

Thematic Dossier

Effects of R-Optimized Administration on Form Quality in the Zulliger-SC Test

Anna Elisa de Villemor-Amaral¹ 

Cato Grønnerød² 

Gabriel Vitor Acioly Gomes³ 

Gisele Magarotto Machado¹ 

Juliana de Almeida Araújo⁴ 

Mayara Salgado de Moraes⁵ 

Scarlett Borges Fernandes⁶ 

Lucila Moraes Cardoso⁷ 

Ana Cristina Resende⁸ 

Silvana Alba Scortegagna⁹ 

Abstract: The Zulliger test, inspired by the Rorschach method, is widely used in Brazil to assess personality, cognition, and emotional aspects, particularly in contexts requiring quick evaluations. Recent adaptations, such as the R-Optimized administration based on the (Rorschach Performance Assessment System (R-PAS), aim to enhance its validity and psychometric properties. We examined whether controlling the number of responses (3–5 per card) in the Zulliger test leads to an increase in distorted form responses (FQ-). A total of 64 adults participated (M age = 32.9 years, $SD = 10.8$; 57.8 % women), with analyses conducted using repeated-measures ANOVA and paired t-tests. The result ($F(1, 63) = 3.60, p = .015, \eta^2 = .07$) showed no clinically significant increase in FQ-, indicating that the R-Optimized administration does not pathologize results. These findings reinforce the effectiveness of controlling the number of responses, highlight the need for updated normative data, and support adopting R-Optimized instructions which strengthen the reliability of the Zulliger test in psychological assessments.

Keywords: psychological assessment, validity, projective techniques, performance tests

Efeito da Aplicação R-Otimizada na Qualidade Formal no Zulliger-SC

Resumo: O teste Zulliger, inspirado no método Rorschach, é amplamente usado no Brasil para avaliar personalidade, cognição e aspectos emocionais, especialmente em contextos de avaliações breves. Recentes adaptações, como a administração R-Otimizada baseada no R-PAS, visam melhorar sua validade e propriedades psicométricas. Nosso estudo investigou se o aumento controlado de respostas (3-5 por cartão) no Zulliger gera aumento significativo de respostas com forma distorcida (FQ-). Sessenta e quatro adultos participaram (M idade = 32,9; $DP = 10,8$; 57,8% mulheres) e as análises foram conduzidas por meio de ANOVA de medidas repetidas e testes t pareados. O resultado ($F(1, 63) = 3,60; p = .015; \eta^2 = .07$) revelou ausência de aumento significativo em FQ-, indicando que a administração R-Otimizada não patologiza os resultados. Esses achados reforçam a eficácia do controle de respostas, destacam a necessidade de dados normativos atualizados e sustentam a adoção das instruções R-Otimizadas, fortalecendo a confiabilidade do Zulliger.

Palavras-chaves: avaliação psicológica, validade, técnicas projetivas, teste de desempenho

¹Universidade São Francisco, Campinas-SP, Brazil.

²University of Oslo, Oslo, Norway.

³Universidade Federal de Roraima, Boa Vista-RR, Brazil.

⁴Universidade Estácio de Sá, Resende-RJ, Brazil.

⁵Pontifícia Universidade Católica de Campinas, Campinas-SP, Brazil.

⁶Pontifícia Universidade Católica de São Paulo, São Paulo-SP, Brazil.

⁷Universidade Estadual do Ceará, Fortaleza-CE, Brazil.

⁸Pontifícia Universidade Católica de Goiás, Goiânia-GO, Brazil.

⁹Universidade de Passo Fundo, Passo Fundo-RS, Brazil.

Correspondence address: Anna Elisa de Villemor-Amaral. Universidade São Francisco. Rua Waldemar César da Silveira, 105, Campinas - SP, Brazil. CEP 13045-510. E-mail: anna.villemor@usf.edu.br e aevillemor@gmail.com

Efecto de la Aplicación R-Optimizada en la Calidad Formal del Zulliger-SC

Resumen: El test Zulliger, inspirado en el método de Rorschach, se utiliza ampliamente en Brasil para evaluar personalidad, cognición y aspectos emocionales, especialmente en contextos que requieren evaluaciones rápidas. Adaptaciones recientes, como la administración R-Optimizada basada en el R-PAS, buscan mejorar su validez y propiedades psicométricas. Nuestro estudio investigó si el control del número de respuestas (3–5 por lámina) en el Zulliger conduce un aumento significativo de respuestas con forma distorsionada (FQ-). Participaron sesenta y cuatro adultos (M edad = 32.9 años, $DE = 10.8$; 57.8 % mujeres), con análisis realizados mediante ANOVA de medidas repetidas y pruebas t pareadas. El resultado ($F(1, 63) = 3.60$, $p = .015$, $\eta^2 = .07$) no mostró un aumento clínicamente significativo en FQ-, lo que indica que la aplicación R-Optimizada no patologiza los resultados. Estos hallazgos refuerzan la eficacia del control del número de respuestas, destacan la necesidad de datos normativos actualizados y respaldan la adopción de instrucciones R-Optimizadas, fortaleciendo la fiabilidad del Zulliger en evaluaciones psicológicas.

Palabras clave: evaluación psicológica, validez, técnicas proyectivas, test de desempeño

Introduction

The Zulliger test is one of the most researched methods in the Brazilian context for assessing personality. The primary aim of the test is to assess the structure and dynamics of an individual's personality, while it also aims at assessing cognitive and emotional aspects of the individual (Zulliger, 1969; Vergati et al. 2019). One of the main advantages of this instrument is its reduced time for application and analysis of responses. This makes it particularly useful for assessments which need to be conducted quickly, such as in situations with a large volume of people assessed, or when applied to children and adolescents (Fazendeiro & Novo, 2012; Caporale et al., 2023).

The Zulliger test was developed in 1948 by Hans Zulliger, inspired by the Rorschach method, and was initially created to select officers for the Swiss armed forces. The test consists of three cards, each presenting different symmetrical ink blots similar to Rorschach inkblots. The evaluator presents each card individually during administration of the test, and then asks the examinee to describe what it might be. After the administration, the responses provided are then coded and interpreted based on the interpretation system used (Caporale et al., 2023).

Several administration and coding systems have been proposed since the Zulliger test was created to adapt it to different contexts of personality assessment. These adaptations have reflected a shift from the original Swiss approach to more contemporary models developed to meet evolving assessment needs over time (Grazziotin & Scortegagna, 2024). These systems evolved from the original Swiss method to contemporary approaches which integrate R-Optimized procedures and performance-based logic.

In addition, the Z-SC is currently approved for professional use by the Brazilian Psychological Test Assessment System (SATEPSI), coordinated by the Federal Council of Psychology. This endorsement ensures that the test meets national standards of validity, reliability, and ethical application.

Several studies have been conducted in recent years using a new version for administering the Zulliger instrument, named R-Optimized administration (Gonçalves et al., 2021; Hammarström & Grønnerød, 2023; Seitl et al., 2018). This administration is based on the Rorschach

Performance Assessment System (R-PAS), as described by Meyer et al. (2017), which originated from research conducted within the Comprehensive System. It aims to standardize usage of the instrument internationally and to enhance the psychometric rigor of its interpretations. One of the key differences of this system compared to other administration, coding, and interpretation methods of the Rorschach method is the limitation on the number of responses. In R-PAS, the respondent is asked to provide two or three responses per card, with a maximum of four. This is intended to reduce variability in the number of responses and consequently its impact on the psychometric qualities of the instrument (Meyer et al., 2017). The R-PAS proposes an ideal range of responses to optimize statistical analyses, while ensuring the quality of the protocols and the validity of data interpretations. The R-PAS authors justified this change in instructions to improve the accuracy of assessments and to ensure that analyses are based on a sufficient but also a maximum number of responses, without compromising the integrity of the results. The intention was to reduce distortions in interpretation of other variables and to enhance the stability of the indicators (Meyer et al., 2017).

Historically, adaptations of the systems used in the Rorschach method have also been made for the Zulliger due to the similarities between the two instruments (Hammarström & Grønnerød, 2023; Seitl et al., 2018). Additionally, the results presented by Villemor-Amaral et al. (2016), which compared the coding of responses from both tests applied to the same person, showed that the most significant and consistent correlations occurred in protocols with more than nine responses on the Zulliger. It was identified that very short protocols provide a limited sample of an individual's performance on the test, weakening the analysis and conclusions about the examinee's personality, as a reduced number of responses affects other variables of the instrument and compromises the validity of the results (Cárpio & Cubas Lugón, 2011; Grazziotin et al., 2023; Seitl et al., 2018; Villemor-Amaral et al., 2016).

Considering these findings and the fact that the Zulliger test contains only three cards, it was decided to increase the number of responses proposed in the R-PAS by adopting the procedure of requesting three to five responses per card.

However, this raised two important issues: first, the need to establish new normative data with this new procedure, and second, the importance of verifying the validity of this new administration procedure. Thus, several studies have been conducted in recent years to achieve these objectives, such as the studies performed by Gonçalves et al. (2021), Gonçalves and Villemor-Amaral (2020), and Graziotin et al. (2023).

Gonçalves et al. (2021) sought validity evidence for the R-Optimized administration of the Zulliger test in a sample of 41 participants. Administration of the Zulliger and Rorschach tests were alternated between participants to minimize the influence of one test on the other. As a result, 25% of all the protocols administered were recoded by independent raters, with inter-rater reliability considered adequate to proceed with the analyses. The correlation between the coding of the two instruments showed satisfactory results, with some codes demonstrating improved associations compared to the existing literature, providing validity evidence for the R-Optimized application of the Zulliger test.

In addition, Gonçalves and Villemor-Amaral (2020) conducted a study aimed at comparing the frequency of depression-related codes in the Zulliger test using the R-Optimized administration between a group of individuals with depression and a control group without diagnoses. The sample consisted of 86 participants, with 43 diagnosed with depression (88.0% female, with a mean age of 35.8 years) and 43 without diagnoses (84.4% female, with a mean age of 35.7 years). The results revealed statistically significant differences between the groups in variables such as mixed determinants, total achromatic color responses (C'), pure color responses (C), and special codes for morbid (MOR) and aggressive (AG) content. The authors concluded that the R-Optimized administration method of the Zulliger test provided deeper understanding on the functioning of individuals diagnosed with depression.

Finally, the study conducted by Graziotin and Scortegana (2023), demonstrate the stability of the Zulliger-SC when using the R-optimized administration. However, the effects of requiring a higher number of responses in the test needed further scrutiny. Asking at least three responses per card would increase the minimum number of total responses in the protocol from three to nine, and the highest number of allowed responses to five would increase the maximum protocol number from 12 to 15.

Based on the results of the discussed studies, the question arose as to what extent requiring three and allowing five responses could increase distortion in the last responses given to each card, considering that this number exceeds the average number of responses typically given in the conventional application of the Zulliger. The concern raised was: "could encouraging a higher number of responses pathologize the results, leading to an increase in distorted form quality (FQ-)?" The FQ- indicator reflects that the examinee's response shows an overly subjective and generally distorted perspective of reality (Ghirardello et al., 2020; Pignolo et al., 2021) which means higher pathology levels. Motivated by these issues, we aimed at investigating whether increasing the number of

responses in the Zulliger test leads to an increase in distorted form responses.

The concern that increasing the number of responses per card could lead to artificial inflation of perceptual distortions (e.g., an increase in distorted form quality, or FQ-) has also been raised in the context of the Rorschach Performance Assessment System (R-PAS) (Kleiger & Mihura, 2021). While the current study does not apply the R-PAS system, it draws on the rationale of encouraging three to five responses per card. Prior findings suggested that such procedural changes do not necessarily lead to pathological results, for example in the study by Viglione et al. (2015), but this had to be verified in the Zulliger test.

Method

Participants

A total of 64 volunteers over the age of 18 participated in this study, recruited in person by researchers from the Psychological Assessment in Mental Health Laboratory (LAPSaM I) at the University of São Francisco, Brazil. Of these participants, 37 were women (57.8%) and 27 were men (42.2%), with an average age of 32.9 years ($SD = 10.8$ years). Of the total participants, 45.3% ($n = 29$) had completed high school, and 54.7% ($n = 35$) had attended higher education. The exclusion criteria comprised individuals who were undergoing psychological or psychiatric treatment at the time of the study.

Measures

A form developed specifically for this study was applied containing questions about the participants' sociodemographic information, such as age, gender, marital status, educational level, profession, and known psychiatric diagnosis.

The Zulliger test is used to assess aspects of personality, similar to the Rorschach test. It consists of three cards with inkblots, which are presented to the respondent, who must describe what they might look like. The test was administered and coded in this study according to the Comprehensive System, with the initial instructions modified to control the number of responses. Participants were asked to provide between three and five responses per card, with a maximum of six, whereas the standard R-PAS instruction is to give two to three responses, with a maximum of four.

When participants provided fewer than three responses, the administrator prompted "Can you see anything else?", in order to administrate contingencies; if a sixth response was given, the examiner accepted it, replied "That's enough, thank you" and immediately turned the card to end that phase.

Procedures

Data Collection. Participants were invited to voluntarily join the research by the team of researchers from the



Psychological Assessment in Mental Health Laboratory (LAPSaM I), through personal contacts and using the “snowball” strategy, in which one participant could refer another potential participant.

The data collection was conducted individually according to the availability of both the participant and the researcher, and took place at the university’s facilities, which had appropriate rooms for psychological testing. Each test administration lasted approximately 40 minutes. The Zulliger test responses were coded by the researcher immediately after the administration (or as soon as possible), following the guidelines established in the Zulliger Comprehensive System manual. After the data collection was fully completed by all administrators, 25% of the Zulliger test protocols were recorded by other independent judges to conduct coding agreement analyses. The inter-rater agreement for FQ- was $\kappa = 0.56$ (95 % $CI = 0.41-0.71$), classified moderate as per Landis & Koch (1977).

Data analysis. We performed the coding of the Zulliger Test using the Chesss 1.52 Paris program following the Comprehensive System (SC). We used the JASP software (JASP Team, 2023) to conduct descriptive statistical analyses to characterize the sample, as well as repeated measures ANOVA analyses focusing on the number of FQ- occurrences in each response (e.g., first, second, third) given by a participant to each card.

We conducted inferential statistical analyses to determine whether the R-Optimized administration would increase distortion of the percepts seen in the cards. This distortion would be identified by an increase in the coding of form quality minus (FQ-). Therefore, we conducted a repeated measures ANOVA comparing responses 1, 2, 3, and 4 across all cards simultaneously to assess whether there were changes in FQ- across the increase in responses given, as well as responses 1, 2, 3, and 4 on each card separately. We selected the participants who gave at least four responses on one card for the first analysis, totaling 51 participants.

Ethical considerations

The project was approved by the Ethics Committee at a university in Brazil (CAAE: 04269818.7.0000.5514; Approval: 3.083.562). Participation was voluntary, and we did not consider participation as involving risk or harm in any

way. The participants were not given feedback or any other reward for their participation.

Results

The number of responses provided by participants for each of the three cards varied between three and six responses. It was noted that most participants provided between three (26 in card I, 28 in Card II and 36 in card III) and four (25 in card I, 24 in Card II and 16 in card III) responses per card. A smaller number of people gave five (12 in card I, 10 in Card II and 9 in card III) and six responses (1 in card I, 2 in Card II and 3 in card III). Card I had an average of 3.81 responses ($SD = 0.79$), Card II had an average of 3.78 responses ($SD = 0.826$), and Card III had an average of 3.67 responses ($SD = 0.89$). All cards had an average of 3.75 responses ($SD = 0.84$).

The analysis conducted with all cards simultaneously revealed a statistically significant difference in the amount of FQ- across responses 1, 2, 3, and 4 ($F = 3.598$; $p = .015$; $\eta_p^2 = .070$). We performed a Bonferroni post-hoc test to identify the source of the difference, which indicated a statistically significant difference in the amount of FQ- only between the 1st and 3rd responses ($t = -3.373$; $d = -0.59$; $p_{\text{bonf}} = .018$), with a higher mean in the 3rd response ($M = 0.76$, $SD = 0.68$, $d = |0.59|$). While not significant, it is important to note that a moderate effect size was also observed in the amount of FQ- between the 2nd and 3rd responses ($t = -2.541$, $p_{\text{bonf}} = .072$, $d = |-0.50|$), also with a higher mean for the 3rd response. Further details can be seen in Table 1.

Following this initial analysis, we examined the possibility of differences in the level of FQ- for each card separately. We first analyzed card I. The criterion of selecting participants who provided four responses on card I was maintained ($n = 38$). We observed a statistically significant difference in the amount of FQ- across responses 1, 2, 3, and 4 ($F = 5.456$; $p = .002$; $\eta_p^2 = .129$).

Next, a Bonferroni post-hoc test was conducted to determine the source of the differences in responses on Card I. A statistically significant difference in the amount of FQ- was found between the 1st and 3rd responses ($t = -3.071$; $d = -0.694$; $p_{\text{bonf}} = .016$), and between the 2nd and 3rd responses ($t = -3.350$; $d = -0.757$; $p_{\text{bonf}} = .007$). Additionally, two comparisons showed a considerable effect size, although they were not statistically

Table 1.
Bonferroni Post-Hoc Test for ANOVA across all card responses simultaneously

		Mean Difference	Standard Error	t	p_{bonf}	Cohen's d
	FQ- 2nd response	-0.059	0.123	-0.476	1.000	-0.094
FQ- 1st response	FQ- 3rd response	-0.373	0.123	-3.018	0.018	-0.595
	FQ- 4th response	-0.196	0.123	-1.588	0.686	-0.313
FQ- 2nd response	FQ- 3rd response	-0.314	0.123	-2.541	0.072	-0.501
	FQ- 4th response	-0.137	0.123	-1.112	1.000	-0.219
FQ- 3rd response	FQ- 4th response	0.176	0.123	1.429	0.930	0.282

significant. These results were found in the comparison between FQ- in the 1st and 4th responses ($d = -0.505$; $p_{\text{bonf}} = .165$), and between the 2nd and 4th responses ($d = -0.568$; $p_{\text{bonf}} = .080$). The 4th response had a higher mean FQ- ($M = 0.34$) in both comparisons. These results are detailed in Table 2.

Next, we conducted an analysis for the responses on Card II, selecting 36 protocols from participants who provided at least four responses. Although the mean scores for each response differed according to this analysis (except between the second and third responses), no statistically significant differences were found in the amount of FQ- across responses 1, 2, 3, and 4 ($F = 1.444$; $p = .234$; $\eta_p^2 = .040$).

We conducted an analysis to determine whether there were statistically significant differences in the amount of FQ- across responses 1, 2, 3, and 4 on Card III. A total of 28 cases from participants who provided at least four responses were selected. No statistically significant differences were found ($F = 0.849$; $p = .471$; $\eta_p^2 = .030$).

In the final step, we compared the frequency of FQ- in the last response a person gave (this last response could be the third, fourth, or fifth) with the frequency of FQ- in the penultimate response (second, third, or fourth). The participants were subsequently grouped according to the number of responses provided for each card (three, four, or five), and the frequency of FQ- of the penultimate response was compared with that of the final response for each group. We performed paired sample t-tests to conduct these comparisons.

The results of the paired sample t-test comparing the FQ- frequency between the penultimate and last responses across

cards are presented in Table 3. Table 4 presents the results of the comparisons made for each card and each group according to the number of responses provided by the participants. Most of the results were non-significant and showed small effect sizes (except for individuals who gave five responses on Card I), suggesting no meaningful differences in the frequency of FQ- between the last response and the penultimate response.

Discussion

Despite the evidence that administering the Zulliger test with a controlled number of responses optimizes interpretations, there was concern about whether this administration would increase perceptual distortion, specifically distorted form quality (FQ-), in the later responses for each card. Our results showed that the increase in FQ- did not occur in the fourth response of the cards, indicating that requesting three to five responses per card did not lead to an increase in FQ- frequency in subsequent responses. Comparisons of the last and penultimate responses also confirmed that there was no significant increase in FQ- responses.

These findings are consistent with those reported by Viglione et al. (2015), who compared the standard Comprehensive-System administration with an R-Optimized administration (2–3 responses per card, maximum 4). The optimized method reduced overly short and overly long protocols in a psychiatric and mixed-clinical sample and did not increase distorted form quality responses (FQ-; Kleiger

Table 2.
Bonferroni Post-Hoc Test for ANOVA for Card I responses

		Mean Difference	Standard Error	<i>t</i>	p_{bonf}	<i>Cohen's d</i>
FQ- 1st response	FQ- 2nd response	0.026	0.094	0.279	1.000	0.063
	FQ- 3rd response	-0.289	0.094	-3.071	.016	-0.694
	FQ- 4th response	-0.211	0.094	-2.234	.165	-0.505
FQ- 2nd response	FQ- 3rd response	-0.316	0.094	-3.350	.007	-0.757
	FQ- 4th response	-0.237	0.094	-2.513	.080	-0.568
FQ- 3rd response	FQ- 4th response	0.079	0.094	0.838	1.000	0.189

Table 3.
Paired sample t-test results comparing FQ- frequency between penultimate and last responses across cards

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i> (<i>df</i> =63)	<i>p</i>	<i>d</i>
Penultimate Response - Card 1	64	0.313	0.467	1.093	.279	0.137
Last Response - Card 1	64	0.234	0.427			
Penultimate Response - Card 2	64	0.297	0.460	-0.728	.470	-0.091
Last Response - Card 2	64	0.359	0.484			
Penultimate Response - Card 3	64	0.188	0.393	-1.426	.159	-0.178
Last Response - Card 3	64	0.281	0.453			
Penultimate Response - All cards summed	64	0.797	0.694	-0.659	.512	-0.082
Last Response - All cards summed	64	0.875	0.745			

Table 4.*Paired sample t-test results comparing FQ- frequency between the penultimate and last responses on Card I, II, and III*

Card	Number of responses	Response	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t (df)</i>	<i>p</i>	<i>d</i>
Card I	People who gave three responses	Penultimate Response	26	0.12	0.326	-0.811 (25)	.425	-0.159
		Last Response	26	0.19	0.402			
	People who gave four responses	Penultimate Response	25	0.48	0.510	1.281 (24)	.212	0.256
		Last Response	25	0.32	0.476			
	People who gave five responses	Penultimate Response	12	0.42	0.515	2.345 (11)	.039	0.677
		Last Response	12	0.08	0.289			
Card II	People who gave three responses	Penultimate Response	28	0.32	0.476	0.000 (27)	1.000	0.000
		Last Response	28	0.32	0.476			
	People who gave four responses	Penultimate Response	24	0.21	0.415	-0.700 (23)	.491	-0.143
		Last Response	24	0.29	0.464			
	People who gave five responses	Penultimate Response	10	0.40	0.516	-0.361 (9)	.726	-0.114
		Last Response	10	0.50	0.527			
Card III	People who gave three responses	Penultimate Response	36	0.19	0.401	-0.627 (35)	.535	-0.105
		Last Response	36	0.25	0.439			
	People who gave four responses	Penultimate Response	16	0.25	0.447	-0.436 (15)	.669	-0.109
		Last Response	16	0.31	0.479			
	People who gave five responses	Penultimate Response	9	0.11	0.333	-1.512 (8)	.169	-0.504
		Last Response	9	0.33	0.500			

& Mihura, 2021). Although our study applied the Zulliger rather than the Rorschach, the convergent result is the same: encouraging additional responses within a reasonable limit enriches data without inflating perceptual distortion.

Consistent with previous evidence on the Zulliger's psychometric properties, our findings are conceptually aligned with the validity patterns described by Gonçalves et al. (2021), although their study did not compare R-Optimized versus non-optimized protocols. Those authors provided validity evidence under an R-Optimized administration and highlighted the importance of controlling response quantity to avoid compromising data interpretation. Nevertheless, they relied on a relatively small sample, and larger studies are needed to determine whether psychiatric populations (who typically show higher mean FQ- scores) respond differently under the same procedure.

The findings of the present study complement and extend the existing validity evidence for the Zulliger test, particularly regarding its psychometric robustness. While Grazziotin et al. (2023) demonstrated reasonable to excellent temporal stability for most Zulliger-CS variables, even across long intervals between administrations, our study addressed a complementary aspect of validity by examining the impact of standardizing the number of responses on the occurrence of form distortion (FQ-). Our results specifically indicate that requesting three to five responses per card, as proposed in the R-Optimized administration, does not significantly increase distorted form responses. Thus, while Grazziotin et al. (2023) reinforce the test's reliability over time, our findings suggest that standardizing

administration procedures maintains the integrity of clinical indicators, thereby strengthening the Zulliger's applicability in evidence-based psychological assessments.

These outcomes align with the work of Seidl et al. (2018), who found that the R-Optimized instruction effectively standardizes the number of responses compared to traditional and Comprehensive System instructions. This standardization reduces variability across protocols and prevents the number of responses from confounding test results. Our study consistently showed that requiring three to five responses per card does not lead to an artificial increase in distorted form responses, indicating that the R-Optimized administration is not only methodologically sound, but also clinically safe. Whereas Seidl et al. (2018) emphasized the procedural benefits of standardization, our findings contribute additional evidence by demonstrating that the R-Optimized method preserves the clinical validity of the test.

Furthermore, our results reinforce and expand upon the evidence presented by Gonçalves et al. (2021), who reported that the R-Optimized administration significantly improves the formal quality of responses in the Zulliger test. Their findings revealed greater consistency and organization in collected data when employing this methodology. By specifically analyzing the impact of R-optimization on Form Quality within the Zulliger-SC, our study confirms that standardizing and optimizing administration reduces distorted or inconsistent responses, enhancing the instrument's reliability. Taken together, these aligned findings substantiate the incremental validity of the R-Optimized approach and

highlight its potential to increase both rigor and precision in projective psychological assessment.

Finally, the R-Optimized procedure minimizes the risk of excessively short protocols which require re-administration by encouraging a moderate and standardized number of responses (three to five per card), thereby improving clinical efficiency. Importantly, no increase in perceptual distortion (FQ-) was observed, confirming that the procedure is non-pathologizing and preserves interpretive validity. This standardization is intended to promote greater consistency across protocols and to control for potential effects of response quantity on test results. Based on these findings and prior studies, the authors opted to retain the recommendation of a minimum and maximum number of responses per card in the Zulliger-SC. As a result, the next edition of the Zulliger-SC manual in Brazil will include updated normative tables based on a sample tested with R-Optimized instructions. Future studies with larger clinical and non-clinical samples should examine whether R-Optimization also benefits other variables (e.g., special codes).

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

- Caporale, R., Allone, C., Maccarrone, A., & Craparo, G. (2023). The clinical use of the Zulliger Test in the assessment of personality functioning and psychopathology. *Mediterranean Journal of Clinical Psychology, 11*(1), 1-13. <https://doi.org/10.13129/2282-1619/mjcp-3627>
- Cárpio, S. R., & Cubas Lugón, M. (2011). Validación del sistema comprensivo de Exner en el Test de Zulliger [Validation of the Exner comprehensive system in the Zulliger Test]. *Persona, 14*, 145-158. <https://doi.org/10.26439/persona2011.n014.256>
- Fazendeiro, H. M. M., & Novo, R. J. F. (2012). Versão coletiva do Teste de Zulliger segundo o Sistema Compreensivo de Rorschach [Group version of the Zulliger Test according to the Rorschach Comprehensive System]. *Avaliação Psicológica, 11*(3), 407-422. <https://pepsic.bvsalud.org/pdf/avp/v11n3/v11n3a08.pdf>
- Ghirardello, D., Ales, F., Raimondi, V., Viglione, D. J., Zennaro, A., & Giromini, L. (2020). The impact of extrapolation and examiner's judgment on Rorschach form quality coding: Interrater reliability and clinical implications. *Bolettino di Psicologia Applicata, 68*(289), 10-24. <https://doi.org/10.26387/bpa.289.2>
- Gonçalves, A. P., & Villemor-Amaral, A. E. (2020). Zulliger R-optimized in the evaluation of depressive characteristics. *Paidéia (Ribeirão Preto), 30*, e3023. <https://doi.org/10.1590/1982-4327e3023>
- Gonçalves, A. P., Zuanazzi, A. C., & Villemor-Amaral, A. E. (2021). Evidências de validade para Zulliger utilizando a aplicação R-otimizada [Validity evidence for Zulliger using the R-optimized administration]. *Arquivos Brasileiros de Psicologia, 73*(2), 144-155. https://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1809-52672021000200011
- Grazziotin, J. B. D. D., & Scortegagna, S. A. (2024). Propriedades psicométricas do Teste de Zulliger: Revisão sistemática (2009-2022) [Psychometric properties of the Zulliger Test: Systematic review (2009-2022)]. *Avaliação Psicológica, 23*(1), 76-85. https://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712024000100076&lng=pt&nrm=iso
- Grazziotin, J. B. D. D., Tognon, A. P., & Scortegagna, S. A. (2023). Temporal stability of the Zulliger Test in Brazilian adults. *Paidéia (Ribeirão Preto), 33*, e3304. <https://doi.org/10.1590/1982-4327e3304>
- Hammarström, E., & Grønnerød, C. (2023). Pre- and postpsychotherapy assessment of a patient with retrograde amnesia: A Zulliger study. *Rorschachiana, 44*(1), 59-83. <https://doi.org/10.1027/1192-5604/a000167>
- Kleiger, J. H., & Mihura, J. L. (2021). Developments in the Rorschach assessment of disordered thinking and communication. *Rorschachiana, 42*(2), 265-280. <https://doi.org/10.1027/1192-5604/a000132>
- Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics, 33*(2), 363-374. <https://doi.org/10.2307/2529786>
- Meyer, G. J., Viglione, D. J., Mihura, J. L., Erard, R. E., & Erdberg, P. (2017). *Rorschach sistema de avaliação por desempenho: Manual de aplicação codificação e interpretação e manual técnico* [Rorschach performance assessment system: Administration, coding, interpretation manual and technical manual] (D. R. Silva & F. K. Miguel, Trans.). Hogrefe.
- Pignolo, C., Viglione, D. J., & Giromini, L. (2021). How reliably can examiners make form quality (FQ) judgments in the absence of the form quality (FQ) tables? *Rorschachiana, 42*(1), 21-34. <https://doi.org/10.1027/1192-5604/a000135>
- Seitl, M., Dočkalová, J., Dostál, D., Kolařík, M., Hasoňová, V., Palová, K., & Fryšťacká, K. (2018). Standardizace počtu odpovědí v Zulligerově Testu: První krok při vyvoji nového evidence-based systému [Standardization of number of responses in the Zulliger Test: First step in the development of a new coding system]. *E-Psychologie, 12*(4), 1-15. https://www.e-psycholog.eu/pdf/seitl_et al1.pdf
- Vergati, C., Annunziata, S., Baldo, G., Longobardi, T., Ricciardi, O., & Verniti, S. (2019). L'utilità dello Zulliger Test nella valutazione del funzionamento psichico [The usefulness of the Zulliger Test in the assessment of psychic functioning]. *Phenomena Journal: Giornale Internazionale di Psicopatologia, Neuroscienze e Psicoterapia, 1*(1), 68-73. <https://doi.org/10.32069/pj.2019.1.52>

Villemor-Amaral, A. E., Pianowski, G., & Carvalho, L. F. (2016). Issues about color, human movement and number of responses in the Zulliger Test. *Rorschachiana*, 37(2), 95-113. <https://doi.org/10.1027/1192-5604/a000068>

Viglione, D. J., Meyer, G. J., Jordan, R. J., Converse, G. L., Evans, J., MacDermott, D., & Moore, R. (2015). Developing an alternative Rorschach administration method to optimize the number of responses and enhance clinical inferences. *Clinical Psychology & Psychotherapy*, 22(6), 546-558. <https://doi.org/10.1002/cpp.1913>

Zulliger, H. (1969). *The Zulliger individual and group test*. International Universities Press.

Anna Elisa de Villemor-Amaral is a Ph.D. and Associated Professor of the Universidade São Francisco, Campinas-SP, Brazil.

Cato Grønnerød is a Ph.D. and Professor of the University of Oslo, Oslo, Norway.

Gabriel Vitor Acioly Gomes is a Ph.D. and Professor of the Universidade Federal de Roraima, Boa Vista-RR, Brazil.

Gisele Magarotto Machado is a Ph.D. and Professor of the Universidade São Francisco, Campinas-SP, Brazil.

Juliana de Almeida Araújo is a Ph.D. and Professor of the Universidade Estácio de Sá, Resende-RJ, Brazil.

Mayara Salgado de Moraes is a Ph.D. and Professor of the Pontifícia Universidade Católica de Campinas, Campinas-SP, Brazil.

Scarlett Borges Fernandes is a Ph.D. and Professor of the Pontifícia Universidade Católica de São Paulo, São Paulo-SP, Brazil.

Lucila Moraes Cardoso is a Ph.D. and Professor of the Universidade Estadual do Ceará, Fortaleza-CE, Brazil.

Ana Cristina Resende is a Ph.D. and Professor of the Pontifícia Universidade Católica de Goiás, Goiânia-GO, Brazil.

Silvana Alba Scortegagna is a Ph.D. and Professor of the Universidade de Passo Fundo, Passo Fundo-RS, Brazil.

Authors' Contribution:

All authors made substantial contributions to the study conception and design, to data analysis and interpretation, and to the manuscript revision and approval of the final version. All the authors assume public responsibility for the content of the manuscript.

Associate editor:

Samuel Lincoln Bezerra Lins

Received: May. 1, 2025

1st Revision: Jul. 26, 2025

Approved: Sep. 07, 2025

How to cite this article:

Villemor-Amaral, A. E., Grønnerød, C., Gomes, G. V. A., Machado, G. M., Araújo, J. A., Moraes, M. S., Fernandes, S. B., Cardoso, L. M., Resende, A. C., & Scortegagna, S. A. (2025). Effects of R-Optimized Administration on Form Quality in the Zulliger-SC Test. *Paidéia (Ribeirão Preto)*, 35, e3534. <https://doi.org/10.1590/1982-4327e3534>