Schoch, K. (2021). How to measure art: Three validations studies on the RizbA, a quantitative rating instrument to assess pictorial expression. In V. Huet & Kapitan, L. (Eds.), *International advances in art therapy research and practice: The emerging picture* (pp. 149-157). Cambridge Scholars Publishing.

1. How to Measure Art: Three Validation Studies on the RizbA, a Quantitative Rating Instrument to Assess Pictorial Expression

Kerstin Schoch

Why Measure Art?

One of art therapy's basic assumptions is that inner representations are reflected in the pictorial expression. These representations comprise a wide variety of psychological and clinical constructs. Few correlations between visual art and psyche have been researched quantitatively due to a lack of reliable measures for artworks. However, developing greater empirical evidence between psychological variables and artworks could have a positive impact on the credibility of art therapy. If artistic products and processes are found to mirror a person's states and traits, then art therapy could become a powerful tool for accessing and working with them.

As described in this chapter, we developed and validated a rating instrument for two-dimensional pictorial works to address the need for a quantitative measure for artworks. The questionnaire measures pictorial expression in terms of a formal image analysis that includes such aspects as representation, color, shaping, spatiality, motion, and composition. The chapter provides a brief overview of the instrument, three validation studies conducted thus far, and statistical quality criteria.

Previous Approaches to Measuring Art

Although a large proportion of art therapy research is grounded in qualitative methods, a growing number of studies include quantitative data as equally important for refining theory and providing generalization. Qualitative content analysis for systematizing pictorial works (Thyme et al., 2013) and phenomenological picture analysis (Stuhler-Bauer & Elbing, 2003) are two examples that qualitatively assess pictorial works. Among attempts to quantify images, most are specifically designed for therapeutic or clinical use, such as the Diagnostic Assessment of Psychiatric Art (Hacking et al., 1996), the Nürtinger Rating Scale (Elbing & Hacking, 2001*)* and the Systematic Picture Analysis (Gruber et al., 2002). Other instruments, such as the Diagnostic Drawing Series (Cohen & Mills, 2015), the Formal Elements Art Therapy Scale (Gantt, 2016) and the Bird's Nest Drawing (Yoon et al., 2020) have been developed for research and assessment beyond clinical applications, but generally involve specific drawing tasks and therefore are not applicable across all types of images.

Beyond classic paper and pencil tests, there are digital approaches that aim to quantify image analysis, including computerized assessment of art-based instruments (Mattson, 2010), existing measures combined with image analysis software (Mattson, 2009, 2011), computer-based systems that rate basic color-related aspects (Kim et al., 2007), and the use of computational network science for assessing global properties (Hayn-Leichsenring et al., 2020).

Other disciplines are involved with measuring art: art psychology and its subdomains empirical aesthetics and neuroaesthetics. These fields of research have the advantage of evidence based in statistics and psychometrics. However, art has not been operationalized or measured in the same detailed way as psychological variables (e.g., personality, intelligence). This is probably because the contemporary field of art psychology is primarily interested in correlates of art rather than in the art

image itself (e.g., Minissale, 2013). However, one tool stands out: the Assessment of Art Attributes (Chatterjee et al., 2010). Based on a neuropsychological perspective, the tool aims to quantify artworks with 14 items, including both formal-perceptional and content-representational attributes. The results suggest medium to high agreement between raters on nonparametric measures of correlation and a previously conducted rater training with training slides of the attributes. Other statistical quality criteria are not reported. The assessment has been tested on a small sample of prominent paintings from the Western canon, but not on other pictorial material and it has not been further validated on larger samples.

Quality Criteria in Measuring Art

All the above-mentioned instruments are useful for analyzing, documenting, or understanding art images. However, none meet all criteria for measuring pictorial works in a universally applicable, reliably measurable, and validated way. The criteria, which formed the basis on which we developed the new instrument (Schoch et al., 2017), are:

- **Universal applicability** to all types of images (e.g., amateurs' works vs. professional art, clinical vs. nonclinical settings, contemporary vs. historical artworks)
- **Quantitative methodology** (beyond descriptive statistics that are limited to describing features) that allows a calculation of inferential statistics to test hypotheses and a derivation of estimates
- **Reliability** in terms of psychometric criteria (e.g., item difficulty, capacity of differentiation between images, test-retest reliability, inter-rater reliability, factors solution)
- **Validity** based on studies with large, representative sample sizes

To summarize, although there are several instruments that describe art systematically, there is a need for quantitative, standardized, and reliable measures that allow inferential statistics and are validated with and applicable to all types of two-dimensional pictorial works.

The Rating Instrument for Two-Dimensional Pictorial Works (RizbA)

The RizbA (Ratinginstrument für zweidimensionale bildnerische Arbeiten) is a questionnaire to assess pictorial expression, which is defined as artistic creation in the form of a picture. It deploys the concept of a formal picture analysis (Bauer, 1996) by focusing on such formal aspects as representation, color, shape, spatiality, motion, and composition, as generally described in art literature. This approach is rooted in the tradition of phenomenological picture analysis that seeks to overcome accidental judgment, preconception, and association (Streb, 1984). The relevant systemic level is the pictorial representation, which consciously excludes the picture's references to knowledge, associations, emotions, and the like. Instead, the RizbA focuses on visual presentation while leaving out colonializing projections of unintentional identification with the objects (Marotzki & Stoetzer, 2006), immediate associations, and interpretation of formal elements. It is limited to a detailed but classical conception of images without taking into account the creation process. The test does not judge the creator's achievement or mastery and is distinct from aesthetic appreciation. It is neither evaluative nor interpretative or projective, but rather aims for a value-free description of the picture's formal elements. A rater training based on sample images for certain characteristics is deliberately not conducted in order to avoid a manipulation of judgment.

The RizbA questionnaire consists of 26 items (Table 8.1), currently in the German language. The rating is based on a 6-point Likert scale, which is discretely scaled and verbally anchored inshades of agreement (0 = *strongly disagree*, 5 = *strongly agree*). Raters receive a brief instruction to rate the image presented using the questionnaire. They are asked to focus on the predominant overall expression of a picture and not single details, while assured that there is no right or wrong while rating.

Table 8.1

RizbA Items: English Translation and Original German Version

Item No.	English translation	Original version
1	The picture includes graphic elements	Das Bild enthält zeichnerische Elemente
2	The picture includes pictorial elements	Das Bild enthält malerische Elemente
3	The manner of representation is concrete	Die Darstellungsweise ist gegenständlich
4	The manner of representation is abstract	Die Darstellungsweise ist abstrakt
5	The color application is impasto[1]	Der Farbauftrag ist pastos
6	The predominant coloring is vibrant	Die vorherrschende Farbgebung ist leuchtend
7	In the picture primary colors are prevalent	Im Bild befinden sich vorwiegend reine Farben
8	In the picture mixed colors (secondary colors) are prevalent	Im Bild befinden sich vorwiegend Mischfarben (Sekundärfarben)
9	In the picture there are complementary contrasts	Im Bild sind Komplementärkontraste vorhanden
10	In the picture organic shapes are prevalent	Im Bild enthaltene Formen sind vorwiegend organisch
11	In the picture geometric shapes are prevalent	Im Bild enthaltene Formen sind vorwiegend geometrisch
12	The layout of the line is predominantly curved	Die Linienführung verläuft vorwiegend gebogen
13	The layout of the line is predominantly angled	Die Linienführung verläuft vorwiegend eckig
14	The picture includes unworked areas	Das Bild enthält unbearbeitete Flächen
15	The picture appears to be deep	Das Bild wirkt tief
16	The picture is perspectival[2]	Das Bild ist perspektivisch
17	The picture is without perspective (aperspectival)	Das Bild ist frei von Perspektive (aperspektivisch)
18	The picture is restless[3]	Das Bild ist unruhig
19	The picture is wild[4]	Das Bild ist wild
20	The global composition is laid out vertically	Die Gesamtkomposition ist senkrecht angelegt
21	The global composition is laid out horizontally	Die Gesamtkomposition ist waagrecht angelegt
22	The global composition is laid out diagonally	Die Gesamtkomposition ist diagonal angelegt
23	The global composition is laid out area-wide without a main subject (all-over-structure)	Die Gesamtkomposition ist flächendeckend ohne Hauptmotiv (All-Over-Structure)
24	The picture appears to be diffuse	Das Bild wirkt diffus
25	The picture appears to be precise, accurate	Das Bild wirkt präzise, exakt
26	The picture appears to be harmonic	Das Bild wirkt harmonisch

[1] impasto = pasty, thick layers of paint
[2] perspectival = enhances the impression of three-dimensionality on a plain surface
[3] The image suggests a restless effect
[4] The image suggests a wild effect

Method

For empirical testing and validation of the scale, three randomized, online studies in a test-retest design were conducted (Table 8.2). In these studies, experts in picture analysis rated samples of images consisting of pictorial works by nonprofessionals and professional contemporary artists. All artworks were two-dimensional and included drawings, paintings, collages, and mixed techniques. After data collection, statistical quality criteria (i.e., item difficulty, capacity of differentiation, test-retest reliability, and intraclass correlation) were calculated for each item and for the overall test. Principal component analysis and indices of factor similarity also were computed. Detailed study designs, implementations, and statistical procedures can be found in the literature cited on the table when published.

Table 8.2

Validation Studies: Pictorial Material and Raters

	Study 1 (Schoch, Gruber, & Ostermann, 2017)	Study 2 (Schoch & Ostermann, 2020b)	Study 3 (Schoch & Ostermann, 2020a)
Images	Amateurs' pictorial works $N = 12$	Amateurs' pictorial works $N = 294$	Contemporary artworks $N = 318$
Raters	Art therapists $N_{T1} = 12$ $N_{T2} = 8$	Diverse experts (e.g., art therapists, art pedagogues, art historians, artists, designers, restorers) $N_{T1} = 880$ $N_{T2} = 475$	Diverse experts (e.g., art therapists, art pedagogues, art historians, artists, designers, restorers) $N_{T1} = 506$ $N_{T2} = 238$

Note. N = sample size, $T1$ = test, $T2$ = retest.

Results

The current test version was found to yield a medium to high capacity of differentiation between pictorial works in the three samples tested, along with high inter-rater reliability. The test-retest reliability also was found to be highly reliable (Table 8.3). In Study 1 the principal component analyses suggested a four-factors solution, although only exploratory and lacking representativeness of the image sample. Because the other two studies have larger samples, a more plausible eight-factors structure is suggested, consistent across studies. Prospective factor labels might be *picture effect*, *spatiality*, *shaping*, *pictorial elements* (*drawing vs. painting*), *representation*, *color intensity*, *color mixture*, and *composition*. Detailed results can be found in the literature cited on the tables.

Table 8.3

Statistical Quality Criteria: Overall Test

	Time of measurement	Study 1	Study 2	Study 3
Capacity for differentiation between images, $\eta\eta_p^2$	*T1*	.90	.28	.31
	T2	.77	.33	.40
Test-retest reliability, *r*	-	.92	.93	.86
Factors (PCA)	-	4	8	8
Inter-rater reliability, *ICC*	*T1*	.53	.81	.86
	T2	.92	.84	.73

Note. T1 = test, *T2* = retest, η_p^2 = partial eta squared effect size estimator, *r* = *correlation*, *PCA* = principal component analysis, ICC = intra-class correlation coefficient.

Discussion

These results suggest that the scale can be generalized and applied to both nonprofessional and contemporary pictorial art. As a methodically sound, quantitative instrument the meets all quality criteria mentioned above, the RizbA opens up new perspectives for practice and research.

Implications

By being a widely applicable tool, RizbA creates a methodological foundation in measurement for art therapeutic practice and future research. In practice, the questionnaire may be used as a reliable tool for structured documentation in individual and institutional reporting systems, particularly when tracking individual processes through art production. In addition the instrument can support the therapist's routine questioning, reflecting, and objectifying of their own perceptions.

As fundamental research, the RizbA holds promise in investigating hypotheses of correlations between pictorial expression and inner representations (e.g., personality, neurodiversity, resources, or conflicts). A quantitative display of these correlations would imply that pictorial expression is linked to internal cognitive and emotional processes. This potential would acknowledge art therapy as a powerful way of working with such inner representations.

Moreover, it is also a useful tool in applied research. At present, mechanistic and efficacy studies in art therapy research consist of pre-posttest designs that investigate cognitive, emotional, social, and other psychological outcomes (see, e.g., Abbing et al., 2018; Kim, 2013; Maujean et al., 2014; Schouten et al., 2015; Slayton et al., 2010). The RizbA enables further incorporation of the artistic medium itself—not only descriptively but quantitatively by including inferential statistics. Thus the artistic medium may be included in empirical studies and do justice to the actual subject of study.

Limitations

Principal component analysis suggests an interpretable factor structure with eight components. However, not having conducted a confirmatory factor analysis, final conclusions cannot be drawn. An additional limitation is that the study samples of nonprofessionals' art images as well as the raters themselves were European in origin. In particular, the samples of these pictorial works were biased in terms of gender and ethnicity towards women and European culture, which limits the claimto universality. In contrast the sample of contemporary art imagery was mixed with respect to

regions of origin, but also contained a bias towards *white*, cisgender, male artists. Because visual expression as well as its perception can greatly differ depending on the cultural context (Cattaneo, 1994), further studies should also address this aspect within a critical reflection on Eurocentrism (Mosquera, 1992) and paternalistic structures.

As stated earlier, the RizbA only captures one part of the big picture: that of the formal pictorial expression of two-dimensional works. In doing so, it provides only a glimpse of the entirety of variables relevant to art therapy. It leaves aside the choice of material, the creation process, reflections on the picture, the relationship between client and therapist, dynamics within a group, and so on. Even while concentrating on picture analysis, there are many more levels to analyze, evaluate, and understand, such as heuristics, motives, and narrations behind the images. To comprehensively analyze art therapy sessions, we need a battery of different instruments and research approaches to complete the picture.

Future Research

The next obligatory methodical step is to conduct a confirmatory factor analysis, which will help verify the factor structure of the observed variables. Therefore, we have now conducted a fourth validation study using new image material and computed an evolved theoretical model of the construct pictorial expression. We will then need to empirically validate an English translation of the questionnaire to make it accessible to more professionals. To make it applicable for users beyond art therapists and artists, we recently developed and validated a manual that explains the art vocabulary used (Jerusalem, 2020). The manual offers guidance to non-art experts, such as psychologists, pedagogues, and scientists, to apply the RizbA. Additionally, we are working on a machine learning approach that makes it possible to have artificial intelligence raters process the pictorial material. This will be of use for scientific studies where large amounts of image data are gathered. In terms of test validation and examination of convergent and divergent validity, further validation studies are planned. These will compare the RizbA to related but distinct scales, such as the Diagnostic Drawing Series (Cohen & Mills, 2015) and the Assessment of Art Attributes (Chatterjee et al., 2010).

For a generalization to other types of two-dimensional pictorial works and a better understanding of the factor structure, further validation studies on different image samples are needed (e.g., pictures by children and adolescents, non-handmade techniques like photography or printing techniques). Several pilot studies have been implemented on specific samples that compare pictorial expression between clinical subgroups, such as clients with chronic pain (Janßen, 2018) and recurrent depressive disorder (Epstein, 2019), and healthy control groups. Initial results imply differences between groups, suggesting correlations between artworks and clinical constructs.

Conclusion

The RizbA is a pioneering and transdisciplinary means to bridge art science and psychometric research methods, resulting in a quantitative assessment of pictorial expression. As a valid and reliable measurement tool with a broad range of applications, it fosters a more inherently art therapeutic research methodology.

Open Science

Study 3 described in this chapter was funded by the Open Science Fellows Program by Wikimedia Germany, Stifterverband, and Volkswagen Foundation. The current version of the RizbA questionnaire is freely available via https://zenodo.org/record/3765221#.YDxQGy1h3jA. Articles, materials, data, and syntax are freely available and can be accessed via www.kunsthochzwei.com

References

Abbing, A., Ponstein, A., van Hooren, S., de Sonneville, L., Swaab, H., & Baars, E. (2018). The effectiveness of art therapy for anxiety in adults: A systematic review of randomised and non-randomised controlled trials. *PloS ONE, 13*(12), e0208716. https://doi.org/10.1371/journal.pone.0208716

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* (Vol. 6, pp. 157–173). Reimer.

Cattaneo, M. (1994). Addressing culture and values in the training of art therapists. *Art Therapy: Journal of the American Art Therapy Association, 11*(3), 184–186. https://doi/10.1080/07421656.1994.10759081

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

Cohen, B. M., & Mills, A. (2015). The Diagnostic Drawing Series (DDS) at thirty: Art therapy assessment and research. In D. E. Gussak & M. L. Rosal (Eds.), *The Wiley handbook of art therapy* (pp. 558–568). John Wiley & Sons.

Elbing, U., & Hacking, S. (2001). Nürtinger Beurteilungsskala und Diagnostic Assessment of Psychiatric Art: Neue Wege zur Evaluation der Bilder von Kunsttherapie-Patienten [Nürtinger 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/10.1026//0933-6885.12.3.133

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. https://zenodo.org/record/3365921#.X4JWoS2ZPjA

Gantt, L. (2016). The Formal Elements Art Therapy Scale (FEATS). In D. Gussak & M. Rosal (Eds.), *The Wiley handbook of art therapy* (pp. 569–578). Wiley-Blackwell.

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

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

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]. University of Applied Sciences and Arts in Ottersberg. 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://doi.org/10.17605/OSF.IO/FVJZM

Kim, S., Bae, J., & Lee, Y. (2007). A computer system to rate the color-related formal elements inart therapy assessments. *The Arts in Psychotherapy, 34*(3), 223–237. https://doi.org/10.1016/j.aip.2007.02.002

Kim, S. K. (2013). A randomized, controlled study of the effects of art therapy on older Korean-Americans' healthy aging. *The Arts in Psychotherapy, 40*(1), 158–164. https://doi.org/10.1016/j.aip.2012.11.002

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.* VS Verlag für Sozialwissenschaften.

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

Maujean, A., Pepping, C. A., & Kendall, E. (2014). A systematic review of randomized controlled studies of art therapy. *Art Therapy: Journal of the American Art Therapy Association, 31*(1), 37–44. https:/doi.org/10.1080/07421656.2014.873696

Minissale, G. (2013). *The psychology of contemporary art*. Cambridge University Press. 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

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.2017.04.01499

Schoch, K., & Ostermann, T. (2020a). Giving the art greater weight in art psychology: RizbA, a quantitative questionnaire for formal picture analysis. [Manuscript submitted for publication].

Schoch, K., & Ostermann, T. (2020b). Psychometrics of art: Validation of RizbA, a quantitative rating instrument for pictorial expression. [Manuscript submitted for publication].

Schouten, K. A., de Niet, G. J., Knipscheer, J. W., Kleber, R. J., & Hutschemaekers, G. J. M. (2015).The effectiveness of art therapy in the treatment of traumatized adults: A systematic review onart therapy and trauma. *Trauma, Violence, and Abuse, 16*(2), 220–228. https://doi.org/10.1177/1524838014555032

Slayton, S. C., D'Archer, J., & Kaplan, F. (2010). Outcome studies on the efficacy of art therapy: Areview of findings. *Art Therapy: Journal of the American Art Therapy Association, 27*(3), 108– 118. https://doi.org/10.1080/07421656.2010.10129660

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 therapeuticinstrument]. *Musik-, Tanz- und Kunsttherapie, 14*, 32-46. https://doi.org/10.1026//0933- 6885.14.1.32

Thyme, K. E., Wiberg, B., Lundman, B. & Hällgren Graneheim, U. (2013). Content

analysis in artpsychotherapy 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

Yoon, J. Y., Betts, D., & Holttum, S. (2020). The Bird's Nest Drawing and accompanying stories inthe assessment of attachment security. *International Journal of Art Therapy, 25*(2), 76–87. https://doi.org/10.1080/17454832.2019.1697306