# Measuring art:
# Methodical development of a quantitative rating instrument measuring pictorial expression (RizbA)

**Kerstin Schoch[a,b], Thomas Ostermann[b], Harald Gruber[c]**

[a] *Institute for Art Therapy and Research, University of Applied Sciences and Arts Ottersberg, Germany*
[b] *Department of Psychology and Psychotherapy, Faculty of Health, Witten/Herdecke University, Germany*
[c] *Department of Art Therapy, Faculty of Human Sciences and Social Sciences, Alanus University of Arts and Social Sciences, Germany*

**Abstract.** Art therapy assumes that art work is related to differential constructs of the artist. Empirically, this hypothesis has not been proven yet because quantitative methods are rare. The Rating Instrument for two-dimensional Pictorial Work (RizbA) is designed to address this issue. The construct – *pictorial expression* – is theoretically defined by seven content areas (*representation*, *color*, *shape*, *space*, *motion*, *composition*, *expression*), which combined create the overall construct. Test development is based on art historical and art therapeutic theories and supported empirically. Two online studies are conducted using a sample of nine pictures, which are rated by experts ($n_1$ = 12, $n_2$ = 8). In the first study, based on psychometric characteristics, an item pool of 113 items is examined and a preliminary test version is developed. The second study examines quality criteria of the preliminary version. For both studies, factor analyses are computed.

The preliminary version includes 26 items. Its ability for differentiation between pictorial works ranges between .897 (T1) and .766 (T2), its inter-rater reliability between .525 (T1) and .917 (T2). Test-retest reliability is .919. PCA suggests a four-factors solution, which in large part is consistent across studies. As a reliable measurement RizbA opens new perspectives in fundamental art therapeutic and psychological research.

**Keywords:** pictorial expression, art therapy, differential psychology, test construction, quantitative methods

## Introduction

Though art therapy is a heterogeneous discipline, there is one fundamental assumption most experts share: it suggests that art work is related to the artist and his/ her inner representations. From a differential psychological perspective, this relation can be applied to a multitude of constructs, such as needs, motives, coping mechanisms, sense of self, self-efficacy, as well as social and emotional skills. Studies indicate correlations between personality and preferences in art reception (Gridley, 2013), but few studies examine if there is also a correlation between personality and preferences in the way of creating art. In the therapeutic context, art facilitates the display of differential constructs that cannot be expressed verbally and thereby provides another way of understanding. Empirically, these hypotheses have not yet been proven.

## Previous Research

The number of empirical studies about art therapy has increased over the years (Metzl, 2008) and indicates positive effects on specific variables such as mood valence (De Petrillo & Winner, 2005), mentalization (Franks & Whitaker, 2007), PTSD (Gantt & Tinnin, 2007), acute stress symptoms connected to PTSD (Chapman, Morabito, Ladakakos, Schreier & Knudson, 2001), and depression (Gussak, 2006). However, heterogeneous samples, designs and measures complicate the comparison of results (Slayton, D'Archer & Kaplan, 2010). Studies often consist of qualitative data and single case analyses (Maujean, Pepping & Kendall, 2014; Slayton, D'Archer & Kaplan, 2010) and do not allow statistical conclusions. In terms of quantitative research, art therapy is in its infancy (Eitel, Szkura, Pokorny & von Wietersheim, 2008).

There have been different approaches such as regarding artistic and psychiatric aspects of pictures (e.g. *Prinzhorn Collection*, *Cunningham Dax Collection*; Koh, 2014) or considering aesthetics and cognitive schemata (Wagner, Menninghaus, Hanich & Jacobsen, 2014). Still, standards of picture analysis hardly exist (Stuhler-Bauer & Elbing, 2003). Most tests are designed for clinical use, not art in general. Only some describe the construct exhaustively and even fewer fulfill quantitative criteria.

One common method of systematizing artwork in this field is qualitative content analysis such as described by Thyme, Wiberg, Lundman and Graneheim (2013). *Diagnostic Assessment of Psychiatric Art* (Hacking, Foreman, Belcher, 1996) fragments a picture into fields, measures aspects such as color and structure within these fields and results in mean scores. It does not include shape, space or composition. Application is defined clinically. *Nürtinger Rating Scale* (Eber, Müller, Bader und Baukus, 1998; Elbing & Hacking; 2001) is defined for therapeutic use, analyzes the picture in total and covers many aspects of classical picture analysis. It comes with a theory-based structure. Gruber, Frieling und Weis (2002) developed a *Systematic Picture Analysis* using discriminant analysis, which suggested differences between clients with different clinical diagnoses considering such as color effect or shape development (Petersen, Gruber, & Tüpker, 2011). *DokuPro* (Elbing & Hölzer, 2007; Elbing, Hölzer, Danner-Weinberger & von Wietersheim, 2009; Elbing, Neuwirth, Knöbel & Krilles-Mayr, 2010; Oster, Elbing, von Wietersheim & Hölzer, 2010) is an empirically approved documentation of art therapeutic processes, of which one scale addresses art work itself. *Diagnostic Drawing Series* (Mills, 2003; Machiodi, 2012) asks the client to draw three pictures using pastels following different tasks, which then are analyzed by an art interview following a manual. *Formal Elements Art Therapy Scale* (Gantt & Anderson, 2009) refers to global variables in two-dimensional art. Coming from a clinical perspective, its 14 scales aim at structural elements based on the graphic equivalent of different mental states.

A very different approach to objectification of subjective ratings are approaches based on digital technology (Mattson, 2010). Mattson (2009; 2011) uses public domain image analysis software to rate items of the *Formal Elements Art Therapy Scale* (Gantt & Anderson, 2009). Considering color one of the most important factors of a drawing, Kim, Bae and Lee (2007) use a computer-based system that automatically rates various basic color-related aspects.

Qualitatively, *Phenomenological picture analysis* (Stuhler-Bauer & Elbing, 2003) is used in art therapy and features an important theoretical position: the double function of the aesthetic state

serves as a method of gaining cognition as well as a therapeutic agent. Thus, art is result but also impulse for new processes. The method is descriptive, critical and heuristic, pursuing the observer's impartiality. It leads back to phenomenology, examining the subject in an Aristotelian way, in which the whole is more than the sum.

Although several instruments of art therapeutic assessment exist, there is a severe lack of standardized, non-projective and sufficiently reliable measures that objectively describe *pictorial expression* based on quantitative methods. This is where *RizbA* contributes.

## Theory

Construct definition is theory-based and focuses on *pictorial expression,* which is defined as artistic creation in the form of a picture. It involves several components that can be found in literature (e.g. Arnheim, 2000; Kandinsky, 1955; Meyer, 2011; Vollmar, 2008) and is reflected in an a priori designed meta structure of seven theoretical content areas (*representation*, *color*, *shape*, *space*, *motion*, *composition*, *expression*). Description is supposed to be exhaustive but non-redundant, objective, value-free and conducted in terms of a phenomenological analysis. It is neither evaluative nor interpretative and designed for rating two-dimensional manually created pictorial work (e.g. drawings, paintings). RizbA aims to be an instrument for experts. It can be used for investigating and measuring correlations of *pictorial expression* with many other constructs for which valid instruments exist, such as personality or clinical diagnoses. As such it is a valuable method for future art therapeutic research.

Defining the nomological network, the test is unendorsed for machined (e.g. photography) or three-dimensional work (e.g. sculpture). It does not judge mastery nor is it projective like drawing tests (e.g. House-Tree-Person test; Buck & Warren, 1992). The subject is to be distinguished from experimental aesthetics, in which the perceiver becomes the subject. Theoretical background overlaps with art psychology but differs substantially because the processes of the artist's or observer's psyche are not relevant in the rating itself.

## Materials and Method

Test construction is led by theory and thus is empirically-based but non-empirical in a traditional sense. The underlying measurement model is similar to formative models. *Pictorial expression* is not a one-dimensional latent variable, but rather a profile of several characteristics, which sum up to the whole construct. Creating a construct map is hardly possible. To obtain an exhaustive description, the item pool is deductively developed based on theories of art historical and art therapeutic picture analysis and supplemented by aspects which proved relevant in practical artistic and art therapeutic work. The instruments mentioned above are examined for suitable items as well. Seven content areas as mentioned above are proposed. They are to exhaustively cover the whole scope of classical picture analysis. They are not to be seen as independent factors but rather as systemization with overlaps. Items sharing the same item stem are combined in question blocks. The questionnaire is in German and uses a bipolar six-point Likert-scale, which is discretely scaled and verbally anchored in shades of agreement.

Data was collected in two online studies using *SoSci Survey* (Leiner, 2014). Participants were instructed not to rate single details but the pre-dominant overall expression of pictures. Instructions indicated there is no right or wrong but personal valuation. Demographic variables were collected to confirm the participant's state of expertise (e.g. university degree, university and working experience as an art therapist).

Since art therapy usually focuses on art created without concern for one's professional artistic identity or the application of proficient techniques, stimulus material consisted of nine two-dimensional pictorial works of amateurs (2 male, 7 female), who were between 24 and 62 years old ($M$ = 35.4, $SD$ = 12.73) and differed in level of education and occupation. To generate a sample, participants were asked to create a picture. The pictures were selected in such a way as to be heterogeneous and representative of all types of genres (painting vs. drawing; representational vs. abstract). Pictures were photographed and used as digital images.

Analysis focuses on two units. Pictorial works serve as unit for item difficulty, capacity of differentiation, inter-item correlations, test-retest reliability and factor analysis. Raters serve as unit for inter-rater reliability. Since the data includes three levels (items, pictures, raters), for each analysis the relevant levels are chosen and the others are averaged.

## Study 1

### *Sample and design*

The first study is designed to examine items, exclude insufficient ones and create a preliminary test version. A total of 126 art therapists with graduate degrees at institutions certified by the *German professional association of art and gestalt therapy* (DFKGT) were contacted via email. Each of them received a personalized link to the online survey whereupon 23 participated. Ten of them completed the whole survey, two more than half. These 12 raters (1 male, 11 female) were aged between 29 and 63 years ($M$ = 42.41, $SD$ = 11.77) with average work experience of 9.73 years ($SD$ = 8.15). The raters received their art therapy education at different institutions with different functional focuses. No rater training was conducted.

### *Procedure*

The survey successively presented the whole sample of pictorial work whereby each picture was to be rated by all raters using the item pool of 113 items. Order of pictures, areas of content and items within question blocks were randomized. Because of the large number of items, the option was given to fill out the survey in two stages.

### *Statistical Analysis*

Due to the formative model, statistical quality criteria such as corrected item-total correlation and Cronbach's α are not expedient. Instead, item difficulty, inter-item correlations, inter-rater reliability and differentiation capability are calculated for each item. Factor analyses are computed to examine the factors structure and the hypothesized formative model. Interval level of data is assumed. Analyses are conducted using SPSS and Excel.

As a theoretical criterion, it is stated that each area of content should be represented by three or more items. Psychological standards are used as a statistical criterion for item difficulty and inter-item correlation. Since no standards exist in this field, criteria for the items' inter-rater reliability and capacity for differentiation are determined a posteriori based on analysis of the empirical distribution.

## Study 2

### *Sample and design*

The goal of the second study is to examine the preliminary version's quality criteria. Via email, all former participants were sent a personalized link to the online survey. Eight of them participated and completed the whole survey, i.e. there were no missing data. Raters (1 male, 7 female) were aged between 33 and 65 years ($M$ = 46, $SD$ = 12.67). Their work experience averaged 11.25 years ($SD$ = 10.25).

### *Procedure*

The survey successively presented the same sample of pictorial work whereby each picture was to be rated by all raters using the 26 items which were identified as relevant in the first study. The order of pictures, areas of content and items within question blocks were randomized. There was no information given about filling out the survey in two stages.

### *Statistical analysis*

Overall inter-rater reliability, capacity of differentiation and test-retest reliability is calculated. Factor analyses are computed to examine the factor structure. Interval level of data is assumed. Analyses are conducted using SPSS and Excel.

## Results

### Study 1

Since two raters did not complete the whole survey, a few pictures were not rated by all raters. Based on the feedback raters gave, it is assumed that values are not missing systematically but because of a lack of time in consequence of the time-consuming first survey. Missing values are estimated using multiple imputation.

Exploratory factor analysis in terms of a PCA (principal factor analysis) followed by varimax rotation is computed including the whole item pool. Eight factors with eigenvalues above one are extracted. A scree plot suggests four factors. Compared to Horn's parallel analysis using O'Connor's syntax (2000; University of Columbia, 2014), three factors have larger eigenvalues than parallel components. Therefore, another PCA followed by varimax rotation is conducted extracting three factors accounting for 63.26 % of the common variance. Considering the items assignment to factors as well as a large amount of low and double loadings, the meanings of factors is hardly interpretable.

Dahl's difficulty index for interval-levelled data is computed. A priori, a cut-off point is determined, which excludes all items with an item difficulty lying beneath .10 and above .90. All items lie within this range thus no items are excluded.

For inter-item reliability *ICC*s (intra-class correlations), two-way random single measures with absolute agreement are computed. Cut-off points are determined a posteriori based on the empirical distribution. Since values differ between areas of content, different minimum limits of correlation coefficients ranging from .35 to .55 are determined to cover each area. A total of 80 items do not fulfil the specified criteria and are excluded.

Capacity of differentiation of the items is examined by item variance, item effect size and item set effect size. Cut-off points are determined based on the distribution. The minimum criterion for item variance is .10. Nine items lie beneath that threshold and are excluded. Item effect size

is computed using MANOVA, in which partial eta squared is interpreted as an effect size estimator. The lowest ten percent of the distribution are excluded, which applies to eleven items resulting in a cut-off point of .189. Thus, *shape* and *movement* now include fewer items than other areas. To examine if an item extension explains more variance, present item sets are compared to extended ones. Extended sets include one more item that would cover for more theoretical aspects. Partial eta squared is calculated for item sets using MANOVA. The extended set *shape* (Pillai's trace: $F(48, 594) = 7.74$, $p < .001$, $\eta^2 = .385$) explains as much variance as the present set (Pillai's trace: $F(40, 495) = 7.75$, $p < .001$, $\eta^2 = .385$). The extended set *movement* (Pillai's trace: $F(24, 297) = 4.46$, $p < .001$, $\eta^2 = .265$) explains less variance than the present set (Pillai's trace: $F(16, 198) = 5.87$, $p < .001$, $\eta^2 = .322$). Therefore, an extension is statistically not justified.

To examine inter-item correlations, Pearson product-moment correlations with two-tailed tests of significance are calculated. The a priori criterion indicates the exclusion of items correlating higher than .80 or lower than -.80 due to redundancy. In addition to statistical analysis, correlations are investigated based on theoretical aspects by differentiating whether correlations are based on measuring the same issue or are due to characteristics related in the sample. In the former case, the item with better statistical values with respect to an underrepresented area is kept. In the latter case, both items are kept. A total of 13 inter-item correlations are statistically and theoretically redundant. Within these, five couplets are already excluded due to other criteria. In three cases one of the items is already excluded. Eventually, three items are eliminated due to their inter-item correlations.

Starting with an item pool of 113 items, 87 are excluded. The remaining 26 items represent the preliminary version of RizbA (see Tab.1). To determine its overall capacity of differentiation a MANOVA is conducted and yields a partial eta squared of .897 (Pillai's trace: $F(112, 49) = 1.46$, $p = .068$, $\eta^2 = .770$). Inter-rater reliability was computed using *ICC*s two-way random

single measures with absolute agreement. Overall reliability yields .525 ($ICC_{unjust,random}$ = .525, $F(233, 2563) = 14.35$; $p < .001$).

PCA followed by varimax rotation is performed using the preliminary version. Six factors with eigenvalues above one are extracted. A scree plot suggests four factors. Compared to Horn's parallel analysis three factors have larger eigenvalues than parallel components. A fourth factor is legitimate due to the theoretical rationale and justified by marginal statistical data. Therefore, another PCA followed by varimax rotation is conducted, extracting four factors (see Tab.2) accounting for 84.22 % of the common variance.

**Study 2**

To determine overall capacity of differentiation a MANOVA is conducted using the preliminary version. It yields a partial eta squared of .766 (Pillai's trace: $F(208, 360) = 6.126$, $p < .001$, $\eta^2 = .366$). Inter-rater reliability was computed for the preliminary version using *ICC*s two-way random single measures with absolute agreement. Overall reliability yields .917 ($ICC_{unjust,random}$ = .917, $F(234, 1638) = 12$; $p < .001$). To examine test-retest reliability between both studies, Pearson product-moment correlations are calculated. Overall reliability yields a correlation of .919 ($r(232) = .919$, $p < .001$).

PCA followed by varimax rotation is conducted using preliminary version. Seven factors with eigenvalues above one are extracted. A scree plot suggests five factors. Compared to Horn's parallel analysis three factors have larger eigenvalues than parallel components. A fourth factor is legitimate due to the theoretical rationale and justified by marginal statistical data. Therefore, another PCA followed by varimax rotation is conducted extracting four factors (see Tab.3) accounting for 82.36 % of the common variance. Comparing both studies' four-factor structure and content, assignments of 20 items remain consistent and are interpretable. Six items differ in their assignment to factors.

## Discussion

### *Results*

The studies result in a preliminary version of RizbA consisting of 26 items with sufficient statistical values. Although varying between studies, capacity of differentiation and inter-rater reliability are promising – especially regarding inter-rater agreement, which often is barely satisfactory in this field (Eitel et. al., 2008). Addressing the factorability of data, multicollinearity among the variables is assumed to be given since there are several sizeable inter-item correlations. Due to a non positive define matrix, Bartlett's test for sphericity could not be conducted. This issue should be addressed in further studies. No outliers were excluded from the analysis. Test-retest reliability is relatively high even though the studies cannot be seen as a proper test-retest since they differ in design and number of items used. New samples are needed for further examination and the development of a final version.

Given that there is no explicit theory about the structure of statistical factors yet, PCA is used as an exploratory method of analysis as recommended by Brown (2009). Since the main purpose of the content areas is to gain an exhaustive item pool describing all facets of the construct, they are not expected to occur in the data. Aside from that a factor model based on data might exist. For this reason, varimax rotation is used to keep factors independent as recommended by Osborne and Costello (2009).

The factor analysis enforces the hypothesis of *pictorial expression* consisting of compounded aspects. It indicates not only single items but factors behind the data. Structure in large part remains consistent in both studies. As a first attempt at interpretation, the first factor (see Tab.3) can be summed up to painting versus drawing as well as description of color and color application. The second factor refers to representation and perspective. The third clearly reflects shape while the fourth is about movement. Still, the analyses must be seen as an exploratory

attempt. Further studies are needed to investigate if there is a constant solution. Then the a priori defined areas of content could be replaced by empirically supported factors.

***Limitations***

External validity was pursued by selecting preferably heterogeneous and representative samples of pictures and raters. However, nine pictures cannot depict every possible aspect. To examine whether results can be generalized, further samples are needed. Furthermore, correlations to other related measures have to be investigated.

Since quantitative research on this topic is relatively new, a conservative approach is hardly applicable. The scarcity of data allows only few and mostly unspecified a priori hypotheses and thus these studies have to be seen as an exploratory attempt.

Statistical analysis showed varying item quality. Areas such as *shape* owe a larger number of high-quality items whereas in *expression* items are rather low regarding inter-rater reliability and other statistical criteria. These differences might result from different capabilities of objective rating.

Finally, pictorial work cannot be pictured only by quantitative methods, especially in an art therapeutic setting where qualitative criteria must be considered as well. Phenomenological picture analysis with aspects such as *spontaneous expression* offers a reasonable extension to quantitative description. Both quantitative and qualitative methods have to be used complementary.

***Future research***

Quantitative research in art therapy is in its infancy, even more with regard to fundamental research. As a result, there is a lack of methodical standards, also in picture analysis. By developing a quantitative instrument for measuring *pictorial expression,* this work provides a new method for fundamental psychological and art therapeutic research. The next step is to replicate results and examine the factor structure, including a proper test-retest design with new

samples of raters and pictures. Samples can be extended to pictorial work of children or artists and other raters (e.g. artists, art historians, psychologists).

## Tables

*Table 1.* RizbA preliminary version: items, translation of content and item values

| No. | Area of content | Item and translation of content | *M* | *SD* | $s_i^2$ | $p_i$ | *ICC* | $\eta_p^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | RE | *Das Bild enthält zeichnerische Elemente*<br>The picture includes graphic elements | 2.68 | .43 | .19 | .54 | .567 | .573 |
| 2 | RE | *Das Bild enthält malerische Elemente*<br>The picture includes pictorial elements | 3.28 | .51 | .26 | .66 | .470 | .482 |
| 3 | RE | *Die Darstellungsweise ist gegenständlich*<br>The manner of representation is concrete | 2.14 | .38 | .14 | .43 | .704 | .705 |
| 4 | RE | *Die Darstellungsweise ist abstrakt*<br>The manner of representation is abstract | 2.55 | .35 | .13 | .51 | .609 | .615 |
| 5 | RE | *Der Farbauftrag ist pastos*<br>The color application is pastose | 2.48 | .58 | .34 | .50 | .519 | .526 |
| 6 | CO | *Die vorherrschende Farbgebung ist leuchtend*<br>The predominant coloring is vibrant | 2.62 | .33 | .11 | .52 | .581 | .588 |
| 7 | CO | *Im Bild befinden sich vorwiegend reine Farben*<br>In the picture primary colors are prevalent | 2.39 | .37 | .13 | .48 | .521 | .531 |
| 8 | CO | *Im Bild befinden sich vorwiegend Mischfarben (Sekundärfarben)*<br>In the picture mixed colors (secondary colors) are prevalent | 2.40 | .39 | .15 | .48 | .557 | .564 |
| 9 | CO | *Im Bild sind folgende Farbkontraste vorhanden: Komplementärkontrast*<br>In the picture there are following color contrasts: complementary contrast | 2.43 | .42 | .18 | .49 | .582 | .588 |
| 10 | SH | *Im Bild enthaltene Formen sind* vorwiegend *organisch*<br>In the picture organic shapes are prevalent | 2.45 | .39 | .15 | .49 | .555 | .563 |
| 11 | SH | *Im Bild enthaltene Formen sind vorwiegend geometrisch*<br>In the picture geometric shapes are prevalent | 2.47 | .36 | .13 | .49 | .570 | .578 |
| 12 | SH | *Die Linienführung verläuft vorwiegend gebogen*<br>The layout of the line is predominantly curved | 2.31 | .47 | .22 | .46 | .573 | .578 |
| 13 | SH | *Die Linienführung verläuft vorwiegend eckig*<br>The layout of the line is predominantly angled | 1.78 | .53 | .28 | .36 | .580 | .584 |
| 14 | SH | *Das Bild enthält enthält unbearbeitete Flächen*<br>The picture includes unworked areas | 2.33 | .34 | .11 | .47 | .616 | .621 |
| 15 | RA | *Das Bild wirkt tief*<br>The picture appears to be deep | 2.12 | .35 | .13 | .42 | .419 | .437 |
| 16 | RA | *Das Bild ist perspektivisch*<br>The picture is perspectival | 2.03 | .44 | .20 | .41 | .609 | .613 |
| 17 | RA | *Das Bild ist frei von Perspektive (aperspektivisch)*<br>The picture is without perspective (aperspectival) | 2.15 | .40 | .16 | .43 | .457 | .473 |
| 18 | MO | *Das Bild ist unruhig*<br>The picture is restless | 2.35 | .42 | .17 | .47 | .420 | .437 |
| 19 | MO | *Das Bild ist wild*<br>The picture is wild | 1.86 | .51 | .26 | .37 | .427 | .440 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 20 | COM | *Die Gesamtkomposition ist senkrecht angelegt*<br>The global composition is laid out vertically | 2.34 | .51 | .26 | .47 | .527 | .535 |
| 21 | COM | *Die Gesamtkomposition ist waagrecht angelegt*<br>The global composition is laid out horizontally | 1.60 | .55 | .31 | .32 | .451 | .463 |
| 22 | COM | *Die Gesamtkomposition ist diagonal angelegt*<br>The global composition is laid out diagonally | 2.30 | .49 | .24 | .46 | .451 | .466 |
| 23 | COM | *Die Gesamtkomposition ist flächendeckend ohne Hauptmotiv (All-Over-Structure)*<br>The global composition is laid out area-wide without a main subject (All-Over-Structure) | 1.60 | .53 | .28 | .32 | .405 | .423 |
| 24 | EX | *Das Bild wirkt diffus*<br>The picture appears to be diffuse | 1.80 | .37 | .14 | .36 | .398 | .419 |
| 25 | EX | *Das Bild wirkt präzise, exakt*<br>The picture appears to be precise, accurate | 2.40 | .42 | .18 | .48 | .437 | .452 |
| 26 | EX | *Das Bild wirkt harmonisch*<br>The picture appears to be harmonic | 2.60 | .49 | .24 | .52 | .369 | .351 |

RE = *representation*, CO = *color*, SH = *shape*, SP = *space*, MO = *motion*, COM = *composition*, EX = *expression*; *M* = mean (0 = *strongly disagree*, 5 = *strongly agree*), *SD* = item standard deviation, $s_i^2$ = item variance, $p_i$ = item difficulty, *ICC* = intra-class correlation coefficient, $\eta_p^2$ = partial eta squared effect size estimator

*Table 2.* Study 1: Rotated component matrix, factor loadings

| **Items** | **Components** | | | |
|---|---|---|---|---|
| | **1** | **2** | **3** | **4** |
| The picture is restless | .983 | | | |
| The picture appears to be diffuse | .937 | | | |
| The picture is wild | .898 | | | |
| The picture appears to be precise, accurate | -.808 | | .415 | |
| The global composition is laid out vertically | -.788 | | | |
| The picture appears to be harmonic | -.650 | .445 | | |
| In the picture primary colors are prevalent | .612 | .422 | | |
| The picture includes pictorial elements | | .930 | | |
| The picture includes unworked areas | | -.896 | | |
| The picture includes graphic elements | | -.881 | | |
| The predominant coloring is vibrant | | .847 | | |
| In the picture there are following color contrasts:<br>complementary contrast | .509 | .771 | | |
| In the picture mixed colors (secondary colors) are prevalent | .444 | .767 | | |
| The picture appears to be deep | | .597 | | -.576 |

| | | | | |
|---|---|---|---|---|
| The global composition is laid out area-wide without a main subject (All-Over-Structure) | | .559 | .432 | |
| In the picture organic shapes are prevalent | | | -.927 | |
| The layout of the line is predominantly angled | | | .925 | |
| In the picture geometric shapes are prevalent | | | .915 | |
| The layout of the line is predominantly curved | | | -.901 | |
| The manner of representation is abstract | | | .414 | .792 |
| The global composition is laid out horizontally | | | | -.777 |
| The manner of respresentation is concrete | | | | -.764 |
| The picture is perspectival | -.572 | | | -.742 |
| The picture is without perspective (aperspectival) | .499 | | | .724 |
| The global composition is laid out diagonally | | | -.489 | .668 |
| The color application is pastose | | .525 | | .597 |

Method of extraction: PCA; method of rotation: varimax, Kaiser normalization

*Table 3.* Study 2: Rotated component matrix, factor loadings

| **Items** | **Components** | | | |
|---|---|---|---|---|
| | **1** | **2** | **3** | **4** |
| The picture includes pictorial elements | .986 | | | |
| The picture includes graphic elements | -.970 | | | |
| The predominant coloring is vibrant | .888 | | | |
| In the picture there are following color contrasts: complementary contrast | .872 | | | |
| The picture includes unworked areas | -.822 | | | |
| The color application is pastose | .752 | | | |
| In the picture mixed colors (secondary colors) are prevalent | .700 | | | |
| The global composition is laid out area-wide without a main subject (All-Over-Structure) | .522 | .506 | | |
| The picture is perspectival | | -.938 | | |
| The picture is without perspective (aperspectival) | | .937 | | |
| The manner of representation is concrete | | -.885 | | |

| | | | | |
|---|---|---|---|---|
| The manner of representation is abstract | | .850 | | |
| The picture appears to be deep | .433 | -.783 | | |
| In the picture primary colors are prevalent | | .687 | | |
| The global composition is laid out vertically | | -.603 | | -.495 |
| In the picture geometric shapes are prevalent | | | .921 | |
| The layout of the line is predominantly angled | | | .899 | |
| The layout of the line is predominantly curved | | | -.886 | |
| In the picture organic shapes are prevalent | | | -.876 | |
| The global composition is laid out horizontally | | | .675 | |
| The global composition is laid out diagonally | | .463 | -.629 | |
| The picture appears to be diffuse | | | | .988 |
| The picture appears to be precise, accurate | | | | -.907 |
| The picture is wild | | | | .879 |
| The picture is restless | | .493 | | .816 |
| The picture appears to be harmonic | | | | -.434 |

Method of extraction: PCA; method of rotation: varimax, Kaiser normalization

## References


Arnheim, R. (2000). *Kunst und Sehen: Eine Psychologie des schöpferischen Auge*s [Art and seeing: A psychology of the creative eye]. Berlin: De Gruyter.

Brown, J. D. (2009). Principal components analysis and exploratory factor analysis – definitions, differences, and choices. *Shiken: JALT Testing & Evaluation SIG Newsletter, 13(1)*, 26-30.

Buck, J. N., & Warren, W. L. (1992). *The House-Tree-Person projective drawing technique: Manual and interpretive guide, revised ed.* Los Angeles: Western Psychological Services.

Chapman, L., Morabito, D., Ladakakos, C., Schreier, H., & Knudson, M. (2001). The effectiveness of art therapy interventions in reducing post traumatic stress disorder (PTSD):

Symptoms in pediatric trauma patients. *Art Therapy: Journal of the American Art Therapy Association, 18(2),* 100-104. doi:10.1080/07421656.2001.10129750

De Petrillo, L., & Winner, E. (2005). Does art improve mood? A test of key assumptions underlying art therapy. *Art Therapy: Journal of the American Art Therapy Association, 22(4),* 205-212.

Eber, P., Müller, S., Bader. R., & Baukus, P. (1998). *Die Nürtinger Beurteilungsskala zur Einschätzung von Patientenbildern im Therapieverlauf* [The Nürtinger Rating Scale for evaluating pictures of patients in the therapeutic process]. Nürtingen: Fachhochschule für Kunsttherapie.

Eitel, K., Szkura, L., Pokorny, D., & von Wietersheim, J. (2008). Do you see what I see? A study about the inter-rater reliability in art therapy. *Psychotherapy Research, 18(2),* 216-224. doi:10.1080/10503300701441995

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. doi: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]. *Zeitschrift für Musik-, Kunst und Tanztherapie, 18(2),* 85-99. doi: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“ [Reliability and validity of the instrument „DokuPro: Documentation of art therapeutic processes“]. *Musik, Tanz und Kunsttherapie, 20(1)*, 1-7. doi:10.1026/0933-6885.20.1.1

Elbing, U., Neuwirth, S., Knöbel, A., & Krilles-Mayr, M. (2010). *DokuPro: Dokumentation kunsttherapeutischer Prozesse: Handbuch zum Dokumentationsbogen* [DokuPro: Documentation of art therapeutic processes: Manual for the documentation form].

Franks, M., & Whitaker, R. (2007). The image, mentalisation and group art psychotherapy. *International Journal of Art Therapy, 12(1),* 3-16.

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. doi: 10.1080/07421656.2009.10129372

Gantt, L., & Tinnin, L. (2007). Intensive trauma therapy of PTSD and dissociation: An outcome study. *The Arts in Psychotherapy, 34*, 69-80. doi:10.1016/j.aip.2006.09.007

Gridley, M. C. (2013). Preference for abstract art according to thinking styles and personality. *North American Journal of Psychology, 15(3),* 463.

Gruber, H., Frieling, E. & Weis, J. (2002). Kunsttherapie: Entwicklung und Evaluierung eines Beobachtungsinstruments zur systematischen Analyse von Patientenbildern aus der Onkologie und Rheumatologie [Art therapy: Development and evaluation of a observational instrument for the systematic analysis of pictures by patients in oncology and rheumatology]. *Forschende Komplementärmedizin - Klassische Naturheilkunde, 9,* 138-146.

Gussak, D. (2006). Art therapy with prison inmates: A pilot study. *The Arts in Psychotherapy, 33,* 188-198.

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, 184*, 425-430. Retrieved June, 27, 2016, from www.researchgate.net/publication/14511452_The_descriptive_assessment_for_psychiatric_art._A_new_way_of_quantifying_paintings_by_psychiatric_patients/file/e0b4952a5d61b044a9.pdf

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]. Bern: Benteli.

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

Koh, E. (2014). The Cunningham Dax Collection: A unique mental health resource. *Australasian Psychiatry, 22(1),* 41-43.

Leiner, D. J. (2014). SoSci Survey (Version 2.5.00-i) [Computer software]. Retrieved August, 8, 2016, from https://www.soscisurvey.de

Machiodi, C. A. (2012). A brief overview of art-based assessments. In Machiodi, C.A. (Ed.). *The handbook of art therapy* (2nd ed.) (435-445). New York: The Guilford Press.

Mattson, D. C. (2009). Accessible image analysis for art assessment. *The Arts in Psychotherapy, 36(4*)*,* 208-213.

Mattson, D. C. (2010). Issues in computerized art therapy assessment. *The Arts in Psychotherapy, 37(4)*, 328-334.

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.

Metzl, E. S. (2008). Systematic analysis of art therapy research published in *ArtTherapy: Journal of AATA* between 1987 and 2004. *The Arts in Psychotherapy, 35,* 60–73. doi:10.1016/j.aip.2007.09.003

Meyer, G. (2011). *Sprache der Bilder: Kunst verstehen: Form, Farbe, Komposition* [Language of the pictures: Understanding art: Shape, color, composition]. Leipzig: Seemann.

Mills, A. (2003). The Diagnostic Drawing Series. In Machiodi, C. A. *The handbook of art therapy (*401-409). New York: The Guilford Press.

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, Instrumentation, and Computers, 32,* 396-402. Retrieved June, 27, 2016, from http://link.springer.com/article/10.3758/BF03200807#page-1

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.

Oster, J., Elbing, U., von Wietersheim, J., & Hölzer, M. (2010). Weiterentwicklung eines Instrumentes zur Dokumentation Kunsttherapeutischer Prozesse [Further development of an instrument for documentation of art therapeutic processes]. *Psychologische Medizin, 21, Sondernummer 2010*, 76.

Petersen, P., Gruber, H., & Tüpker, R. (Eds.) (2011). *Forschungsmethoden Künstlerischer Therapien* [Research methods in creative arts therapies]. Wiesbaden: Reichelt.

Stuhler-Bauer, A., & Elbing, U. (2003). Die phänomenologische Bilderfassung: Ein kunsttherapeutisches Experiment [The phenomenological picture survey: An art therapeutic experiment]. *Zeitschrift für Musik-, Kunst- und Tanztherapie, 14(1)*, 32-46. doi:10.1026//0933-6885.14.1.32

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.

University of Columbia (2014). SPSS, SAS, and MATLAB Programs for Determining the Number of Components and Factors Using Parallel Analysis and Velicer's MAP Test. Retrieved June, 27, 2016, from https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html.

Vollmar, K. (2008). *Die faszinierende Welt der Farben: Ein Glossar von A-Z* [The fascinating world of colors: A glossary from A to Z]. Witten: Ars Momentum.

Wagner, V., Menninghaus, W., Hanich, J., & Jacobsen, T. (2014). Art schema effects on affective experience: The case of disgusting images. *Psychology of Aesthetics, Creativity, and the Arts, 8(2),* 120.