam segundo um grau em que neles se manifesta a variável única. Assemelha-se muito mais a um processo científico de validação de hipóteses. (p. 227)

Assim, Villemor-Amaral (2006) ressalta que um teste não é válido por si só, mas apresenta indicadores de validade para as inferências feitas a partir dos resultados, considerando o contexto e o objetivo da avaliação. Afirma ainda que, para as técnicas projetivas, isso também seria pertinente, pois nelas são encontrados alguns indicadores, de tal modo que, à medida que estes aumentem, o grau de confiabilidade da técnica será maior.

Segundo Alves (2006), a sensibilidade nas técnicas projetivas é uma constatação de que elas podem refletir as mudanças que ocorrem na personalidade em função da idade, das doenças e de circunstâncias excepcionais, sempre à luz da teoria psicológica. Nesse sentido, a fidedignidade poderia ser investigada por meio da estabilidade das respostas às aplicações sucessivas, isolando a variável aprendizagem. Ao lado disso, também poderia ser obtida pela concordância de juízes diferentes ao avaliarem os mesmos protocolos, trabalhando separadamente.

Anastasi e Urbina (2000), ao descreverem os métodos para avaliar a precisão dos instrumentos, destacam que teste e reteste, apesar de aparentemente simples e diretos, não são adequados à maioria dos testes psicológicos. O intervalo entre as testagens e a possível memorização das respostas pode influenciar nos resultados. As autoras afirmam que o intervalo não deve ultrapassar seis meses, evitando assim a influência de mudanças desenvolvimentais. Em um intervalo curto, o sujeito pode recordar as respostas dadas na aplicação anterior. Destacam ainda que somente os testes que não são afetados pela repetição são indicados ao reteste.

Para as referidas autoras, verificar a fidedignidade de técnicas projetivas por meio de teste-reteste dificilmente seria adequado, uma vez que raramente as pessoas dão respostas semelhantes em aplicações sucessivas. Exner (1994), entretanto, fez estudos para verificar a fidedignidade do Rorschach-sistema compreensivo com teste e reteste em um grupo de não pacientes, adultos, com intervalo de 36 e 39 meses, e, em outro momento, fez pesquisas com intervalos de 12 e 14 meses, também com adultos não pacientes. O pesquisador partiu da hipótese de que os indivíduos apresentam estilos de respostas claramente preferentes, que se manifestam na maioria de suas respostas, e que a evidência de tais estilos deve aparecer consistentemente no transcorrer de administrações repetidas. Suas pesquisas mostram que uma grande quantidade de dados indica que as chamadas marcas de personalidade permanecem consistentes ao longo do tempo.

Quanto à variável memória, Exner (1994) afirma que terá uma influência menor sobre o que o sujeito viu o que viu do que sobre o que foi dito. Em outras palavras, no caso de testes de percepção de manchas de tinta, como o Rorschach e o Zulliger, importa mais como a pessoa viu do que o que ela viu para

se considerar a fidedignidade do método, já que se analisa a estrutura perceptiva, e não a resposta em si.

O método de Rorschach e o Teste de Zulliger são instrumentos que, por meio de manchas de tinta não estruturadas, provocam no sujeito associações com registros de experiências que geram as respostas, que, por sua vez, serão transformadas em categorias e interpretadas no todo, sob os múltiplos aspectos do dinamismo perceptivo-associativo, que refletem a personalidade. É um conjunto de variáveis por meio das quais se podem quantificar os aspectos funcionais e dinâmicos da personalidade (Weiner, 1994).

Desde a década de 70, John Exner empreendeu esforços para reunir os conhecimentos e as investigações até então desenvolvidas para o método de Rorschach pelos principais sistematizadores nos Estados Unidos, tais como Samuel Beck, Marguerite Hertz, Bruno Klopfer, Zigmunt Piotrowski e também Rapaport, com o objetivo de unificar as principais contribuições em um só sistema que pudesse superar as complicações advindas de um paralelismo que impossibilitava a troca entre os estudiosos. O autor também se empenhou em selecionar as características empiricamente aceitáveis do ponto de vista psicométrico, ou seja, dos itens que se demonstrassem válidos e precisos, garantindo com isso uma uniformidade da metodologia e do sistema de codificação das respostas.

Como discutido por Nascimento e Güntert (2000), Exner estabeleceu critérios de classificação claros e facilmente aplicáveis aos seguintes elementos: localização das respostas (W, D, Dd, S), qualidade evolutiva (DQ), que complementa a localização, diferenciando os processos perceptivos envolvidos na formação da resposta e visando a discriminar respostas vagas, respostas simples que envolvem uma forma e respostas que introduzem uma relação entre os elementos (v, o, + e v/+).

Os determinantes são os elementos mais complexos da classificação, e indicam quais características da mancha estimularam o sujeito e influenciaram a formação das respostas; são eles: forma (F), movimento (M, FM, m – ativo e passivo), cor (C, CF, FC, Cn), cor acromática (C', C'F, FC'), sombreado (Y, YF, FY, V, VF, FV, T, TF, FT), forma e dimensão (FD), pares (2) e reflexos (Fr e rF). A composição dos determinantes também é importante, classificando-se em determinante único ou misto.

Para a qualidade formal, segundo Nascimento e Güntert (2000), Exner atribui um código para as respostas superiores/superelaboradas (+), para as ordinárias (o), para as incomuns (u) e para as que distorcem a forma em maior ou menor grau (-). Existem 26 categorias de conteúdo: H, (H), Hd, (Hd), Hx, A, (A), Ad, (Ad), An, Art, Ay, Bl, Bt, Cg, Cl, Ex, Fi, Fd, Ge, Hh, Ls, Na, Sc, Sx, Xy, todas de fácil identificação, inclusive os conteúdos idiossincráticos (Id). As respostas populares (P) são respostas que aparecem com muita frequência na população. É atribuído um código quando diferentes elementos da resposta são integrados pela atribuição de uma relação entre eles, e é dado um valor conforme o grau de facilidade ou de dificuldade para se estabelecer essa relação (Z). Em cada protocolo, estima-se a frequência com que tais valores aparecem (Zf) e a soma desses valores (Zsum). Ainda compõem os elementos do sistema os códigos especiais, que são atribuídos pela presença de uma característica incomum na fala do sujeito ao emitir as respostas, sendo, para verbalizações inusuais, 6 códigos (DV, DR, INCOM, FABCOM, CONTAM e ALOG), para perseveração ou falha na integração, 2 códigos (PSV e CONFAB), para características especiais do conteúdo, 4 códigos (AB, AG, COP e MOR), para respostas personalizadas, 1 código (PER), e, para fenômeno especial de cor, 1 código (CP). Os elementos da classificação das respostas são organizados no sumário estrutural, que representa a

composição das freqüências dos códigos, mais as proporções, razões, porcentagens e derivações numéricas. A partir desses dados, são geradas muitas hipóteses importantes relativas às características e ao funcionamento psicológico. Há mais de 25 anos se tem trabalhado com o sistema compreensivo, e este continua a se desenvolver de maneira dinâmica, com integrações, revisões e acréscimos que são feitos constantemente por meio de investigações fundamentadas em metodologias de pesquisa atualizadas (Nascimento & Güntert, 2000).

O Teste de Zulliger (Z-Teste) foi construído durante a segunda grande guerra por Hans Zulliger a partir de estudos com o Rorschach. Nesse contexto, sendo responsável pela seleção de oficiais para as forças armadas suíças, constatou a inviabilidade da aplicação individual de 10 cartões, porque despendia muito tempo. Dessa experiência resultou um conjunto de três figuras, que foi aplicado individualmente, pareado com o Rorschach, em uma amostra de 800 pessoas, o *Zulliger Diapositive-Test (Z-Test)* (Vaz, 2000). A vantagem de permitir aplicações e análises mais rápidas possibilita sua utilização em um maior contingente de indivíduos, o que o torna um instrumento aplicável em uma variedade de contextos e circunstâncias mais diversificados do que o Rorschach. Esse fato justifica os esforços de pesquisa que visem ao seu aprimoramento.

Sendo assim, nas duas últimas décadas, pesquisadores têm se empenhado em adaptar o Teste de Zulliger para o sistema compreensivo que Exner desenvolveu. Brinkmann (1998), do Departamento de Psicologia da Universidade de Concepción, no Chile, estudou numa amostra de 300 sujeitos em situação não clínica, adultos em processo de seleção de pessoal e jovens postulantes a carreiras universitárias, com a aplicação coletiva do Zulliger, a freqüência das respostas globais (G), detalhes grandes (D), detalhes pequenos ou de respostas

pouco usuais (Dd), espaços brancos (S) e conteúdo humano (H e Hd) e animal (A e Ad), e apresentou uma lista de respostas populares para cada diapositivo. Para tabulação e codificação das respostas, o pesquisador utilizou o sistema compreensivo de Rorschach, por considerar que é atualmente o sistema de codificação mais difundido para trabalhar com o Rorschach, e afirma que não vê sentido utilizar outro método no Zulliger quando este tem o mesmo propósito que o Rorschach.

Na Argentina, Zdunic (1999) desenvolveu estudos com o objetivo de adaptar o Teste de Zulliger para o sistema compreensivo na aplicação individual, o que resultou na publicação de um manual. A autora destaca que o teste é um instrumento de reconhecida importância, pois tem se mostrado válido na avaliação da personalidade e é amplamente utilizado na seleção de pessoal na área do trabalho.

Pesquisas para adaptar o Zulliger ao sistema compreensivo também estão sendo feitas no Brasil. Villemor-Amaral, Primi e Cardoso (2004) organizaram inicialmente um atlas com áreas de localização das respostas e suas freqüências para atribuição dos escores D e Dd. Posteriormente, elaboraram uma lista de respostas para cada área, visando a estudar suas freqüências para elaborar a lista de qualidade formal e popularidade.

Outros estudos foram desenvolvidos por Franco e Villemor-Amaral (2004) para a adaptação do Teste de Zulliger conforme os critérios de classificação de respostas e análise dos resultados do sistema compreensivo. As pesquisadoras verificaram quais os indicadores do teste são característicos para o diagnóstico psicopatológico da esquizofrenia, identificando como indicadores significativos os códigos especiais: DV, INCOM, CONTAM, ALOG, AB e PER.

Outros estudos de validade do Zulliger sistema compreensivo foram realizados

O Teste de Zulliger é composto de um conjunto de três lâminas com manchas de tinta, de 18,5 por 25 cm, sendo uma acromática, uma policromática e outra em vermelho e preto. As manchas são estímulos não estruturados que têm por objetivo promover a projeção de aspectos da dinâmica subjetiva do indivíduo. A aplicação é individual, e leva em média de 30 a 40 minutos. Os estudos de validade e precisão desse instrumento estão em fase final de análise bem como os dados normativos e de padronização para publicação do manual, porém, no momento, ainda não é possível descrevê-los (Villemor-Amaral et al., no prelo).

por Villemor-Amaral e Lamounier (2006), Ferreira e Villemor-Amaral (2005), Machado e Villemor-Amaral (2006), Villemor-Amaral, Farah e Machado (2007).

Como o Zulliger – sistema compreensivo tem sido submetido a uma série de estudos visando a atender aos critérios científicos e ao desenvolvimento do próprio instrumento, o presente estudo tem por objetivo verificar a sua fidedignidade por meio do teste-reteste, uma vez que não há publicação sobre sua precisão com esse método. A fim de contribuir com esse parâmetro psicométrico, selecionou-se uma amostra específica, estudantes de teologia, com características sociodemográficas niveladas.

Método

Participantes

Participaram da pesquisa 25 sujeitos do sexo masculino, não pacientes, estudantes de uma instituição privada de ensino superior de teologia, de uma cidade do interior do Estado de São Paulo. A média de idade na primeira aplicação foi de 28,32 anos (DP=7,739), variando de 20 a 47 anos. A média de escolaridade foi de 13,60 anos (DP 2,021), com o mínimo de 12 anos de estudo e o máximo de 20 anos, sendo que, de 12 a 17 anos, correspondem 96% da amostra, e 20 anos, a 4%, ou seja, a somente um sujeito. Os voluntários participaram da primeira aplicação do Teste de Zulliger em novembro de 2006, e da segunda aplicação, em abril de 2007.

Instrumentos

Foi utilizado um questionário para obter informações sobre dados socioeconômico-familiares e o Teste de Zulliger. O questionário continha perguntas sobre mudanças na constituição familiar, necessidade de procura para tratamento de problemas emocionais ou

psiquiátricos, consumo de drogas, álcool ou medicação, alteração de humor e dificuldades profissional ou financeira causadoras de estresse.

O Teste de Zulliger é composto de um conjunto de três lâminas com manchas de tinta, de 18,5 por 25 cm, sendo uma acromática, uma policromática e outra em vermelho e preto. As manchas são estímulos não estruturados que têm por objetivo promover a projeção de aspectos da dinâmica subjetiva do indivíduo. A aplicação é individual, e leva em média de 30 a 40 minutos. Os estudos de validade e precisão desse instrumento estão em fase final de análise bem como os dados normativos e de padronização para publicação do manual, porém, no momento, ainda não é possível descrevê-los (Villemor-Amaral et al., no prelo).

Procedimento

Foram contatados os voluntários que haviam participado da primeira aplicação do teste para uma segunda aplicação, em um intervalo de cinco meses, considerando o fato de que um intervalo de até seis meses é o sugerido pela literatura (Anastasi & Urbina, 2000). O Termo de Consentimento Livre e Esclarecido foi assinado em duas vias, ficando uma com o voluntário e outra com a examinadora. Aplicou-se o questionário, para obter informações de possíveis alterações socioeconômico-familiares entre as duas aplicações que pudessem interferir nos escores dos testes. A aplicação foi realizada de forma verbal, e as respostas foram anotadas pela aplicadora, e, em seguida, foi aplicado o teste.

A aplicação do Zulliger foi individual, em um tempo aproximado de 40 minutos, e consistiu em apresentar as lâminas com manchas de tinta e perguntar *o que poderia ser* para o sujeito. Após a apresentação das três lâminas,

iniciou-se o inquérito, com a reapresentação das lâminas na mesma seqüência, repetindo-se o que foi dito e perguntando onde o sujeito via determinado conteúdo, para que, em seguida, fosse destacado na folha de localização. Logo após, perguntou-se o que na lâmina lhe havia dado aquela idéia. As respostas foram registradas na íntegra, bem como as reações do sujeito.

As codificações dos protocolos foram feitas pela pesquisadora após cada etapa. Da soma total de protocolos (50), 25% foi submetido à codificação de um juiz independente. Além desse procedimento, as respostas que geraram dúvidas na codificação entre os juízes foram submetidas à análise de um terceiro juiz especialista, que definiu a melhor codificação em cada discordância entre os dois juizes anteriores.

Os dados foram tabulados e submetidos à análise da média e desvio-padrão, comparação com o grupo normativo de universitários e por fim, foi feita a correlação dos resultados do teste-reteste, utilizando-se o método de Pearson.

Resultados

Os protocolos foram codificados no que se refere à localização, qualidade evolutiva, determinantes, qualidade formal, conteúdos, popularidade, reflexos/pares, atividade organizativa e códigos especiais.

Um percentual de aproximadamente 25% (13) do total de protocolos referentes à primeira e à segunda aplicações foi sorteado e submetido à análise independente de um juiz. Na avaliação da precisão, focaram-se os indicadores referentes aos determinantes e conteúdos, principais elementos avaliados nesse trabalho. A escolha dos indicadores para a presente análise se deu pelo fato de eles serem passíveis de um maior grau de interferência subjetiva, e buscou-se reduzir tal efeito. A Tabela 1 apresenta os resultados da precisão entre juízes.

Tabela 1. Resultados das análises de concordância entre juízes.

Número de protocolo	Quantidade de indicadores avaliados	% de concordância entre juízes
1	10	60%
2	14	100%
3	14	78,57%
4	12	91,66%
5	12	91,66%
6	20	95%
7	14	92,86%
8	16	68,75%
9	14	92,86%
10	10	70%
11	12	91,66%
12	10	90%
13	16	87,50%

Conforme a Tabela 1, o valor mínimo de concordância foi de 60%, e o máximo foi de 100%. O cálculo da média dos valores de concordância entre juízes foi de 85,42%. Um terceiro juiz especialista analisou as respostas que geraram dúvidas na codificação em cada discordância entre os juízes anteriores, definindo qual seria a melhor codificação.

Entre os vários indicadores passíveis de análise no Zulliger-sistema compreensivo, foram selecionados, para as análises estatísticas, o indicador R, que se refere ao número de respostas dadas, o indicador S, que aparece quando há integração do espaço em branco nas respostas, e os indicadores W, D, Dd, que correspondem às áreas de localização das respostas. Os indicadores M, C, CF, FC fazem parte dos determinantes, e o H, Hd, (H), (Hd), dos conteúdos. Os indicadores Sum_SH, EB e H:Hd+(H)+(Hd) são proporções, frações e soma ponderada, realizadas para uma compreensão de aspectos dinâmicos do sujeito. Considerando que estes são alguns dos principais indicadores para a análise da dinâmica individual relacionada aos componentes cognitivos, afetivos e da relação interpessoal, que constituem pontos centrais na interpretação de um protocolo (Exner, 1994), neste trabalho, as análises restringiram-se a eles. A Tabela 2 apresenta os valores da média e desvio-padrão dos indicadores selecionados, referentes à primeira e à segunda aplicações.

Tabela 2. Média, desvio-padrão e valores mínimos e máximos para os indicadores da 1ª e 2ª aplicações do Zulliger (n=25).

Indicadores	Freq.	Freq.	Freq.	Freq.	Média	DP –	Média	DP –
	mínima 1ª aplic.	mínima 2ª aplic.	máxima 1ª aplic.	máxima 2ª aplic.	1ª Aplic.	1ª aplic.	2ª aplic.	2ª aplic.
R – resposta ^o	4	4	21	25	8,00	4,272	9,00	4,983
S – espaço branco ^o	0	0	9	10	2,36	2,177	2,20	2,398
W – global ^o	0	0	3	5	1,80	1,000	1,68	1,145
D – detalhe ^o	1	2	10	14	4,80	2,432	5,24	3,179
Dd – detalhe incomum ^o	0	0	10	12	1,44	2,382	2,08	2,827
M – movimento humano ¹	0	0	4	6	1,36	1,221	1,52	1,636
C – cor ¹	0	0	2	2	0,16	0,473	0,16	0,473
CF – cor e forma ¹	0	0	2	1	0,24	0,523	0,16	0,374
FC – forma e cor ¹	0	0	2	2	0,40	0,577	0,56	0,651
H – humano ²	0	0	4	4	1,04	1,098	1,20	0,957
HD – detalhe humano ²	0	0	2	4	0,48	0,770	0,52	0,918
(H) – Para-humano ²	0	0	2	2	0,36	0,569	0,44	0,583
(Hd) – detalhe para-hum. ²	0	0	2	4	0,32	0,557	0,48	0,918
Sum_SH ³								
Soma sombreado	0	0	7	8	1,84	1,625	2,24	2,047
EB – Tipo de vivência ³								
M:C+CF+FC	1	1	4	4	2,04	1,241	1,68	1,108
H:Hd+(H)+(Hd) ³								
Interesse pelas pessoas	1	1	4	4	1,96	1,098	2,20	1,080

Número de respostas e localização; ¹- Determinantes; ²- Conteúdos; ³- Proporções, frações e somas ponderadas.

Na Tabela 2, pode-se observar um único indicador – CF – com uma frequência máxima inferior na segunda aplicação se comparado à primeira aplicação, sendo que a maioria dos indicadores obteve uma frequência máxima superior na segunda aplicação. O indicador R, que corresponde ao número de respostas, mostra que houve um aumento de aproximadamente 20% na segunda aplicação. A Tabela 3 mostra a comparação das médias e o desvio-padrão dos indicadores da Tabela 2 com o grupo normativo de universitários (Villemor-Amaral et al., no prelo).

Tabela 3. Comparação da média e desvio-padrão da 1ª e 2ª aplicações com dados normativos de universitários.

Indicadores	1ª aplicação		2ª aplicação		Dados normativos	
	Média	DP	Média	DP	Média	DP
R – resposta ^º	8,00	4,272	9,00	4,983	7,23	2,712
S – espaço branco ^º	2,36	2,177	2,20	2,398	1,00	1,061
W – global ^º	1,80	1,000	1,68	1,145	2,15	1,476
D – detalhe ^º	4,80	2,432	5,24	3,179	4,28	2,560
Dd – detalhe incomum ^º	1,44	2,382	2,08	2,827	0,59	0,848
M – movimento humano ¹	1,36	1,221	1,52	1,636	0,96	0,941
C – cor ¹	0,16	0,473	0,16	0,473	0,16	0,432
CF – cor e forma ¹	0,24	0,523	0,16	0,374	0,56	0,652
FC – forma e cor ¹	0,40	0,577	0,56	0,651	0,46	0,742
H – humano ²	1,04	1,098	1,20	0,957	1,11	0,987
HD – detalhe humano ²	0,48	0,770	0,52	0,918	0,32	0,566
(H) – Para-humano ²	0,36	0,569	0,44	0,583	0,27	0,500
(Hd) – detalhe para-hum. ²	0,32	0,557	0,48	0,918	0,12	0,331
Sum_SH ³						
soma sombreado	1,84	1,625	2,24	2,047	1,074	1,243
EB – Tipo de vivência ³						
M:C+CF+FC	2,04	1,241	1,68	1,108	1,260	1,055
H:Hd+(H)+(Hd) ³						
Interesse pelas pessoas	1,96	1,098	2,20	1,080	1,335	0,742

^º- Número de respostas e localização; ¹- Determinantes; ²- Conteúdos; ³- Proporções, frações e somas ponderadas.

Conforme os dados da Tabela 3, pode-se observar que o indicador S ficou um DP acima da média nas duas aplicações. Já o Dd ficou a um DP acima da média na primeira aplicação e aproximadamente dois DP acima da média na segunda aplicação.

Os indicadores W e CF ficaram abaixo da média do grupo normativo em menos de um DP. O indicador C ficou na média do grupo normativo, e os indicadores (Hd) e Sum_SH ficaram na média na primeira e na segunda aplicações. A Tabela 4 descreve os resultados da correlação de Pearson e os valores de p entre as médias dos indicadores da primeira e da segunda aplicações.

Tabela 4. Correlação de Pearson (*r*) e valores de *p* do teste e reteste.

Indicadores	(<i>r</i>)	<i>p</i>
R – respostas ^º	0,94	0,000
S – espaço em branco ^º	0,91	0,000
W – global ^º	0,52	0,007
D – detalhe ^º	0,82	0,000
Dd – detalhe incomum ^º	0,97	0,000
M – movimento humano ¹	0,78	0,000
C – cor ¹	1,00	-
CF – cor e forma ¹	0,44	0,030
FC – forma e cor ¹	0,38	0,063
H – humano ²	0,82	0,000
HD – detalhe humano ²	0,63	0,001
(H) – para-humano ²	0,76	0,000
(Hd) – detalhe para-humano ²	0,83	0,000
Sum_SH – soma sombreado ³	0,51	0,009
EB – Tipo de vivência ³		
M:C+CF+FC	0,28	0,171
H:Hd+(H)+(Hd) ³		
Interesse pelas pessoas	0,53	0,006

^º- Número de respostas e localização; ¹- Determinantes; ²- Conteúdos;

³- Proporções, frações e somas ponderadas.

Pode-se observar que os indicadores R, S, Dd, D, C, H e (Hd) obtiveram correlações muito altas, acima de 0,80 (Sisto, 2007). Para Anastasi e Urbina (2000), é desejável que uma correlação atinja coeficientes de precisão aproximados de 0,80 ou 0,90. Já os indicadores M, (H) e Hd alcançaram correlação entre 0,60-0,80, o que indica uma associação alta. Uma correlação moderada, entre 0,40-0,60, foi obtida por H:Hd+(H)+(Hd), W, Sum_SH e CF. Por fim, a correlação do indicador FC foi baixa e sua significância marginalmente significativa foi de 0,063, e o indicador EB foi o único que não apresentou significância estatística.

Uma das possibilidades para um valor não significativo de EB é que ele é a resultante das somas de respostas M com a soma ponderada de cor cromática (M:WSumC). A soma das respostas com cor é obtida multiplicando-se cada tipo de resposta de cor cromática por um peso: $WSumC = (0,50) \times FC + (1,0) \times CF + (1,5) \times C$. Pode-se observar que a correlação de M foi alta, e C obteve uma correlação perfeita, porém CF teve uma correlação que, mesmo sendo significativa, foi moderada, e FC teve uma correlação baixa com $p = 0,063$, marginalmente significativa, o que indica que a dispersão entre os dados foi muito maior, e isso pode ter influenciado na correlação do indicador EB.

Villemor-Amaral e Lamounier (2006) estudaram a validade convergente do indicador EB no Zulliger e Rorschach, e concluíram que é necessário estabelecer normas e proporções específicas para o Zulliger, porque, quanto às respostas M, existe uma semelhança maior entre

os instrumentos, mas o mesmo não ocorre com as respostas de cor, devido ao número de pranchas coloridas do Rorschach e do Zulliger, sendo que, no primeiro, das dez pranchas, três são coloridas, e uma, a última da série, proporciona um estímulo muito maior de respostas coloridas. A prancha II do Zulliger é a mais colorida, não tão fragmentada, e, em consequência, propicia menos respostas.

Discussão

O sistema compreensivo possui um número considerável de variáveis a serem examinadas na tarefa proposta. Foram selecionadas algumas das variáveis consideradas fundamentais na proposta do método, que é avaliar a dinâmica da personalidade, com o objetivo de verificar, por meio do teste-reteste, a fidedignidade Zulliger - sistema compreensivo, por ser esse um instrumento e método de credibilidade clínica e foco de pesquisa em alguns países, inclusive no Brasil (Brinkmann, 1998; Ferreira & Villemor-Amaral, 2005; Franco & Villemor-Amaral, 2004; Machado & Villemor-Amaral, 2006; Villemor-Amaral, Farah, & Machado, 2007; Villemor-Amaral & Lamounier, 2006; Villemor-Amaral, Primi, & Cardoso, 2004; Zdunic, 1999).

Os resultados revelaram que, dos 16 indicadores analisados, a maioria obteve uma alta precisão com nível de significância de 0,01, o que atribui credibilidade ao método. No entanto, outros indicadores, presentemente não avaliados em razão de não estarem relacionados diretamente com os objetivos deste estudo, poderão ser analisados posteriormente em outras pesquisas.

O indicador EB precisa ser estudado mais criteriosamente, e as proporções para o Zulliger precisam ser apuradas, pois, de todos

os indicadores estudados, foi o único que não apresentou correlação, tal como já observado por Villemor-Amaral e Lamounier (2006) no estudo comparativo de validade convergente com o indicador EB no Zulliger e Rorschach. A hipótese levantada pelas pesquisadoras era de que as diferenças dos instrumentos poderiam favorecer o aparecimento de indicadores em proporções distintas das do Rorschach. Os resultados mostraram que as correlações de Pearson do indicador EB foram baixas, apontando assim, a necessidade de se estabelecer normas e proporções específicas para o Zulliger.

O presente estudo traz uma perspectiva positiva dos métodos projetivos, em especial ao Zulliger no sistema compreensivo, quanto aos critérios psicométricos exigidos pela comunidade científica, porém tem seus limites, uma vez que o tamanho da amostra e sua especificidade dificultam as generalizações desejáveis (Anastasi & Urbina, 2000; Pasquali, 1999). Cabe destacar que a especificidade da amostra proporciona certo controle de variáveis, e isso pode ter contribuído positivamente para este estudo, já que os resultados da média dos indicadores do grupo teste-reteste se encontraram dentro do intervalo médio do grupo normativo de universitários (Villemor-Amaral et al., no prelo), indicando que, apesar da especificidade do grupo voluntário, os resultados permitem certo grau de generalização. Entretanto, o estudo de fidedignidade do instrumento pode ser ampliado em outra ocasião, procurando-se aumentar tanto o número e a heterogeneidade da amostra para verificar se os resultados se mantêm quanto ampliando a quantidade de indicadores avaliados.

Quanto às dificuldades dos procedimentos científicos para as técnicas projetivas,

Villemor-Amaral (2006) fala sobre a natureza dos fenômenos estudados, pois nem sempre são observáveis diretamente pelo comportamento e são singulares, podendo não se repetir nem mesmo no indivíduo. Tais questões foram observadas no processo teste-reteste, pois, por ocasião do teste, alguns participantes expressaram acontecimentos vividos recentemente na elaboração das respostas, dado esse percebido em comentários muito pessoais durante a entrevista inicial, como um parente hospitalizado ou a gravidez da esposa, dentre outros. Já no reteste, depois de decorridos cinco meses, a elaboração da resposta mostrava-se diferente, alterando certos conteúdos e articulações dos mesmos, mas não necessariamente as características do estímulo percebido.

A singularidade do fenômeno e as vivências afetivas do período, somadas ao estímulo pouco definido, de fato proporcionam um número maior ou menor de respostas e a fragmentação ou integração destas. Alguns sujeitos, no primeiro momento (teste), deram um número de respostas em uma prancha, e, no segundo (reteste), integraram algumas delas, o que resultou em um número menor de respostas naquela prancha; também fragmentaram respostas que, na primeira aplicação, estavam integradas. Esse é um fator que pode ter contribuído para um aumento na frequência da maioria dos indicadores, principalmente do indicador R, que diz respeito ao número de respostas.

Além disso, o aumento do indicador R pode sugerir que, em um segundo momento, os indivíduos produzem mais, talvez por

estarem familiarizados com a tarefa e com a aplicadora. Nesse sentido, a padronização dos procedimentos, desde o ambiente de testagem até o *rapport*, bem como da aplicação e do inquérito do Zulliger, foram fundamentais para diminuir a variância de erro (Anastasi & Urbina, 2000).

Em uma análise rápida dos indicadores, pode-se destacar que o indicador R diz respeito ao número de respostas dadas pelos indivíduos, tal como já afirmado, e relaciona-se à capacidade de produzir idéias e à produtividade em si, sendo que, indiretamente, associa-se à energia e à boa motivação. Esse elemento ficou, neste estudo, acima da média do grupo normativo. Ao lado disso, respostas que incluam ou se localizam no espaço branco (S) indicam assertividade, busca de independência e auto-afirmação ou sentimento de hostilidade e oposicionismo, raiva e ressentimento contra pessoas ou eventos considerados injustos (Villemor-Amaral et al., no prelo). Nesse indicador, os participantes encontraram-se acima da média do grupo normativo (um DP). Convém ressaltar que R e S tiveram um índice de precisão acima de 0,90.

A localização W está associada ao esforço em processar os dados despertados pelos aspectos do estímulo, e revelou-se abaixo da média do grupo normativo. Por sua vez, as respostas D indicam que o aspecto mais comum, convencional e fácil do campo de estimulação chamou a atenção do sujeito, e Dd seria a atenção dirigida para os aspectos mais incomuns da experiência, com forte tendência a uma análise mais cuidadosa do estímulo, podendo estar associada ao

perfeccionismo e a tendências obsessivas. Os dois indicadores revelaram-se acima da média do grupo normativo, ao mesmo tempo em que alcançaram um índice alto de precisão.

As respostas com determinante M estão associadas aos recursos intelectuais, relacionando-se ao raciocínio abstrato, às fantasias, à criatividade, às formas superiores de conceituação e à tendência a usar o pensamento para a solução dos problemas (Villemor-Amaral et al., no prelo). M, além de estar acima da média do grupo normativo, também foi outro indicador preciso.

Os determinantes C, CF e FC estão relacionados com o desenvolvimento emocional, sendo que C indica menos esforço cognitivo, descarga irrestrita das emoções, CF seria a vazão dos estímulos afetivos com pouca modulação cognitiva e FC indica que as experiências afetivas foram controladas, direcionadas por elementos cognitivos, tal como apontado por Villemor-Amaral et al. (no prelo). Os indicadores FC e CF foram os aparentemente mais instáveis, sendo que C se manteve estável, com uma correlação positiva perfeita e na média. CF se manteve abaixo da média e com precisão significativa, porém FC ficou abaixo da média na 1ª aplicação e acima na 2ª, não alcançando um índice de precisão satisfatório (Conselho Federal de Psicologia [CFP], 2003).

Os conteúdos H, Hd, (H) e (Hd), componentes do índice de relação interpessoal, referem-se ao interesse pelas pessoas. H e Hd são indicadores de identificação clara e estável com os outros. De acordo com Villemor-Amaral et al. (no prelo), os indicadores (H) e

(Hd) mostram uma tendência à identificação com personagens com os quais não se interage realmente, o que pode sugerir um desconforto na identificação com pessoas reais e distanciamento do mundo real. Tanto (H) como (Hd) encontram-se acima da média e alcançaram alta precisão. Os indicadores estão relacionados com uma visão parcial, cautelosa, reservada e desconfiada dos outros. O indicador H, na primeira aplicação, ficou abaixo da média, e, na segunda, acima, quando comparado com o grupo normativo. Se sua frequência for menor que os demais indicadores da relação interpessoal (Hd, (H), (Hd)), isso pode revelar uma tendência ao isolamento.

Em síntese, os resultados corroboram as asserções de Güntert (2001), ao afirmar que os métodos projetivos permitem lidar com os dados de forma quantitativa e qualitativa. O conjunto de variáveis discutido pelo método lança luz sobre a dinâmica do indivíduo, no caso dos protocolos individuais, assim como nos do grupo, conforme os dados estatísticos.

Apesar de todas as dificuldades envolvidas no contexto dos procedimentos científicos para métodos projetivos, o presente estudo proporcionou um panorama positivo resultante dos esforços de outros pesquisadores, das discussões dos prováveis obstáculos e dos meios possíveis para se alcançar as qualidades imprescindíveis para os instrumentos de avaliação psicológica. Os resultados desta pesquisa, ainda que modestamente, puderam contribuir com dados sobre a fidedignidade de alguns indicadores do Zulliger-sistema compreensivo, proporcionando, dessa forma, mais credibilidade ao instrumento.

Anna Elisa de Villemor-Amaral*

Doutora em Ciências pela UNIFESP/EPM, docente do Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco, Itatiba, SP - Brasil; bolsista Produtividade em Pesquisa do CNPq.

Maria Aparecida dos Santos Machado

Graduada em Psicologia pela Universidade São Francisco, Itatiba, SP - Brasil.

Ana Paula Porto Noronha

Doutora em Psicologia: Ciência e Profissão pela Pontifícia Universidade Católica de Campinas; docente do Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco, Itatiba, SP - Brasil; bolsista Produtividade em Pesquisa do CNPq.

***Endereço para envio de correspondência:**

Rua Alexandre Rodrigues Barbosa, 45, Centro, Itatiba, SP - Brasil – CEP: 13251-900
E-mail: aevillemor@terra.com.br

Recebido 04/03/2008, Reformulado 29/10/2008, Aprovado 25/11/2008

Referências

- Alves, I. C. B. (2006). Considerações sobre a validade e precisão nas técnicas projetivas. In A. P. P. Noronha, A. A. A. Santos, & F. F. Sisto (Orgs.), *Facetas do fazer em avaliação psicológica* (pp. 173-176). São Paulo: Vetor.
- Alves-Mazzotti, A. J. (2001). O método nas ciências sociais. In A. J. Alves-Mazzotti, A. J., & F. Gewandszader (Orgs.), *O método nas ciências naturais e sociais* (p. 160). São Paulo: Pioneira.
- Anastasi, A., & Urbina, S. (2000). *Testagem psicológica*. Porto Alegre: Artes Médicas Sul.
- Anzieu, D. (1978). *Os métodos projetivos*. Rio de Janeiro: Campus.
- Brinkmann, H. (1998). Proposição de parâmetros para el test de Zulliger (Z). *Revista Chilena de Psicología*, 9(2).
- Conselho Federal de Psicologia. (2003). *Resolução nº 002/2003*. Recuperado em 13 de novembro de 2007, de <http://www.pol.org.br>
- Exner, J. E., Jr. (1994). *El Rorschach: un sistema comprehensivo* (3a ed., Vol. 1: fundamentos básicos, pp. 45-47). Madrid: Psimática.
- Ferreira, M. E. A., & Villemor-Amaral, A. E. (2005). O teste de Zulliger e avaliação de desempenho. *Paidéia: Caderno de Psicologia e Educação*, 15(32), 367-386.
- Franco, R. R. C., & Villemor-Amaral, A. E. (2004). Adaptação do sistema compreensivo de Exner para o Teste de Zulliger: estudo com pacientes esquizofrênicos [Resumo]. In *Anais do III Congresso Nacional da Sociedade Brasileira de Rorschach e Métodos Projetivos* (Vol. 3, p. 52).
- Güntert, A. E. V. A. (2001). Técnicas projetivas: o geral e o singular em avaliação psicológica. In F. F. Sisto, & E. T. B. Sbardelina, & R. Primi, R. (Orgs.), *Contextos e questões da avaliação psicológica* (pp. 77-84). São Paulo: Casa do Psicólogo.
- Machado, M. A. S., & Villemor-Amaral, A. E. (2006). Indicadores de depressão no Zulliger no sistema compreensivo [Resumo]. In *Anais do XII Encontro de Iniciação Científica* [CD-ROM] (p. 56).
- Nascimento, R. S. G. F., & Güntert, A. E. V. A. (2000). Novas tendências: introdução ao sistema compreensivo de Exner. In J. A. Cunha, *Psicodiagnóstico-V* (5a ed., pp. 368-377). Porto Alegre: ArtMed.
- Noronha, A. P. P., Sartori, F. A., Freitas, F. A., & Ottati, F. (2001). Informações contidas nos manuais de testes de inteligência publicados no Brasil. *Psicologia em Estudo*, 6(2), 101-106.
- Pasquali, L. (1999). Histórico dos instrumentos psicológicos. In L. Pasquali (Org.), *Instrumentos psicológicos: manual prático de elaboração* (pp. 13-21). Brasília, DF: LabPAM; IBAPP.
- Pasquali, L. (2001). Testes psicológicos: conceitos, história, tipos e usos. In L. Pasquali (Org.), *Técnica de exame psicológico-TEP: manual* (pp. 13-56). São Paulo: Casa do Psicólogo; Brasília, DF: Conselho Federal de Psicologia.
- Sisto, F. F. (2007). Delineamento correlacional. In M. N. Baptista (Org.), *Metodologias de pesquisas em ciências: análises quantitativas e qualitativas* (pp. 90-101). Rio de Janeiro: LTC.
- Vaz, C. E. (2000). A técnica de Zulliger no processo de avaliação da personalidade. In J. A. Cunha, *Psicodiagnóstico-V* (5a ed., pp. 386-387). Porto Alegre: ArtMed.
- Villemor-Amaral, A. E. (2006). Desafios para a cientificidade das técnicas projetivas. In A. P. P. Noronha, A. A. A. Santos, & F. F. Sisto (Orgs.), *Facetas do fazer em avaliação psicológica* (pp. 163-171). São Paulo: Vetor.
- Villemor-Amaral, A. E., Farah, F. H. Z., & Machado, M. A. (2007). *Zulliger – sistema compreensivo: indicadores de depressão*. Trabalho apresentado em III Congresso de Avaliação Psicológica e XII Conferência Internacional de Avaliação Psicológica: Formas e Contextos, João Pessoa.
- Villemor-Amaral, A. E., & Lamounier, R. (2006). *Evidence of convergent validity of the EB in the Rorschach and Zulliger*. Trabalho apresentado em VII ERA Congress, Pádua, Itália.
- Villemor-Amaral, A. E., Primi, R., & Cardoso, L. M. (2004). Adaptação do Teste de Zulliger para o sistema compreensivo numa amostra brasileira [Resumo]. In *Anais do III Congresso Nacional da Sociedade Brasileira de Rorschach e Métodos Projetivos* (Vol. 3, p. 121). Porto Alegre.
- Villemor-Amaral, A. E., Primi, R., Franco, R. R. C., Nuniz, M., Cardoso, L. M., & Machado, M. A. S. (no prelo). *O Teste de Zulliger no sistema compreensivo – forma individual*.
- Weiner, I. B. (1994). The Rorschach Inkblot Method (RIM) is not a test: Implications for theory and practice. *Journal of Personality Assessment*, 62(3), 494-504.
- Zdunic, A. L. (1999). *El Test de Zulliger en la evaluación de personal. Aportes del sistema comprehensivo de Exner*. Argentina: Paidós.