**Psychometrics of Art –
Validation of *RizbA*,
a quantitative Rating Instrument for pictorial Expression**

Kerstin Schoch^{1,2}, Thomas Ostermann²

¹ The Pop-up Institute, HKS – University of Applied Sciences and Arts Ottersberg, Germany,

² Department of Psychology and Psychotherapy, Faculty of Health, Witten/ Herdecke University, Germany

Kerstin Schoch

The Pop-up Institute

HKS – University of Applied Sciences and Arts Ottersberg

Am Wiestebruch 68, 28870 Ottersberg, Germany

kerstin.schoch@hks-ottersberg.de

(corresponding author)

Thomas Ostermann

Department of Psychology and Psychotherapy

Faculty of Health, Witten/ Herdecke University

Alfred-Herrhausen-Straße 50, 58448 Witten, Germany

thomas.ostermann@uni-wh.de

Acknowledgements

The authors gratefully acknowledge Thomas Gengenbach for providing theoretical and practical support regarding IT, the implementation of the online survey and data structuring.

Thanks also to Lily Martin and Rebecca Kahn for their valuable feedback on the manuscript.

Ethics approval

The authors received a positive vote of the Ethics Committee for Creative Arts Therapies at the Nürtingen-Geislingen University, Germany (Reference number 19C27J101W).

Abstract

Although art has been subject to psychological research for some time, the artwork itself received little attention in quantitative research. The rating instrument for two-dimensional pictorial works (*RizbA*) fills this gap by providing a tool for formal picture analysis. This study validates the questionnaire on 294 images created by 147 non-artists. In an online test-retest study the material was rated by 880 (T1) and 475 (T2) experts in visual art using *RizbA*. Statistical quality criteria, Principal component analysis (PCA) and indices of factor similarity were computed. The overall test's capacity of differentiation yields a partial eta-squared of .28 (T1) and .33 (T2). Test-retest reliability is .93. PCA reveals an eight-factors solution. Tucker's coefficients of congruence range between |0.82| and |1.00|. Intraclass correlation coefficients are .81 (T1) and .84 (T2). Results indicate generalizability to amateurs' works. As the first reliable tool for picture analysis, *RizbA* allows a more detailed examination of art and its correlates.

Keywords: visual art, formal picture analysis, pictorial expression, psychometrics, questionnaire

Introduction

Analyzing Art

Psychology of art, including concepts such as creativity and aesthetics, has been subject to psychological research for some time (Commare et al., 2018). Still, visual arts, art-making and the artwork itself have barely gotten attention in quantitative research. One of the reasons for this might be the different scientific disciplines from which art and psychology emerged. Although there are many ways to analyze artworks, it is inherent to the system that most of them come from the humanities. It is probably due to the complexity of art, that research usually applies qualitative methods. Another reason might be rifts between disciplines and a defense against statistics. However, for a complex understanding of aesthetic and creative processes it is crucial to use both idiographic and nomothetic data (Machotka, 2012). Idiographic approaches explore the individual and its introspection revealing personal motives, capacities, life trajectories etc. Nomothetic methods investigate regularities across subjects and seek for generalizability. While the first offers specification and richness of data, group data clarifies underlying structures and allows generalizability of phenomena. And when it comes to investigating the interrelationships between art and psychological variables, psychometrics is particularly relevant.

It has been 50 years since Benjamin (Benjamin, 1969) urged a deliberate handling of traditional concepts like creatorship and ingenuity, eternal value and mystery. Still today, research (Leder et al., 2012) states, that art is a mysterious aspect of human experience and most studies focus on measuring constructs other than the image and its individual expression, largely focusing on appreciation and other responses to art (e.g., Martindale, 2007; North & Hargreaves, 1996; Specht, 2007; Tröndle & Tschacher, 2012). Those focusing on artworks are often insufficient in their operationalization. Most studies do not deal with art in a dimensional way, like we do with psychological scales. Instead they are operationalized in terms of a dichotomous categories such as epochs (e.g., expressionism vs. impressionism) or one-dimensional classifications (e.g.,

abstract vs. representational; Gridley, 2006, 2013; Leder et al., 2012; Mastandrea & Crano, 2019; Meskin et al., 2013). Detailed approaches are needed to measure artworks in nuanced profiles, because only by addressing more components inherent to art, can justice to its complexity be done. This might gradually solve the mystery of art and the ways in which it affects us.

Arts and humanities offer manifold approaches to analyze images. Still, they do not meet methodical principles of disciplines predominantly using quantitative methods (e.g., psychology, sociology). Social sciences, in turn, lack the comprehensive art-theoretical basis that fine arts and art history have. Transdisciplinary approaches are needed to bridge this gap (Leder et al., 2018) and provide a reliable, psychometric tool that is based on art-theoretical contents.

Measuring Instruments

A domain that offers a range of methods to describe images with regard to psychological variables is art therapy, in which qualitative content analysis for systematizing pictorial works (Thyme et al., 2013) or the phenomenological picture analysis (Stuhler-Bauer & Elbing, 2003) serve as examples for systematically assessing pictorial works with qualitative methods. But there are also attempts to quantify images. Several of them are not designed for art in general, but rather for therapeutic or clinical use (e.g., *Diagnostic Assessment of Psychiatric Art* (Hacking et al., 1996), *DokuPro* (Elbing & Hölzer, 2007; Elbing et al., 2009), *Nürtinger Rating Scale* (Eber et al., 1998; Elbing & Hacking, 2001), *Systematic Picture Analysis* (Gruber et al., 2002)). Other instruments like the *Diagnostic Drawing Series* (Cohen & Mills, 2015; Mills, 2003) and the *Formal Elements Art Therapy Scale* (Gantt, 2016; Gantt & Anderson, 2009) can be used for assessment beyond clinical applications, but concentrate on specific drawing tasks and therefore are not suitable for all types of images. Other methods like the *Consensual Assessment Technique* (Amabile, 1982) or the *Creative Solution Diagnosis Scale* (Cropley &

Kaufman, 2012) can also be used to assess artistic products. However, their main construct of interest is creativity of a product or a person, not the product itself with regard to its pictorial expression. Other approaches use digital technology to objectify image analysis, for example in terms of computerized assessment of art-based instruments (Mattson, 2010), combining existing measures with public domain image analysis software (Mattson, 2009, 2011), computer-based systems rating basic color-related aspects (Kim et al., 2007) or using computational network science for assessing global image properties (Hayn-Leichsenring et al., 2020).

A tool that shares a similar approach with this work, is the *Assessment of Art Attributes* (Chatterjee et al., 2010; Penn Center for Neuroaesthetics, 2019) although it stems from a neuropsychological perspective. By connecting art and science it aims to quantify artworks including to both, formal-perceptual and content-representational attributes. The scale consists of 14 items assessing broad categories. The results suggest medium to high agreement between raters, while using non-parametric measures of correlation and having previously conducted a rater training with training slides regarding the attributes. Other statistical quality criteria, such as the test's capacity to differentiate between pictures or subgroups of pictures or factors analysis, are not reported. It has been tested on a small sample of prominent paintings from the western canon, and two individuals with Alzheimer's disease (van Buren et al., 2013). But as far as we can tell, it was not further validated on a larger sample of other pictorial material. To summarize, several instruments exist, but there is a lack of quantitative, standardized and reliable measures, that allow inference statistics and are validated on and meant to be applicable to all sorts of two-dimensional pictorial works – whether professional or non-professional art.

Goal of the Study

This study aims to validate the rating instrument for two-dimensional pictorial works (*RizbA*) with a focus on its generalizability to the entirety of handmade two-dimensional pictorial works created by nonprofessional artists. For test validation, the present study uses a large representative sample of pictorial works. Statistical parameters are calculated for examining the quality criteria of single items as well as for the overall test.

Material and Methods

Given the diversity of methods available to analyze images, this work focuses on formal picture analysis, a phenomenological method (Stuhler-Bauer & Elbing, 2003) originating from the humanities and the arts. It refers to a factual, objective analysis of the visual aspects in the image, leaving aside issues such as the observer's subjective associations, reference to its specific era of creation, or its societal relevance (Bauer, 1996).

The Rating Instrument for two-dimensional pictorial Works (*RizbA*)

RizbA is a questionnaire for formal picture analysis of two-dimensional pictorial works. In a previous pilot study (Schoch et al., 2017) a preliminary test version was developed. Test construction was theory-led and empirically-based using an underlying measurement model similar to formative models. The theoretical model and the resulting pool of 113 items was based on art theory, art historical, and art therapeutic literature as well as existing instruments. For the final version items statistically fulfilling criteria of validity and reliability were selected.

Defining the nomological network, the instrument assesses the construct of pictorial expression, which is defined as artistic creation in the form of a picture. The instrument focuses on the concept of a formal picture analysis (Bauer, 1996), analyzing formal aspects like representation, color, shaping, motion, and composition, which can be found across art literature (e.g., Arnheim, 2013; Bauer, 1996; Kandinsky, 1955; Meyer, 2011; Vollmar, 2008). This approach is rooted in the tradition of phenomenological picture analysis (Stuhler-Bauer & Elbing, 2003,

2004), seeking to overcome accidental judgement, preconception, and association (Streb, 1984). The relevant systemic level is the pictorial representation consciously excluding the picture reference, e.g., knowledge, associations, emotions (Huber, 2005). It focuses on visual presentation (Wiesing, 2005) leaving out the colonializing view of unintentional identification of objects (Marotzki & Stoetzer, 2006). It is limited to a detailed but classical conception of images, not taking into account the creation process (Uzelac, 1998). The test does not judge the creator's achievement or mastery and is distinct from aesthetic appreciation. It is neither evaluative nor interpretative nor projective, but aims for a value-free description of its formal elements. A rater training or sample images for certain characteristics are deliberately dispensed to avoid a manipulation of judgement.

The questionnaire includes 26 items (see Table 3), currently in German language (see Table) and uses a bipolar six-point Likert-scale, which is discretely scaled and verbally anchored in shades of agreement (0 = *strongly disagree*, 5 = *strongly agree*). Raters using the questionnaire receive a brief instruction to rate the image presented using the *RizbA* questionnaire. They are asked to focus on the pre-dominant overall expression of a picture and no single details, while endorsing there is no right or wrong, but only their valuation.

Pictorial Material

In art psychology, often both experts and non-experts serve as participants in studies on art reception. However, art psychology should representatively address both, the viewer and the artwork. Unfortunately, studies in the field do not offer a broad range when it comes to pictorial works. Image samples are mostly limited to paintings by white cis male professional artists recognized by an Euroamerican historiography. This practice disregards current discourses on inclusion and diversity, but also a rather old discourse on self-taught and Outsider Art (Cardinal, 2009; Tansella, 2007). To broaden these perspectives and aiming for a general applicability to all two-dimensional pictorial works, whether professional or non-professional, this study uses

stimuli by non-artists. This makes particularly sense if we consider pictorial expression in terms of a psychological construct, of which we aim to assess the entire scope of possible manifestations, not only the upper end. This is also consistent with art therapy assessment, in which non-artists are the primary target group.

Although referring to a democratic view on creativity (Kampylis & Valtanen, 2010), we do not want to open the discourse on whether nonprofessional's works should be regarded as art. Therefore, hereinafter the term 'pictorial work', instead of artwork is used. Until further evaluation studies have been conducted, the test is unendorsed for machined works (e.g., photography), three-dimensional work (e.g., sculpture) or printing techniques (e.g., etching). Therefore, only handmade two-dimensional works were used.

Pictures were photographed and used as digital images at a size of 500 x 357 pixel. The stimulus material consisted of works created by nonprofessionals and was collected in a prestudy (see Figure 1). As a criterion for exclusion we defined expertise in visual arts. Participants never received a professional art training, neither in terms of an academic degree nor ever having studied a subject related to visual arts.

The pictorial data was gathered at four universities and included contributions from students of psychology, human medicine, social work, pedagogy, and theatre pedagogy. Further participants were acquired by an appeal, distributed via social media, asking participants to join the study at a non-university location. As an incentive, vouchers were raffled among the subjects. Since the images should reflect a realistic spectrum of non-professional pictorial expression, the subjects could choose freely from a wide range of artistic materials. Materials offered included different sizes of paper, pencils, colored pencils, graphite, charcoal, felt pen, oil pastels, watercolor, opaque white, tempera, gouache, acrylic color, paintbrushes in various sizes, palette-knives, sponges, glue, scissors, crepe tape, cardboard, and magazines. The subjects were additionally invited to bring their own material. Each participant received an instruction,

in which they were asked to create two pictures, choosing freely from the available materials, followed by a brief naming of the different materials. The manner and content of their work was explicitly left to them, emphasizing spontaneous expression as the only purpose. For economic reasons and to gain a large sample, each participant created two pictures, one after another or concurrently, however they chose.

Picture Sample Size

Regarding the required sample size of pictorial material, we oriented towards Guadagnoli and Velicer (1988) as well as Osborne and Costello (2009), in which two important criterions for predicting outcomes are stated. First, a common practice and rule-of-thumb is a subject to item ratio of at least 10:1. In this case, it would be ten pictures per item, resulting in a minimum sample size of at least 260 images. Second, since PCA is a large sample procedure (Suhr, 2005) and the bigger the sample the better, a sample of 300 might provide even better adequacy (Osborne & Costello, 2004).

Figure 1. Study design

Study

Study Design

Data was collected in an online study using SoSci Survey (Leiner, 2018) with additional use of Hypertext Preprocessor (PHP) elements. The study consisted of a test (T1) and a retest (T2) at intervals of a minimum of four weeks (see Figure 1), in which experts were asked to rate the

pictorial material using *RizbA*. So far the questionnaire uses art-related terminology, which is well-known to art professionals but might not be familiar to others. Therefore, only raters with expertise in the field of visual art were invited.

Raters were acquired via email, websites, and social media. In total 122 universities, 44 associations and 23 art museums in Germany, Austria, and Switzerland were contacted. As a criterion for inclusion, only experts who have an academic degree related to visual art or are students, who have been studying an art-related subject for at least one year were requested to join. No rater training was conducted. The instructions were brief. Raters were asked to rate each picture presented by using the *RizbA* questionnaire followed by the specific instructions mentioned above. As an incentive, vouchers were raffled among participants.

Sample Size

There are only a few empirically based guidelines regarding the statistically required sample size of raters. Some literature recommends two or three observations per subject, given $\alpha = .05$ and a required reliability of at least .40 (Walter et al., 1998). Since this is not an application study, but the constitutive calibration of a new questionnaire, either way we opted for a larger number of raters. Addressing sample size and ICCs, for an exploratory approach, an a priori set cut-off point and a pilot study are recommended (LeBreton & Senter, 2008). The previous pilot study (Schoch et al., 2017) yielded moderate to very strong rater agreement with eight to twelve raters per image. We now aimed for 20 raters per picture at T1, so that despite a dropout rate, there would be still ten raters per picture at T2.

Procedure

In the online survey for each rater seven pictures were randomly, successively drawn without returning. When all images were rated once, the process began anew. The order of items was randomized between subjects, but remained the same within subjects at T1 and T2. For anonymously assigning raters' data sets from the test and the retest, self-generated anonymous

personal codes (Pöge, 2011) were used. Based on these, a database was established, which allowed to identify the pictures, rated by each subject and to individually assign and present the exact same stimuli in the retest. After completing the test, raters could agree to be invited to the retest, leaving their email address, which was stored separately from the data on the grounds of maintaining anonymity. If raters skipped items, they were forced to answer all questions to be able to continue. At T1 a bug in the PHP code occurred: If subjects did not answer all items, but intentionally or accidentally skipped one or more, another image was drawn, presented, and additionally rated by them. Therefore, few subjects processed more than the intended seven pictures and several missing values occurred in the T1 data set. To maintain equal conditions, only each rater's first seven pictures were included in the statistical analysis. Missing values were not imputed. At T2 this was fixed.

Statistical Analysis

Statistical analysis of the questionnaire's quality criteria focused on two units. Pictorial works served as the unit for item means, item standard deviation, item difficulty, capacity of differentiation, test-retest reliability, factor analysis, and index of factor similarity. Raters served as the unit for inter-rater reliability. Since the data included three levels (items, pictures, raters), for each analysis, the relevant levels were chosen and the others averaged. The dataset for the rater level inherently included a large number of empty cells, because each rater assessed only a small selection of the total pictorial material. For evaluation, individual raters were merged to new generic raters. After restructuring the dataset, there were different amounts of raters per image, which lead to a few missing values. Missing values were not imputed. Excel including Visual Basic for Applications was used for structuring the data sets. All statistical analyses were analogically conducted for both times of measurements, test (T1) and retest (T2). The data was analyzed using SPSS Statistics 24.

Descriptive item statistics were calculated. As a measure for item difficulty, Dahl's difficulty index for interval-levelled data (Moosbrugger & Kelava, 2007) was computed. Capacity of differentiation between images was examined for each item as well as the overall test by determining criteria, such as item variance and item effect size. The latter was computed using MANOVA, in which with the images served as independent variables, the items as dependent variables and the partial eta-squared (Levine & Hullett, 2002) was interpreted as an effect size estimator. To examine test-retest reliability between T1 and T2, Pearson product-moment correlations were calculated.

So far there is a lack of quantitative data on pictorial expression to which one could refer. This is why to start with an exploratory approach of test construction instead of a CFA was chosen. Since there was no plausible theory for the factor structure, and based on recommendations (Brown, 2009; Suhr, 2005) a PCA followed by varimax rotation was conducted as an exploratory data reduction procedure (Osborne & Costello, 2004) to reduce the number of variables in a large sample of interrelated measures (Suhr, 2006). Results were compared to Horn's parallel analysis, using O'Connor's syntax (O'Connor, 2000, 2018). Factors with larger eigenvalues than parallel components were extracted in another PCA followed by varimax rotation.

As a further index on accepting or rejecting the given factors solution, Tucker's coefficient of congruence (Tucker & Lewis, 1973) for factor similarity was calculated, which represents the cosine of the angle between two vectors and a standardized measure of proportionality of elements (Lorenzo-Seva & Ten Berge, 2006). The procedure was followed as described by Wünsch (2016) and Fetz et al. (2018). The dataset was split into two random samples, a PCA followed by varimax rotation was calculated and the suggested eight factors were extracted. The resulting component loading matrices were both compared with each other. Two-tailed Pearson correlations ($\alpha = .05$) were calculated. Based on significant correlations between

factors from both samples, factors were accordingly paired and Tucker's coefficient of congruency was calculated. Values of this statistical parameter per definition range between +1 to -1 and can be interpreted similar to the Pearson's product-moment correlation (Chan et al., 1999). Since it depicts the degree of deviation between two factors, negative coefficients can be similarly interpreted to positive values. Values were interpreted as outlined by Chan et al. (1999) and Lorenzo-Seva and Ten Berge (2006), where a critical congruence level rather than unity for indication of identical factors is suggested. Accordingly, values between .85 - .94 indicate fair similarity, whereas values higher than .95 suggest two factors or components being equal.

For measuring relative consistency between the expert ratings (LeBreton & Senter, 2008), inter-rater reliability was computed for each item and the overall test. Since not all pictures were rated by each subject and values were averaged, inter-rater reliability was computed using intraclass correlation coefficients (*ICCs*) one-way random average measures with absolute agreement (McGraw & Wong, 1996; Shrout & Fleiss, 1979).

Results

The final stimulus material included in the study consisted of 294 pictures by 147 nonprofessionals (107 female, 39 male, one diverse) at an age between 17 and 71 years ($M = 33.86$, $SD = 14.02$), who differed in levels of education and occupation. T1 was completed by a total of 895 raters. Two raters were excluded, due to not fulfilling the criterion for inclusion. 13 more raters were excluded for not seriously processing the questionnaire, which was discernible in their random or extreme patterns of answering. The final sample of valid raters at T1 consisted of 880 experts (see Table 1). All of them had an academic degree in an art-related subject (65.5 %) or were studying at least in their second year (34.5 %). 475 of these raters participated in T2. Each picture was rated by 16 to 31 raters ($M = 20.35$, $SD = 5.23$) at T1 and by one to 20 raters ($M = 10.33$, $SD = 3.36$) at T2.

Table 1. Rater sample (T1)

Gender	Age	Discipline
750 female	18 and 74 years	art history (30.6 %), art therapy (16.8 %), art pedagogy (13.0 %),
122 male	($M = 31.59$, $SD =$	visual art (11.5 %), graphic design (5.3 %), design (5.3 %), art
eight diverse	10.18)	history and aesthetics (4.8 %), restoration (1.3 %), various others (11.4 %)

The calculated item means range from 0.30 to 3.07 ($M = 2.23$, $SD = .60$) at T1 and from 1.51 to 3.01 ($M = 2.34$, $SD = 0.47$) at T2. Item standard deviations lie between 0.49 and 1.31 ($M = 0.93$, $SD = 0.21$) at T1 and between 1.15 and 1.82 ($M = 1.45$, $SD = 0.16$) at T2 (see Table 3). Item difficulty p_i ranges from .30 to .61 ($M = .47$, $SD = 0.09$) at T1 and from .31 to .60 ($M = .47$, $SD = 0.09$) at T2 (see Table 3). Capacity of differentiation in the form of items' variance s_i^2 ranges from 0.24 to 1.72 ($M = 0.91$, $SD = 0.38$) at T1 and from 1.31 to 3.32 ($M = 2.12$, $SD = 0.46$) at T2 (see Table 3). Partial eta-squared ranges from .18 to .63 ($M = .41$, $SD = 0.11$) at T1 and from .25 to .69 ($M = .46$, $SD = 0.11$) at T2 (see Table 3). The overall test yields a partial eta-squared of .28 (Pillai's trace: $F(7618, 137540) = 6.86$, $p = .000$) at T1 and .33 (Pillai's trace: $F(7618, 70616) = 4.47$, $p = .000$) at T2. Overall test-retest reliability yields a Pearson product-moment correlation of .93 ($r(7644) = .93$, $p = .000$).

A PCA followed by varimax rotation for both, T1 and T2, results in eight factors with eigenvalues above one. A scree plot suggests eight factors. Compared to Horn's parallel analysis eight factors have larger eigenvalues than parallel components at both times of measurement. In another PCA followed by varimax rotation eight factors are extracted accounting for 62.18 % (T1) and 63.95 % (T2) of the common variance (see Table 4 and 5). Tucker's coefficient of congruence ranges between |0.98| and |1.00| at T1 and between |0.82| and |1.00| at T2 (see Table 2).

Items' ICCs range from .81 to .98 ($M = .93$, $SD = 0.04$) at T1 and from .52 to .98 ($M = .84$, $SD = 0.12$) at T2. Overall ICC are .81 ($F(59, 1440) = 5.22$, $p = .000$, 95% CI [0.73, 0.87]) at T1 and .84 ($F(3, 68) = 6.25$, $p = .001$, 95% CI [0.47, 0.99]) at T2 (see Table 3).

Discussion

This article outlines the validity of *RizbA* by means of a statistical approach. Difficulty indices of all items are approximately medium-sized and therefore provide an excellent mechanism for to differentiating between characteristics. Another criterion that can be used to indicate capacity of differentiation between pictures, are the moderate to large item variances. Consequently, the overall test explains 27.5 % (T1) and 32.5 % (T2) of the variance. By comparing results to common effect sizes (Coster, 2012), large effects can be assumed. Still, compared to the previous pilot study on *RizbA* (Schoch et al., 2017) the capacity of differentiation is substantially smaller. This might be explained by the prior study using a sample of only nine heterogeneous images. The bigger the sample, the more likely it is to contain similar images, especially when they are created by nonprofessionals with a technical and artistic repertoire that is probably not as diverse as an artist's. Relating to common effect size conventions (Cohen, 1992), the overall test-reliability suggests a large effect and implies *RizbA* being a stable measure across time.

The results empirically support the existence of an interpretable factor structure. Factor analysis suggests an eight-factors solution, which is largely consistent across both measurement points and accounts for a large part of common variance. Prospective factor labels were a posteriori designated based on the empirical data and theoretical reasoning. This involved an iterative discussion process of expert opinions referring to art theory, which included experts of the HKS – University of Applied Sciences and Arts in Ottersberg and the German Creative Arts Therapies Research Association (FVKT). Resulting factor labels (in corresponding order to

Table 4) might be *picture effect*, *spatiality*, *shaping*, *pictorial elements* (*drawing vs. painting*), *representation*, *color intensity*, *color mixture*, and *composition*.

Picture effect refers to the perceived atmosphere of the images, *spatiality* to terms of perspective, and *shaping* to organic, round vs. angled and geometric shapes. *Pictorial elements* (*drawing vs. painting*) points out graphic and pictorial features, while *representation* reflects the degree of abstraction. *Color intensity* and *color mixture* refer to vibrance and the use of primary and secondary colors. Finally, *composition* relates to the dominant layout being horizontal or vertical. While the *spatiality*, *shaping*, *representation*, and *composition* are relatively easy to be interpreted, others raise more questions. *Pictorial elements* (*drawing vs. painting*) includes two categories, drawing and painting, which depending on the image can potentially differ, but also merge into each other. *Color intensity* and *color mixture* surprisingly result in two – instead of as theoretically hypothesized – one factor. Also the three 2-item factors prove to be difficult with regard to covering a whole dimension and must be further examined in future studies.

Another question to be clarified is the meaning of a total scale score. The formation of a sum value across all items is not meaningful, since a high or low proficiency can only be interpreted at an item or factor level. Instead RizbA provides data profiles, that capture pictorial characteristics of an image. Regarding further research, these can be combined with profile similarity measures (Furr, 2008; Helmstadter, 1957) and hereby combined with other scales.

The number of extracted factors is different from the prior pilot study, which indicated only four factors but in an exploratory approach with only a small sample of 12 images. However, several factors are similar across both studies. In particular, *shaping* and *picture effect* are stable dimensions, while other, such as *representation* and *spatiality*, change into multiple, more differentiated factors. The first factor *picture effect* is stable across times of measurement and studies and contains high factor loadings. However, in terms of content validity it does not

clearly measure the defined concept of formal picture elements. Thought must be given to exclude this factor in the final test version.

Nearly all of Tucker's coefficients of congruence at both times of measurement indicate a high factor similarity, except for one at T2 suggesting fair similarity. The values obtained by this procedure to calculate coefficients are rather conservative (Wünsch, 2016). The usual procedure stipulates a Procrustean factor rotation with orthogonal rotation to maximum similarity (Chan et al., 1999; Cliff, 1966; McCrae et al., 1996; Raykov & Little, 1999; Schönemann, 1966; Zumbo et al., 2003) using specific syntaxes (Fisher, 2012; Zumbo et al., 2003). However, we decided to go for the conservative approach, because this study already contains several exploratory aspects due to the lack of paradigms in quantitative picture analysis.

Based on statistical guidelines (LeBreton & Senter, 2008), inter-rater reliabilities imply a moderate to very strong agreement on the item level. For the overall test, results can be interpreted as strong consistency among raters. Especially compared to the few existing studies on inter-rater reliability in art therapy assessment (Eitel et al., 2008; Mattson, 2011), results are substantial.

Limitations

In the prestudy, each nonprofessional created two pictures. For economic reasons this was a feasible approach, but from an artistic point of view, the images cannot be regarded as two independent products of spontaneous expression, but rather as a short-term artistic process. This might reduce the heterogeneity of the sample and consequently result in a lower capacity of differentiation.

The stimulus material was collected in Germany. About 73 % of the participants are female, 27 % male and less than 1 % diverse. Considering the rater sample, experts stem from German speaking countries (Germany, Austria, Switzerland). Here about 85 % are female, 16 % male and less than 1 % diverse. Thus, the image sample is biased in terms of gender and ethnicity

towards a western, in particular Eurocentric, and probably objectivistic cultural context. Since art, as well as several psychological variables differ across cultures (Cattaneo, 1994), it can be assumed that there might be differences as well, when it comes to art-making and rating art. Further studies should investigate the instrument's validity by taking into account diversity with a critical view on Eurocentrism (Mosquera, 1992). This refers to gender, ethnicity, cultural background and other contexts relevant to both, pictures and raters.

On account of the maximum data capacity allowed by SoSci Survey (Leiner, 2018), the size of each digital image was rather small, a fact some raters noted to be irritating. A different rating experience may be expected, when original works instead of digital images are used. For economic reasons, based on the sample size needed, this was discarded.

PCA indicates that the questionnaire is measuring a construct consisting of several interpretable components. However, without conducting a Confirmatory factor analysis (CFA), no final conclusions can be drawn. Also, due to the non-conclusive state of the instrument, other statistics which are usually used in test validation, could not appropriately be conducted at that time. This includes statistical values for internal consistency, that require a test scale to measure one construct (e.g., split-half reliability, Cronbach's α), likely but uncertain condition so far.

For pragmatic reasons, in this article ancient concepts are excluded, for example theories of ingenuity (Schneider, 2017), which belong to a more theoretical art-historical discourse, but also historical foundations such as Fechner's experimental aesthetics (Fechner, 1876). Instead, reference is made to contemporary scientific discourses, whereby the focus lies solely on visual arts, excluding other arts such as music or performing arts.

Implications

The instrument was successfully validated on a sample of nonprofessionals' handmade pictorial works and might soon be generalized to the entirety of two-dimensional pictorial works. As a novelty in research, *RizbA* enables a reliable capture of artworks. This allows for a more detailed

examination of psychological correlates, giving visual art a greater role to play when it comes to considering the psychology of art. However, one thing is certain: It is now possible to show that formal picture analysis can be measured quantitatively, resulting in a plausible factors solution. Offering a new way of reliably assessing artworks and pictorial expression, *RizbA* opens new perspectives in psychometrics and transdisciplinary research between psychology and art.

This work has various implications when it comes to artworks, art-making, and creativity, in terms of documentation and research. Disciplines like art therapy share the idea of client-art relations (Bat Or & Zilcha-Mano, 2018) and the artwork being related to inner representations of its creator. A tool for image analysis can be used to objectively document and analyze artistic processes and test these hypotheses. Although creativity, aesthetics, and art-making overlap to a large extent (Zaidel et al., 2013), research usually focuses on these topics separately (Tinio, 2013). In terms of psychological research such a tool fosters a more holistic view on art by not only considering art reception, but also interrelationships between artworks, viewing and creating them. Considering creativity a relevant factor for artistic creation, concepts like creative expression and creative judgment (Amendt-Lyon, 2001) are likely to manifest in artworks. Combining *RizbA* with existing psychometric instruments would allow an investigation of a multitude of the interrelations, for example with creativity, personality traits or aesthetics.

Artworks

Art therapy is a discipline which, among other things, focuses on client-art relations (Bat Or & Zilcha-Mano, 2018). One of its fundamental assumptions is that the artwork is related to its creator as well as to a multitude of her or his psychological constructs, such as personality, emotions, and social skills. Although pictures' formal elements are rarely understood in a direct diagnostic way, certain psychological aspects are associated with it (Penzes et al., 2018).

Instead of focusing only on psychological and physiological correlates variables, which are expected to relate to art, research could provide a more detailed view by involving numeric measurements of images. Beyond research *RizbA* is an economic tool for documenting and analyzing artworks in the art therapy process, in art education or other art-related practices. For example, during an art therapeutic treatment both, changes in the pictures as well as in psychological variables, can be measured. That way, the client's coping mechanisms and health improvement or even effectiveness of therapeutic methods can be assessed in detail. Another implication is teaching art theory, art history, or art pedagogy. Here it can be beneficial to grasp a phenomenological concept of processing art by using the instrument for a critical reflection on one's own subjective views and aesthetic appreciation while striving for objectivity.

Also, several correlations have been found between art preferences and personality traits. For example, when it comes to perceiving art, individuals enjoying novelty, ambiguity and dissonance, who are open-minded and sensation seeking, prefer abstract over representational art. The same accounts for people, who are abstract, random, and field-independent thinkers (Gridley, 2006, 2013). Using *RizbA* these results can be further examined going beyond the categories of abstract vs. representational. Studies might find out, which actual formal criteria make these individuals like abstract art better. Maybe it is the composition?

Art-making

Although creativity, aesthetics, and art-making have to be differentiated, they correlate and overlap to a large extent (Zaidel et al., 2013). Still, research usually focuses on these topics separately (Tinio, 2013) and ignores artistic creation itself, which limits the possible view on experiencing art. But there are approaches to overcome this divide. One is the Mirror Model of Art, connecting art-making and art-viewing and which proposes that aesthetic experiences mirror the art-making process (Tinio, 2013). There is also research investigating affective, evaluative, and neurophysiological correlates of perceiving and interacting with art (Pelowski

et al., 2017). *RizbA* could contribute to these more holistic views on art, investigating interrelationships between artworks, viewing and creating them.

Further research might thus ask participants not only to look at art, but to create their own art. If abstract thinkers prefer abstract over representational art (Gridley, 2013), they might also be more likely to create abstract than representational art. For assessing resulting images as detailed as personality traits, *RizbA* is a useful tool.

Creativity

The understanding of art has changed considerably since the early 20th century. But even nowadays art-making can hardly be imagined without the concept of creativity. Leaving aside other heterogeneous and complex perspectives on creativity (Glăveanu, 2013; Rhodes, 1961; Said-Metwaly et al., 2017), we hereby refer to a contemporary conception: A democratic view, in which anyone is considered able to create from anything (Kampylis & Valtanen, 2010). This view is particularly interesting when dealing with nonprofessionals' works. Creativity is a relevant theoretical factor for artistic creation. When it comes to art-making, concepts such as creative expression and creative judgment (Amendt-Lyon, 2001) are aspects that are likely to manifest themselves in artworks. It is likely then that creativity, creative expression, creative judgement, and other related aspects, such as complexity and openness to experience, are expressed in pictures. However, possible interrelations between creativity and artworks have barely been investigated quantitatively. Although several instruments exist, assessment of creativity is both a controversial and unsettled issue (Said-Metwaly et al., 2017). By combining *RizbA* with existing psychometric instruments, a multitude of interrelations could be investigated (e.g., with creativity, personality traits or aesthetics).

Future Research

Further Validation

For a generalization to other sorts of two-dimensional pictorial works and a better understanding of the factor structure, further validation studies on different image samples are needed (e.g., historical and contemporary art, pictures by children and adolescents, non-handmade techniques). For this purpose, another subsequent study on contemporary art with a representative sample of pictures and raters was currently conducted. Furthermore, several pilot studies on specific samples were implemented, comparing pictorial expression between clinical subgroups, such as clients with chronic pain (Janßen, 2018) or recurrent depressive disorder (Epstein, 2019), and healthy control groups. Initial results imply differences between groups suggesting correlations between artworks and clinical constructs.

Future research will also include other populations of raters, for example art experts from other cultural areas or raters from other professions than the arts. Addressing the latter issue, a study has currently been conducted, testing the questionnaire on expert vs. non-expert raters in art (Jerusalem, 2020, February 20). The resulting test manual explains art-related terminology and ensures that the questionnaire is also applicable by researchers and practitioners from other domains without art training (e.g., psychology, pedagogy, psychiatry).

The next obligatory step is a CFA, in order to verify the factor structure of the observed variables (Suhr, 2006). Therefore, we are currently conducting another study, using new image material and computing a CFA based on an evolved theoretical model of the construct pictorial expression. For now, data was averaged for the relevant level of analysis. Future statistical approaches could also attach greater importance to the data in terms of multilevel modelling (Silvia, 2007). Another point will be to translate the questionnaire into English to make it accessible to more users.

In terms of test validation, it is common practice to examine convergent and divergent validity by multitrait-multimethod matrix analyses (Campbell & Fiske, 1959). Further studies are planned using related but distinct scales, such as the *Assessment of Art Attributes* (Chatterjee et al., 2010) and the *Diagnostic Drawing Series* (Cohen & Mills, 2015; Mills, 2003).

Psychological Correlates of Art

As the first quantitative, reliable tool for formal picture analysis, *RizbA* allows a wide range of applications. In fundamental research, the rating instrument offers an additional measurement for analyzing art psychological issues, that have not been assessed quantitatively until now. Psychological correlates of art can be further investigated by giving the art more weight in psychology of art. Application-related, therapeutic processes in diagnostic, therapeutic or art therapeutic settings, can be objectively documented and evaluated in clinical documentation, serving professionals from different disciplines to gain more insight into a client's condition. Assumed interrelationships between pictorial expression and psychological constructs (e.g., personality traits, clinical diagnoses) could be examined. However, for both fundamental and applied research, empirical trials must be implemented. In case empirical proof for such correlations is to be found, *RizbA* might, in perspective, serve as a diagnostic inventory, enhancing non-verbal psychological assessment.

Art and Psychometrics

Until now, there has been a lack of instruments that are capable of reliably assessing all the issues addressed above. *RizbA* makes a difference by providing a tool for an objective analysis of artworks. However, it lies somewhere within an area of transdisciplinary tension. From the perspective of psychological correlates of art, an individual's pictorial expression can be regarded as a psychological construct. If the hypothesis of artworks being related to its creator's inner representations can be replicated during empirical tests, we can assume art to become a predictor for established psychological constructs. An image could be a manifestation of latent

variables such as creativity, personality, intelligence, behavior, clinical diagnoses, and other aspects relevant to psychological research, diagnostics, and practice. Coming from this point of view and seeing the artwork as a characteristic of its creator's psyche, a questionnaire measuring pictorial expression can be regarded in terms of a psychometric test. Also, *RizbA* meets statistical requirements. On the other hand, by primarily collecting the expression of a picture, it is only possible to refer to its creator's mind indirectly. Until now, there are few generalizable results on empirical correlations between images and psyche. From this position, it would be more appropriate to speak of a standardized technique for formal picture analysis, rather than a psychometric test.

In summary, a satisfactory answer cannot be provided yet. *RizbA* has analogies to psychometric tests, but refers to formal picture analysis. And as often happens with transdisciplinary approaches, they do not fit into clear categories. However, referring to different disciplines can foster progress in research, for example when it comes to construction and integration of psychological theories (Gigerenzer, 2010).

Conclusion

In summary, the results are promising and suggest fair to high test quality in terms of reliability and validity. As a result, an application of the instrument on non-professional's handmade two-dimensional pictorial works can certainly be recommended. As a universal tool for quantitative formal picture analysis, it offers a multitude of implications in practice as well as in research. In the future, a generalizability to all types of two-dimensional pictorial expressions will be pursued. This includes works by nonprofessionals as well as artists, images from non-western cultures, handmade works, and those created under conditions using technical devices such as print and photography.

References

- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of personality and social psychology*, 43(5), 997. <https://doi.org/10.1037/0022-3514.43.5.997>

- Amendt-Lyon, N. (2001). Art and creativity in Gestalt therapy. *Gestalt Review*, 5(4), 225-248.
- Arnheim, R. (2013). *Kunst und Sehen: eine Psychologie des schöpferischen Auges [Art and seeing: A psychology of the creative eye]*. Walter de Gruyter.
- Bat Or, M., & Zilcha-Mano, S. (2018). The Art Therapy Working Alliance Inventory: The development of a measure. *International Journal of Art Therapy*, 1-12. <https://doi.org/10.1080/17454832.2018.1518989>
- Bauer, H. (1996). Form, Struktur, Stil: Die formanalytischen und formgeschichtlichen Methoden [Form, structure, style: The methods of form analysis and form history]. In H. Belting, H. Dilly, W. Kemp, W. Sauerländer, & M. Warnke (Eds.), *Kunstgeschichte: Eine Einführung [Art history: An introduction]* (Vol. 6, pp. 157-173). Reimer.
- Benjamin, W. (1969). *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit [The artwork in the age of mechanical reproduction]* (Vol. 2). Suhrkamp.
- Brown, J. D. (2009). Principal components analysis and exploratory factor analysis: Definitions, differences, and choices. *Statistics*, 13(1).
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, 56(2), 81.
- Cardinal, R. (2009). Outsider Art and the autistic creator. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1522), 1459-1466. <https://doi.org/10.1098/rstb.2008.0325>
- Cattaneo, M. (1994). Addressing culture and values in the training of art therapists. *Art Therapy*, 11(3), 184-186. <https://doi.org/10.1080/07421656.1994.10759081>
- Chan, W., Ho, R. M., Leung, K., Chan, D. K.-S., & Yung, Y.-F. (1999). An alternative method for evaluating congruence coefficients with Procrustes rotation: A bootstrap procedure. *Psychological methods*, 4(4), 378.
- Chatterjee, A., Widick, P., Sternschein, R., Smith, W. B., & Bromberger, B. (2010). The assessment of art attributes. *Empirical Studies of the Arts*, 28(2), 207-222. <https://doi.org/10.2190/EM.28.2.f>
- Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika*, 31(1), 33-42.
- Cohen, B. M., & Mills, A. (2015). The Diagnostic Drawing Series (DDS) at thirty: Art therapy assessment and research. *The Wiley handbook of art therapy*, 558-568.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155.
- Commare, L., Rosenberg, R., & Leder, H. (2018). More than the sum of its parts: Perceiving complexity in painting. *Psychology of Aesthetics, Creativity, and the Arts*, 12(4), 380. <https://doi.org/10.1037/aca0000186>
- Coster, J. D. (2012). *Converting effect sizes*. <http://www.stat-help.com/spreadsheets.html>
- Cropley, D. H., & Kaufman, J. C. (2012). Measuring functional creativity: Non-expert raters and the Creative Solution Diagnosis Scale. *The Journal of Creative Behavior*, 46(2), 119-137. <https://doi.org/10.1002/jocb.9>
- Eber, P., Müller, S., Bader, R., & Baukus, P. (1998). Die Nürnberger Beurteilungsskala zur Einschätzung von Patientenbildern im Therapieverlauf [The Nürnberger Rating Scale for evaluating pictures of patients in the therapeutic process]. *Fachhochschule für Kunsttherapie Nürnberg*.
- Eitel, K., Szkura, L., Pokorny, D., & von Wietersheim, J. (2008). Do you see what I see? A study about the interrater reliability in art therapy. *Psychotherapy research*, 18(2), 216-224. <https://doi.org/10.1080/10503300701441995>
- Elbing, U., & Hacking, S. (2001). Nürnberger Beurteilungsskala und Diagnostic Assessment of Psychiatric Art: Neue Wege zur Evaluation der Bilder von Kunsttherapie-Patienten [Nürnberger Rating Scale and Diagnostic Assessment of Psychiatric Art: New ways of evaluating pictures of art therapy patients]. *Musik-, Tanz- und Kunsttherapie*, 12(3), 133-144. <https://doi.org/10.1026/0933-6885.12.3.133>
- Elbing, U., & Hölzer, M. (2007). Entwicklung und erste Evaluation eines Instruments zur kunst- und gestaltungstherapeutischen Prozessdokumentation [Development and first evaluation of an instrument for art and gestalt therapeutic process documentation]. *Musik-, Tanz und Kunsttherapie*, 18(2), 85-99. <https://doi.org/10.1026/0933-6885.18.2.85>
- Elbing, U., Hölzer, M., Danner-Weinberger, A., & von Wietersheim, J. (2009). Reliabilität und Validität des Instruments DoKuPro: Dokumentation Kunsttherapeutischer Prozesse erste Ergebnisse [Reliability and validity of the instrument,

- DokuPro: Documentation of art therapeutic processes]. *Musik-, Tanz und Kunsttherapie*, 20(1), 1-7.
<https://doi.org/10.1026/0933-6885.20.1.1>
- Epstein, C. (2019). *Bildlicher Ausdruck von Depression: Erprobung des Ratinginstruments RizbA an Bildstichproben aus dem klinischen Kunsttherapiesetting [Pictorial expression of depression: Testing the ratinginstrument RizbA on image samples from a clinical art therapy setting]* Witten/Herdecke University. Zenodo.
<https://doi.org/10.5281/zenodo.3365921>
- Fechner, G. T. (1876). *Vorschule der Ästhetik [Preschool of aesthetics]* (Vol. 1). Breitkopf & Härtel.
- Fetz, K., Vogt, H., Ostermann, T., Schmitz, A., & Schulz-Quach, C. (2018). Evaluation of the palliative symptom burden score (PSBS) in a specialised palliative care unit of a university medical centre-a longitudinal study. *BMC palliative care*, 17(1), 92. <https://doi.org/10.1186/s12904-018-0342-0>
- Fisher, R. (2012). *How to do procrustean factor rotation*. <https://culturemindspace.blogspot.com/2012/03/how-to-do-procrustean-factor-rotation.html>
- Furr, R. M. (2008). A framework for profile similarity: Integrating similarity, normativeness, and distinctiveness. *Journal of personality*, 76(5), 1267-1316. <https://doi.org/10.1111/j.1467-6494.2008.00521.x>
- Gantt, L. (2016). The Formal Elements Art Therapy Scale (FEATS). *D. Gussac & ML Rosal, The Wiley Handbook of Art Therapy*, 569-578.
- Gantt, L. M., & Anderson, F. (2009). The Formal Elements Art Therapy Scale: A measurement system for global variables in art. *Art Therapy*, 26(3), 124-129. <https://doi.org/10.1080/07421656.2009.10129372>
- Gigerenzer, G. (2010). Personal reflections on theory and psychology. *Theory & Psychology*, 20(6), 733-743.
<https://doi.org/10.1177/0959354310378184>
- Glăveanu, V. P. (2013). Rewriting the language of creativity: The Five A's framework. *Review of General Psychology*, 17(1), 69.
- Gridley, M. C. (2006). Concrete and abstract thinking styles and art preferences in a sample of serious art collectors. *Psychological reports*, 98(3), 853-857.
- Gridley, M. C. (2013). Preference for abstract art according to thinking styles and personality. *North American Journal of Psychology*, 15(3), 463-481.
- Gruber, H., Frieling, E., & Weis, J. (2002). Kunsttherapie: Entwicklung und Evaluierung eines Beobachtungsinstruments zur systematischen Analyse von Patientenbildern aus der Onkologie und der Rheumatologie [Art therapy: Development and evaluation of a observational instrument for the systematic analysis of pictures by patients in oncology and rheumatology]. *Complementary Medicine Research*, 9(3), 138-146. <https://doi.org/10.1159/000064263>
- Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological bulletin*, 103(2), 265-275.
- Hacking, S., Foreman, D., & Belcher, J. (1996). The descriptive assessment for psychiatric art: A new way of quantifying paintings by psychiatric patients. *The Journal of nervous and mental disease*, 185(7), 425-430.
<https://doi.org/10.1097/00005053-199607000-00005>
- Hayn-Leichsenring, G. U., Kenett, Y. N., Schulz, K., & Chatterjee, A. (2020). Abstract art paintings, global image properties, and verbal descriptions: An empirical and computational investigation. *Acta psychologica*, 202, 102936.
<https://doi.org/10.1016/j.actpsy.2019.102936>
- Helmstadter, G. C. (1957). An empirical comparison of methods for estimating profile similarity. *Educational and Psychological Measurement*, 17(1), 71-82. <https://doi.org/10.1177/001316445701700105>
- Huber, H. D. (2005). Systemische Bildwissenschaft [Systemic visual culture]. In K. Sachs-Hombach (Ed.), *Bildwissenschaft zwischen Reflexion und Anwendung [Visual culture between reflexion and application]*. Von Halem.
- Janßen, B. (2018). *Erprobung des Ratinginstruments für zweidimensionale bildnerische Arbeiten (RizbA): Ansatz zu einer möglichen Untersuchung des bildnerischen Ausdrucks von Schmerz in gemalten Bildern von Menschen mit chronischer Schmerzsymptomatik [Testing the rating instrument for two-dimensional pictorial works (RizbA): Pilot study for*

- investigating pictorial expression in pictures painted by people with chronic pain symptoms] HKS – University of Applied Sciences and Arts Ottersberg. Zenodo. <https://doi.org/10.5281/zenodo.3407808>
- Jerusalem, J. L. L. (2020, February 20). *Rizba für alle: Überprüfung der Einsatzmöglichkeiten eines Ratinginstruments für zweidimensionale bildnerische Arbeiten (RizbA) und des neu dazu erstellten Manuals [Rizba for everyone: Review of the possible uses of a rating instrument for two-dimensional pictorial works (RizbA) and the newly created manual]* Witten/Herdecke University. Open Science Framework. <https://osf.io/nsw49>
- Kampylis, P. G., & Valtanen, J. (2010). Redefining creativity: Analyzing definitions, collocations, and consequences. *The Journal of Creative Behavior*, 44(3), 191-214.
- Kandinsky, W. (1955). *Punkt und Linie zu Fläche: Beitrag zur Analyse der malerischen Elemente [Point and Line to Plane: Contribution to the analysis of pictorial elements]*. Benteli.
- Kim, S.-i., Bae, J., & Lee, Y. (2007). A computer system to rate the color-related formal elements in art therapy assessments. *The Arts in Psychotherapy*, 34(3), 223-237. <https://doi.org/10.1016/j.aip.2007.02.002>
- LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational research methods*, 11(4), 815-852. <https://doi.org/10.1177/1094428106296642>
- Leder, H., Gerger, G., Dressler, S. G., & Schabmann, A. (2012). How art is appreciated. *Psychology of Aesthetics, Creativity, and the Arts*, 6(1), 2-10. <https://doi.org/10.1037/a0026396>
- Leder, H., Tinio, P. P., Brieber, D., Kröner, T., Jacobsen, T., & Rosenberg, R. (2018). Symmetry Is Not a Universal Law of Beauty. *Empirical Studies of the Arts*, 0(0), 1-11. <https://doi.org/10.1177/0276237418777941>
- Leiner, D. (2018). *SoSci Survey*. <https://www.soscisurvey.de>
- Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, 28(4), 612-625.
- Lorenzo-Seva, U., & Ten Berge, J. M. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, 2(2), 57-64. <https://doi.org/10.1027/1614-1881.2.2.57>
- Machotka, P. (2012). Understanding aesthetic and creative processes: The complementarity of idiographic and nomothetic data. *Psychology of Aesthetics, Creativity, and the Arts*, 6(1), 43-56. <https://doi.org/10.1037/a0025243>
- Marotzki, W., & Stoetzer, K. (2006). Die Geschichten hinter den Bildern: Annäherungen an eine Methode und Methodologie der Bildinterpretation in biografie- und bildungstheoretischer Absicht [The stories behind the pictures: Approaches to a method and methodology of image interpretation in biography and educational theory]. In M. Winfried & N. Horst (Eds.), *Bildinterpretation und Bildverstehen: Methodische Ansätze aus sozialwissenschaftlicher, kunst- und medienpädagogischer Perspektive [Image interpretation and image understanding: Methodological approaches from a socioscientific, artistic and media educational perspective]*. VS Verlag für Sozialwissenschaften.
- Martindale, C. (2007). Recent trends in the psychological study of aesthetics, creativity, and the arts. *Empirical Studies of the Arts*, 25(2), 121-141. <https://doi.org/10.2190/B637-1041-2635-16NN>
- Mastandrea, S., & Crano, W. D. (2019). Peripheral factors affecting the evaluation of artworks. *Empirical Studies of the Arts*, 37(1), 82-91. <https://doi.org/10.1177/0276237418790916>
- Mattson, D. C. (2009). Accessible image analysis for art assessment. *The Arts in Psychotherapy*, 36(4), 208-213. <https://doi.org/10.1016/j.aip.2009.03.003>
- Mattson, D. C. (2010). Issues in computerized art therapy assessment. *The Arts in Psychotherapy*, 37(4), 328-334. <https://doi.org/10.1016/j.aip.2010.05.008>
- Mattson, D. C. (2011). Standardizing the Formal Elements Art Therapy Scale (FEATS) rotation scale with computerized technology: A pilot study. *The Arts in Psychotherapy*, 38(2), 120-124. <https://doi.org/10.1016/j.aip.2011.02.003>
- McCrae, R. R., Zonderman, A. B., Costa Jr, P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of personality and social psychology*, 70(3), 552.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, 1(1), 30. <https://doi.org/10.1037/1082-989X.96.S3.00>

- Meskin, A., Phelan, M., Moore, M., & Kieran, M. (2013). Mere exposure to bad art. *British Journal of Aesthetics*, 53(2), 139-164.
- Meyer, G. (2011). *Sprache der Bilder: Kunst verstehen: Form, Farbe, Komposition [Language of the pictures: Understanding art: Shape, color, composition]*. Seemann.
- Mills, A. (2003). The diagnostic drawing series. In *Handbook of art therapy* (pp. 401-409).
- Moosbrugger, H., & Kelava, A. (2007). *Testtheorie und Fragebogenkonstruktion [Test theory and questionnaire design]* (2 ed.). Springer.
- Mosquera, G. (1992). The Marco Polo syndrome: Some problems around art and Eurocentrism. *Third Text*, 6(21), 35-41.
<https://doi.org/10.1080/09528829208576382>
- North, A. C., & Hargreaves, D. J. (1996). Affective and evaluative responses to the arts. *Empirical Studies of the Arts*, 14(2), 207-222. <https://doi.org/10.2190/YQYH-XVMN-21EA-94AC>
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior research methods, instruments, & computers*, 32(3), 396-402.
- O'Connor, B. P. (2018). *SPSS, SAS, and MATLAB programs for determining the number of components and factors using parallel analysis and velicer's MAP test*. University of Columbia.
<https://people.ok.ubc.ca/briocconn/nfactors/nfactors.html>
- Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical assessment, research & evaluation*, 9(11), 8.
- Osborne, J. W., & Costello, A. B. (2009). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pan-Pacific Management Review*, 12(2), 131-146.
- Pelowski, M., Markey, P. S., Forster, M., Gerger, G., & Leder, H. (2017). Move me, astonish me... delight my eyes and brain: The Vienna integrated model of top-down and bottom-up processes in art perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Physics of Life Reviews*, 21, 80-125.
<https://doi.org/10.1016/j.plrev.2017.02.003>
- Penn Center for Neuroaesthetics. (2019). *Resources*. University of Pennsylvania.
<https://neuroaesthetics.med.upenn.edu/research.html#resources>
- Penzes, I., Hooren, S. V., Dokter, D., & Hutschemaekers, G. (2018). How art therapists observe mental health using formal elements in art products: Structure and variation as indicators for imbalance and adaptability. *Frontiers in Psychology*, 9, 1611. <https://doi.org/10.3389/fpsyg.2018.01611>
- Pöge, A. (2011). Persönliche Codes bei Längsschnittuntersuchungen III: fehlertolerante Zuordnung unverschlüsselter und verschlüsselter selbstgenerierter Codes im empirischen Test [Personal codes in longitudinal designs III: Fault-tolerant assignment of unencrypted and crypted self-generated codes in the empirical test]. *Methoden, Daten, Analysen (mda)*, 5(1), 109-134.
- Raykov, T., & Little, T. D. (1999). A note on procrustean rotation in exploratory factor analysis: A computer intensive approach to goodness-of-fit evaluation. *Educational and Psychological Measurement*, 59(1), 47-57.
- Rhodes, M. (1961). An analysis of creativity. *The Phi Delta Kappan*, 42(7), 305-310.
- Said-Metwaly, S., Van den Noortgate, W., & Kyndt, E. (2017). Approaches to Measuring Creativity: A Systematic Literature Review. *Creativity: Theories – Research – Applications*, 4(2), 238-275. <https://doi.org/10.1515/ctra-2017-0013>
- Schneider, N. (2017). *Geschichte der Ästhetik von der Aufklärung bis zur Postmoderne: Eine paradigmatische Einführung [The history of aesthetics from enlightenment to post-modernity: A paradigmatic introduction]* (6 ed.). Reclam.
- Schoch, K., Gruber, H., & Ostermann, T. (2017). Measuring art: Methodical development of a quantitative rating instrument measuring pictorial expression (RizbA). *The Arts in Psychotherapy*, 55, 73-79. <https://doi.org/10.1016/j.aip.2017.04.014>
- Schönemann, P. H. (1966). A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31(1), 1-10.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological bulletin*, 86(2), 420.

- Silvia, P. J. (2007). An introduction to multilevel modeling for research on the psychology of art and creativity. *Empirical Studies of the Arts*, 25(1), 1-20.
- Specht, S. M. (2007). Successive contrast effects for judgments of abstraction in artwork following minimal pre-exposure. *Empirical Studies of the Arts*, 25(1), 63-70.
- Streb, J. H. (1984). Thoughts on phenomenology, education, and art. *Studies in Art Education*, 25(3), 159-166.
<https://doi.org/10.1080/00393541.1984.11650370>
- Stuhler-Bauer, A., & Elbing, U. (2003). Die phänomenologische Bilderfassung: Ein kunsttherapeutisches Instrument [The phenomenological picture survey: An art therapeutic instrument]. *Musik-, Tanz- und Kunsttherapie*, 14, 32-46.
<https://doi.org/10.1026/0933-6885.14.1.32>
- Stuhler-Bauer, A., & Elbing, U. (2004). Kritik standardisierter empirischer Bilderfassung aus kunsttherapeutisch-phänomenologischer Sicht: Die Nürtinger Beurteilungsskala und die Diagnostic Assessment of Psychiatric Art, deutsche Version [Review on standardized empirical picture analysis from an art therapeutic phenomenological view: The Nürtinger Rating Scale and the Diagnostic Assessment of Psychiatric Art, German version]. *Musik-, Tanz- und Kunsttherapie*, 15(1), 5-15. <https://doi.org/10.1026/0933-6885.15.1.5>
- Suhr, D. D. (2005). Principal Component Analysis vs. Exploratory Factor Analysis. *Proceedings of the thirtieth annual SAS® users group international conference*, 203, 30. <http://www2.sas.com/proceedings/sugi30/203-30.pdf>
- Suhr, D. D. (2006). Exploratory or confirmatory factor analysis? *Statistics and Data Analysis*, 31.
- Tansella, C. (2007). The long and winding journey of Outsider Art. An historical perspective. *Epidemiology and Psychiatric Sciences*, 16(2), 133-138.
- Thyme, K. E., Wiberg, B., Lundman, B., & Graneheim, U. H. (2013). Qualitative content analysis in art psychotherapy research: Concepts, procedures, and measures to reveal the latent meaning in pictures and the words attached to the pictures. *The Arts in Psychotherapy*, 40(1), 101-107. <https://doi.org/10.1016/j.aip.2012.11.007>
- Tinio, P. P. (2013). From artistic creation to aesthetic reception: The mirror model of art. *Psychology of Aesthetics, Creativity, and the Arts*, 7(3), 265-275. <https://doi.org/10.1037/a0030872>
- Tröndle, M., & Tschacher, W. (2012). The physiology of phenomenology: The effects of artworks. *Empirical Studies of the Arts*, 30(1), 75-113.
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1-10.
- Uzelac, M. (1998). Art and phenomenology in Edmund Husserl. *Axiomathes*, 9(1), 7-26.
- van Buren, B., Bromberger, B., Potts, D., Miller, B., & Chatterjee, A. (2013). Changes in painting styles of two artists with Alzheimer's disease. *Psychology of Aesthetics, Creativity, and the Arts*, 7(1), 89. <https://doi.org/10.1037/a0029332>
- Vollmar, K. (2008). *Die faszinierende Welt der Farben: ein Glossar von A-Z [The fascinating world of colors: A glossary of A-Z]*. ars momentum.
- Walter, S., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in medicine*, 17(1), 101-110.
- Wiesing, L. (2005). Methoden der Bildwissenschaft [Methods of visual culture]. In K. Sachs-Hombach (Ed.), *Bildwissenschaft zwischen Reflexion und Anwendung [Visual culture between reflexion and application]*. Von Halem.
- Wünsch, K. L. (2016). *Comparing two groups' factor structures: Pearson r and the coefficient of congruence*.
<http://core.ecu.edu/psyc/wuenschk/MV/FA/FactorStructure-TwoGroups.docx>
- Zaidel, D. W., Nadal, M., Flexas, A., & Munar, E. (2013). An evolutionary approach to art and aesthetic experience. *Psychology of Aesthetics, Creativity, and the Arts*, 7(1), 100. <https://doi.org/10.1037/a0028797>
- Zumbo, B. D., Sireci, S. G., & Hambleton, R. K. (2003). Re-visiting exploratory methods for construct comparability: Is there something to be gained from the ways of old. Annual Meeting of the National Council for Measurement in Education (NCME), Chicago, Illinois,

Appendix

Abbreviations

CFA Confirmatory factor analysis

PCA Principal component analysis

PHP Hypertext Preprocessor

RizbA Ratinginstrument für zweidimensionale bildnerische Arbeiten
(rating instrument for two-dimensional pictorial works)

Tables

Table 2. Tucker's coefficient of congruence (T1 and T2)

Time of measurement	Factor	Tucker's coefficient of congruence
T1	factor ₁₁	1.00
	factor ₂₂	0.99
	factor ₃₄	1.00
	factor ₄₃	-1.00
	factor ₅₅	0.98
	factor ₆₆	0.99
	factor ₇₇	1.00
	factor ₈₈	1.00
T2	factor ₁₁	1.00
	factor ₂₂	0.96
	factor ₃₃	-0.96
	factor ₄₈	0.82
	factor ₅₄	0.99
	factor ₆₅	0.96
	factor ₇₆	-0.98
	factor ₈₇	0.99

Table 3. *Rizba*: Translated* items and item values (T1+T2)

No.	Item*		<i>M</i>	<i>SD</i>	<i>s</i> _i ²	<i>p</i> _i	<i>η</i> _p ²	<i>ICC</i>
1	The picture includes graphic elements	<i>T1</i>	2.88	1.14	1.31	0.58	0.52	0.96
		<i>T2</i>	2.87	1.58	2.50	0.57	0.54	0.92
2	The picture includes pictorial elements	<i>T1</i>	3.05	0.96	0.92	0.61	0.41	0.94
		<i>T2</i>	2.98	1.50	2.24	0.59	0.48	0.94
3	The manner of representation is concrete	<i>T1</i>	2.15	1.29	1.67	0.43	0.63	0.98
		<i>T2</i>	2.12	1.64	2.69	0.42	0.69	0.98
4	The manner of representation is abstract	<i>T1</i>	3.07	1.09	1.19	0.61	0.52	0.96
		<i>T2</i>	3.01	1.54	2.38	0.60	0.58	0.95
5	The color application is pastose	<i>T1</i>	1.67	0.79	0.63	0.33	0.31	0.91
		<i>T2</i>	1.73	1.45	2.11	0.34	0.39	0.94
6	The predominant coloring is vibrant	<i>T1</i>	2.91	1.04	1.08	0.58	0.52	0.96
		<i>T2</i>	2.97	1.42	2.03	0.59	0.62	0.52
7	In the picture primary colors are prevalent	<i>T1</i>	2.34	0.72	0.52	0.47	0.28	0.87
		<i>T2</i>	2.34	1.37	1.87	0.47	0.34	0.92
8	In the picture mixed colors (secondary colors) are prevalent	<i>T1</i>	2.48	0.85	0.72	0.50	0.36	0.92
		<i>T2</i>	2.49	1.39	1.92	0.50	0.41	0.96
9	In the picture there are complementary contrasts	<i>T1</i>	2.57	1.16	1.34	0.51	0.51	0.96
		<i>T2</i>	2.54	1.59	2.54	0.51	0.54	0.95
10	In the picture organic shapes are prevalent	<i>T1</i>	2.60	0.80	0.65	0.52	0.34	0.92
		<i>T2</i>	2.67	1.35	1.82	0.53	0.39	0.63
11	In the picture geometric shapes are prevalent	<i>T1</i>	1.97	0.99	0.98	0.39	0.46	0.95
		<i>T2</i>	1.97	1.42	2.01	0.39	0.50	0.92
12	The layout of the line is predominantly curved	<i>T1</i>	2.72	1.04	1.07	0.55	0.50	0.96
		<i>T2</i>	2.76	1.45	2.10	0.55	0.55	0.92
13	The layout of the line is predominantly angled	<i>T1</i>	1.52	0.95	0.90	0.30	0.49	0.96
		<i>T2</i>	1.51	1.32	1.76	0.31	0.51	0.63
14	The picture includes unworked areas	<i>T1</i>	2.64	1.31	1.72	0.53	0.51	0.96
		<i>T2</i>	2.75	1.82	3.32	0.55	0.58	0.81
15	The picture appears to be deep	<i>T1</i>	1.62	0.66	0.44	0.32	0.29	0.87
		<i>T2</i>	1.58	1.20	1.44	0.31	0.32	0.84
16	The picture is perspectival	<i>T1</i>	1.64	0.79	0.62	0.33	0.34	0.92
		<i>T2</i>	1.64	1.32	1.73	0.33	0.37	0.80
17	The picture is without perspective (aperspectival)	<i>T1</i>	0.30	0.84	0.71	0.60	0.31	0.91
		<i>T2</i>	3.00	1.50	2.26	0.60	0.35	0.88
18	The picture is restless	<i>T1</i>	2.61	0.79	0.62	0.52	0.34	0.91
		<i>T2</i>	2.61	1.33	1.76	0.52	0.40	0.89
19	The picture is wild	<i>T1</i>	2.39	0.90	0.81	0.48	0.43	0.94
		<i>T2</i>	2.41	1.36	1.86	0.48	0.48	0.88

20	The global composition is laid out vertically	<i>T1</i>	2.04	1.12	1.26	0.41	0.48	0.95
		<i>T2</i>	2.04	1.60	2.55	0.41	0.55	0.70
21	The global composition is laid out horizontally	<i>T1</i>	2.07	1.04	1.07	0.41	0.43	0.93
		<i>T2</i>	2.02	1.58	2.51	0.41	0.47	0.65
22	The global composition is laid out diagonally	<i>T1</i>	1.79	1.04	1.07	0.36	0.47	0.95
		<i>T2</i>	1.82	1.53	2.35	0.36	0.55	0.87
23	The global composition is laid out area-wide without a main subject (All-Over-Structure)	<i>T1</i>	2.21	1.14	1.29	0.44	0.45	0.96
		<i>T2</i>	2.21	1.70	2.88	0.45	0.48	0.97
24	The picture appears to be diffuse	<i>T1</i>	2.31	0.70	0.50	0.46	0.29	0.89
		<i>T2</i>	2.24	1.30	1.69	0.45	0.36	0.81
25	The picture appears to be precise, accurate	<i>T1</i>	1.95	0.63	0.40	0.39	0.25	0.88
		<i>T2</i>	1.99	1.27	1.60	0.40	0.37	0.82
26	The picture appears to be harmonic	<i>T1</i>	2.61	0.49	0.24	0.52	0.18	0.81
		<i>T2</i>	2.57	1.15	1.31	0.51	0.25	0.84
<i>M</i>		<i>T1</i>	2.23	0.93	0.91	0.47	0.41	0.93
		<i>T2</i>	2.34	1.45	2.12	0.47	0.46	0.84
<i>SD</i>		<i>T1</i>	0.60	0.21	0.38	0.09	0.11	0.04
		<i>T2</i>	0.47	0.16	0.46	0.09	0.11	0.12
<i>Min</i>		<i>T1</i>	0.30	0.49	0.24	0.30	0.18	0.81
		<i>T2</i>	1.51	1.15	1.31	0.31	0.25	0.52
<i>Max</i>		<i>T1</i>	3.07	1.31	1.72	0.61	0.63	0.98
		<i>T2</i>	3.01	1.82	3.32	0.60	0.69	0.98

* = original German version see additional material, *T1* = test, *T2* = retest, *M* = mean (0 = *strongly disagree*, 5 = *strongly agree*), *SD* = item standard deviation, *Min* = minimum, *Max* = maximum, s_i^2 = item variance, p_i = item difficulty, η_p^2 = partial eta-squared effect size estimator, *ICC* = intraclass correlation coefficient average measure

Table 4. *Rizb4*: Rotated component matrix, factor loadings (T1)

Prospective factor label	No.	Item	Components							
			1	2	3	4	5	6	7	8
<i>Picture effect</i>	18	The picture is restless	.843							
	19	The picture is wild	.747							
	24	The picture appears to be diffuse	.737							
	26	The picture appears to be harmonic	-.679							
	25	The picture appears to be precise, accurate	-.570							
<i>Spatiality</i>	16	The picture is perspectival		.841						
	15	The picture appears to be deep		.769						
	17	The picture is without perspective (aperspectival)		-.763						
<i>Shaping</i>	13	The layout of the line is predominantly angled			-.779					
	12	The layout of the line is predominantly curved			.712					
	10	In the picture organic shapes are prevalent			.694					
	11	In the picture geometric shapes are prevalent			-.647					
<i>Pictorial elements (drawing vs. painting)</i>	1	The picture includes graphic elements				.716				
	2	The picture includes pictorial elements				-.588		.412		
	14	The picture includes unworked areas				.570				
	5	The color application is pastose				-.520				
<i>Representation</i>	3	The manner of representation is concrete					-.701			
	4	The manner of representation is abstract					.695			
	23	The global composition is laid out area-wide without a main subject (All-Over-Structure)					.599			
	22	The global composition is laid out diagonally					.573			
<i>Color intensity</i>	9	In the picture there are complementary contrasts						.710		
	6	The predominant coloring is vibrant						.634		

<i>Color mixture</i>	7	In the picture primary colors are prevalent							-.828	
	8	In the picture mixed colors (secondary colors) are prevalent							.815	
<i>Composition</i>	21	The global composition is laid out horizontally								-.827
	20	The global composition is laid out vertically								.810

Method of extraction: PCA; Method of rotation: Varimax, Kaiser normalization; Values < .40 suppressed

Table 5. *Rizb4*: Rotated component matrix, factor loadings (T2)

Prospective factor label	No.	Item	Components							
			1	2	3	4	5	6	7	8
<i>Picture effect</i>	18	The picture is restless	.848							
	19	The picture is wild	.776							
	24	The picture appears to be diffuse	.733							
	26	The picture appears to be harmonic	-.689							
	25	The picture appears to be precise, accurate	-.609							
<i>Spatiality</i>	16	The picture is perspectival		.859						
	17	The picture is without perspective (aperspectival)		-.804						
	15	The picture appears to be deep		.734						
	3	The manner of representation is concrete		.506				-.503		
<i>Pictorial elements (drawing vs. painting)</i>	1	The picture includes graphic elements			-.741					
	14	The picture includes unworked areas			-.618					
	2	The picture includes pictorial elements			.608					
	5	The color application is pastose			.555					
	23	The global composition is laid out area-wide without a main subject (All-Over-Structure)			.423					
<i>Shaping</i>	13	The layout of the line is predominantly angled				-.746				

	12	The layout of the line is predominantly curved			.735			
	10	In the picture organic shapes are prevalent			.722			
	11	In the picture geometric shapes are prevalent			-.637			
<i>Color intensity</i>	9	In the picture there are complementary contrasts				.756		
	6	The predominant coloring is vibrant				.610		
<i>Representation</i>	22	The global composition is laid out diagonally					.727	
	4	The manner of representation is abstract		-.418			.534	
<i>Color mixture</i>	8	In the picture mixed colors (secondary colors) are prevalent						-.833
	7	In the picture primary colors are prevalent						.830
<i>Composition</i>	21	The global composition is laid out horizontally						-.838
	20	The global composition is laid out vertically						.802

Method of extraction: PCA; Method of rotation: Varimax, Kaiser normalization; Values < .40 suppressed

Additional Material

Open Methodology

A paper-pencil-version of the current original German version of *RizbA* questionnaire is freely available:

Schoch, K. (2020, April 24). *Ratinginstrument für zweidimensionale bildnerische Arbeiten (RizbA): Fragebogen mit Erläuterungen in deutscher Sprache* [Rating instrument for two-dimensional pictorial works (RizbA): Questionnaire with explanatory notes in German]. *Zenodo*. <https://doi.org/10.5281/zenodo.3765221>

Table 6. *RizbA*: Original German version

No.	Item
1	Das Bild enthält zeichnerische Elemente
2	Das Bild enthält malerische Elemente
3	Die Darstellungsweise ist gegenständlich
4	Die Darstellungsweise ist abstrakt
5	Der Farbauftrag ist pastos
6	Die vorherrschende Farbgebung ist leuchtend
7	Im Bild befinden sich vorwiegend reine Farben
8	Im Bild befinden sich vorwiegend Mischfarben (Sekundärfarben)
9	Im Bild sind Komplementärkontraste vorhanden
10	Im Bild enthaltene Formen sind vorwiegend organisch
11	Im Bild enthaltene Formen sind vorwiegend geometrisch
12	Die Linienführung verläuft vorwiegend gebogen
13	Die Linienführung verläuft vorwiegend eckig
14	Das Bild enthält unbearbeitete Flächen
15	Das Bild wirkt tief
16	Das Bild ist perspektivisch
17	Das Bild ist frei von Perspektive (aperspektivisch)
18	Das Bild ist unruhig
19	Das Bild ist wild
20	Die Gesamtkomposition ist senkrecht angelegt
21	Die Gesamtkomposition ist waagrecht angelegt
22	Die Gesamtkomposition ist diagonal angelegt
23	Die Gesamtkomposition ist flächendeckend ohne Hauptmotiv (All-Over-Structure)
24	Das Bild wirkt diffus
25	Das Bild wirkt präzise, exakt
26	Das Bild wirkt harmonisch

Open Data

Due to the participant data protection statement, the image material and data are not readily accessible. Both are available upon request to the corresponding author and are to be used for scientific purposes only.