

Ranking pictorial material of a landscape modeling task to detect mental stress levels – a pilot study

Abstract

Objective: Pictorial material is one of the oldest human ways of expressing inner-psychological and social processes. This pilot study investigates pictorial material of a landscape modeling task, movement trajectories, and a Tree-Drawings task using a ranking task approach.

Methods: 15 subjects (8 female, aged 19 to 60) were asked to do a gardening task in an experimentally constructed environment and Koch's Tree-Drawing-Test. The finished piece of landscape art and the drawing were photographed. Movement patterns and the resulting space used for the landscape modeling task were tracked to create spatial movement trajectory maps. Mental state was measured via the Global Severity Index (GSI) of the Brief Symptom Inventory (BSI-18) and the State Trait Anxiety Inventory (STAI-S). Picture ranking (from "healthy" to "mentally stressed") was applied to randomized landscape photographs, movement trajectories, and the Tree-Drawings from five experts (raters) representing the three areas of psychology/psychotherapy, art therapy and horticulture/landscaping. Correlations between ratings and GSI were calculated using Spearman's correlation coefficients. Intra-class coefficients (ICC) were calculated from the ratings to validate the initial assignment by experts. Finally, a multivariate regression model was applied to the data.

Results: Correlation between ratings and GSI are high to moderate with psychology/psychotherapy ($r_{\text{tree drawing}} = .714$; $r_{\text{movement trajectories}} = .611$) and with art therapy/theater education ($r_{\text{garden artwork}} = .803$). Only two ratings (one from psychology/psychotherapy ($r_{\text{LB07}} = .643$; $r_{\text{SH09}} = .818$) and one from art therapy/theater education ($r_{\text{SH14}} = .682$; $r_{\text{U025}} = .560$)) show a significantly high correlation. Multiple regressions implied that expertise in psychology/psychotherapy and art therapy/theater education might serve as a predictor for correct picture ranking.

Conclusions: Further investigations in experts ranking abilities as well as the enhancement of the picture ranking test are recommended directions of future research.

Keywords: pictorial material, projective tests, interrater reliability, rater-measure agreement, ranking task

Christina Niedermann^{1,2}

Thomas Ostermann²

1 Hochschule für Künste im Sozialen (HKS) – University of applied sciences and arts, Ottersberg, Germany

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

Introduction

Pictorial material is one of the oldest human ways of expressing inner psychic dynamics and social processes. Cave painting and engraving from the Early Aurignacian period, are commonly cited as evidence for this human artistic activity [1]. Over time, artistic activity developed from simplicity to technical and graphical complexity. Today, pictorial material is a medium of expression in a vast array of areas, including culture, politics, sociology and psychology. Indeed, pictorial material forms an important topic of psychological research in the fields of episodic memory [2], [3], perception tasks [4], [5], [6] and non-verbal understanding of language [7]. However, pictorial material as part of psychological assessments and projective tests, remains ambiguous.

These tests are summarized as a group of tests "which derive an indirect and global assessment of personality through the analysis of meaning or structure freely imposed by the subject upon unstructured or ambiguous materials" [8]. This material is supposed to project inner feelings, conflicts or interests in the outer structure, interpretation or design. Following the taxonomy of Lindzey, outer forms of expression can be examined by construction techniques, association techniques, completion methods, arrangement or selection methods and expression techniques. Such test procedures have a far-reaching significance in the interdisciplinary collaboration between psychology and art therapy, and numerous modules for evaluating pictorial material exist which have been summarized in [9].

As a special area of projective tests, figure drawings such as the House Tree Person Test (HTP) or the Tree Drawing Test (TDT) are essential parts of assessment processes in the field of clinical psychology [10]. For instance, the HTP shows moderate to good validity in diagnosis of forensic psychological items, such as psychiatry hospitalization [11], or in the diagnosis of child abuse [12] or in the diagnosis of attention-deficit hyperactivity disorder (ADHD) [13]. Likewise, the TDT has served as a diagnostic tool for detecting chronic diseases, dementia, schizophrenia or post-traumatic stress disorder [14]. Further fundamental aspects of projective methods can be found in [15] or [16].

For this study, an innovative interpretation of drawing is considered relating to fundamental aspects of projective testing: Since nature and humans are in an aesthetic relationship [17], the cultivation and forming of a landscape and thus the creation of a “landscape art” is considered to be a special type of landscape “drawing process” [18]. Materials such as plants, bricks or wood are “drawn” and arranged in a plain landscape canvas. In this context, a new kind of landscape picture is created. This aligns with Encyclopædia Britannica, where gardening is defined as “the laying out and care of a plot of ground devoted partially or wholly to the growing of plants such as flowers, herbs, or vegetables”. Encyclopædia Britannica concludes that “gardening can be considered both as an art, concerned with arranging plants harmoniously in their surroundings, and as a science, encompassing the principles and techniques of plant cultivation.” [19]. This idea has been integrated in art therapy, where Land Art is described as a creative therapeutic process in a public space with natural materials, for example in the therapy garden of a hospital [20]. Further aspects of nature in psychotherapy and arts therapies are described in detail in [21] and [22] where the holistic and environmental friendly therapeutic approach is demonstrated in various examples.

Apart from the therapeutic aspect, the aesthetic approach to landscape and gardens is also emphasized in landscape architecture or landscape design. Here, aesthetic sensibility both in composition, production and in the consideration of landscape design is a principle that makes landscape in its interaction with nature and multimedia components a work of art [23]. It is assumed that art triggers aesthetic sensation equal to a fundamental basic emotion in human beings [24]. Thus, this kind of creative landscaping process and its outcome (the final landscape picture) might be used for diagnostic processes [25].

In this study, we present a picture ranking test approach using the photographs of the final art works. In psychological literature, picture ranking tests are a recognized method of collecting intuitive knowledge and are particularly suitable for visual items, compared to card ranking test which use written statements or symbols [26], [27]. Picture ranking tests were first introduced in 1935 [28] as a categorize facial expressions. Due to its low-threshold approach and its non-verbal character the picture ranking

test has gained increased attention in psychology and the social sciences [29]. Currently, ranking tasks are also used in perceptual psychology, such as food preferences [30] or typicality ranking tasks in children, to assess their judgements or implicit knowledge when verbal expression is not as developed as visual perception or ability to judge based on other sensory experiences such as audio [31], [32].

In this study, the picture ranking test is used as an external assessment, during which experts rate the mental state of the drawer.

Aims

This pilot study aims at investigating the diagnostic potential of diverse pictorial material (photographs), e.g. of a landscape modeling, using a picture ranking task approach.

The following hypotheses were tested:

1. Raters of art therapy/theater education or psychology/psychotherapy or horticulture/landscaping can make gradual assessments of a person's condition based on a) photographs of the garden artwork, b) photographs of tree drawings, and c) movement trajectories.
2. Raters of art therapy/theater education make a more precise statement about the condition of a person than raters of horticulture/landscaping or psychology/psychotherapy.
3. There is a gender effect according the picture ranking task, especially with ranking of movement trajectories.

Methods

15 participants (8 women, none diverse) aged 19 and 60 took part in this pilot study. All participants were members or acquaintances of the Witten/Herdecke University (e.g. staff, students). Psychology students received credit points for participating. Detailed information about basic demographic data of the participants is shown in Table 1.

Prior to the gardening and drawing tasks, all participants were asked to complete a questionnaire including demographic categories such as age, and gender, as well as creativity and horticultural skills assessed on a 10-point visual analogue scale. Moreover, the test battery was completed by the Brief Symptom Inventory BSI-18 [33], [34] to detect the current mental state (expressed by the global severity index GSI as a sum value), and the State-Trait Anxiety Inventory (here state version STAI-S) [35] to detect the current anxiety level, expressed in STAI-S total score. The BSI-18 show satisfactory to high internal consistencies across representative German samples with a Cronbach's alpha of 0.88 for students, 0.82 for a non-clinical sample, 0.90 for a clinical sample with somatoform disorders, 0.88 for a clinical sample with anxiety, and 0.89 for a clinical sample with depression [34]. Likewise, the STAI-S shows high consistencies with a

Table 1: Basic demographic data of the participants of the pilot study

Variable	Male (n=7)	Female (n=8)	Total (N=15)
Age			
M (SD)	26.14 (2.73)	30.00 (17.68)	28.2 (12.78)
Median	25.00	20.5	25.0
Range	8	41	4
Missing	0	0	0
Education			
Bachelor	3 (42.9 %)	4 (50.0%)	7 (46.7%)
Master	2 (28.6%)	1 (12.5%)	3 (20.0%)
Not specified	2 (28.6%)	0	2 (13.3%)
Missing	0	3 (37.5%)	3 (20.0%)
Creativity			
M (SD)	7.0 (2.09)	7.5 (1.60)	7.29 (1.77)
Median	6.5	8.0	7.5
Range	5	5	6
Missing	1	0	1
Gardening Skills			
M (SD)	4.29 (2.289)	6.38 (1.19)	5.40 (2.03)
Median	3.0	6.5	6.0
Range	6	4	6
Missing	0	0	0
GSI total score			
M (SD)	9.14 (7.01)	6.88 (5.57)	7.93 (6.16)
Median	8.0	4.0	5.0
Range	19	15	19
Missing	0	0	0
STAI-S total score			
M (SD)	30.14 (7.03)	36.25 (6.32)	37.6 (6.59)
Median	36.0	34.5	36.0
Range	21	17	25
Missing	0	0	0

Cronbach's alpha between 0.88 and 0.94 for patients with anxious or depressive symptoms [36].

Afterwards, all participants were asked to complete the Tree Drawing Test according to Karl Koch [37] on a 40x40 cm white paper with a pencil of hardness 4B and a rubber. After completing these two tasks, participants were asked to do a half-hour landscaping modeling task in the outdoor area. Instructions were given to “create a garden from the given material in the assigned area”. Participants were instructed to start from an identical point of view and give a sign when they had finished the task. They were asked to finish the task with a three minute reminder before the time limit of 30 minutes. Both the tree drawings and the landscaping artwork were photographed with a commercially available camera from an identical point of view (tree drawings – top view; landscaping artworks – 45° from participant’s starting point).

Picture ranking test

The picture ranking test consists of three dimensions:

1. colored photographs of the garden artwork (see Figure 1),
2. black-and-white photographs of the tree drawing (see Figure 2), and
3. colored spatial movement trajectories of the participants (see Figure 3).

Movement trajectories were traced manually by transferring the observed movement to a blank DIN A4 sheet of paper printed with a square in the middle (representing the corner of the sandpit) while watching the video. Therefore, a psychology student (female) was asked, to watch the videos on a notebook via VLC Player in real-time speed. The task was, to observe the person’s movement trajectories based on the position of the persons ankle and copy them on the sheet of paper. Breaks or interruptions were marked with a point. Because of overlapping movement trajectories, the person was asked



Figure 1: Examples of photographs of the garden artwork



Figure 2: Examples of photographs of the tree drawing

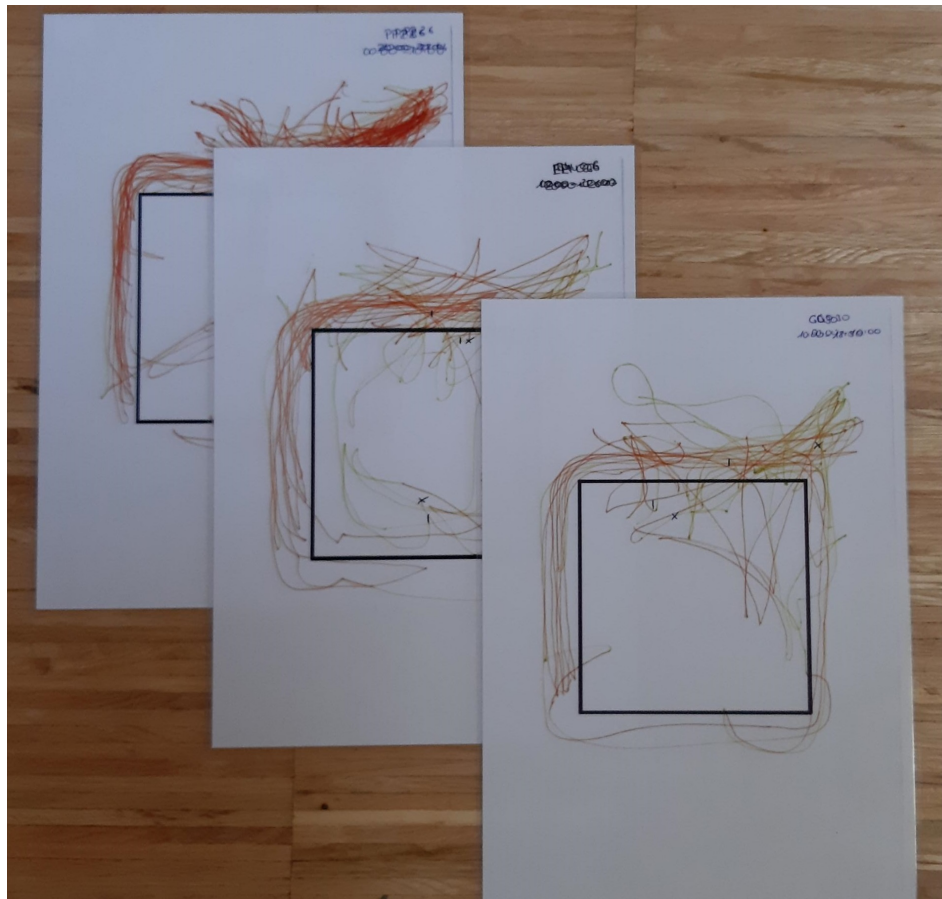


Figure 3: Examples of photographs of the movement trajectories

to take a new blanked sheet of paper every 10th minute. The position at beginning and finishing the task was marked with a cross. The papers were then scanned and placed on top of one another using Photoshop software. This is how the final image of the movement patterns was created, which was used for the picture ranking test. All pictures were printed on DIN A4 cardboard and laminated for protection.

Setting

Five raters each from the three areas of a) psychology/psychotherapy, b) art therapy/theater education, and c) horticulture/landscaping performed the picture ranking test. The raters were acquired from the Witten/Herdecke University, the University of Applied Sciences and Arts, Ottersberg and the Technical College for Horticulture, Essen via social media and personal contact.

The three categories of the picture ranking test were sorted according to a reference system from “0=healthy” to “15=mentally stressed”. This reference is based on the GSI, which was corrected by the STAI-S total sum value ($GSI_{corr.}$) to avoid ranks ties, so that each participant occupies their own rank. Participants were asked to “sort the pictures on a scale from ‘0=healthy’ to ‘15=mentally stressed’”. First, raters ranked the photographs of the garden artwork, followed by the photographs of the tree drawings and finally the movement trajectories. All ranking material was presented in a randomized order. Raters

were told to judge how good the items fit the ranking and not to base their judgement of how much they liked the photographs. As soon as the rater finished, the person gave a sign. There was no time limit for finishing the task. After each completed ranking, the pictures were collected and the next ranking started with the new items of the next category. It was therefore a closed process with simple implementation and without specific direction following the instructions given in the typicality ranking by Djalal and colleagues [32]. However, the time for finishing the task was not registered, nor were the number of corrections to the element’s position.

Evaluation process

Data was analysed by using SPSS Statistics 26. The average rating (MEAN) of the raters of the three professions was calculated for each category of the picture ranking test. Based on these averaged ratings, Spearman’s correlation coefficient between raters’ judgments and $GSI_{corr.}$ was calculated to investigate the agreement of all raters as well as the average rating of all five raters of each professional area. To determine the rater-measure agreement, the intra-class correlation coefficient (ICC) was calculated. The ICCs of the respective rater groups and all rater groups were calculated for each of the three picture ranking categories. To explore the difficulty level of the ranking task, a binary variable of each expert’s ranking was defined by dividing their ranking list into two

categories (“1=healthy”, “0=mentally stressed”) using a median-split of the ranks with the cut-off-score of 8. Inter-rater reliability on this data was calculated using Cohen’s Kappa coefficient. Finally, a multiple linear regression was used to predict the original ranking based on the GSI_{corr} including the independent variables of the three rater’s profession, average ratings, and gender of the pilot study participants. The level of significance was $\alpha=0.05$.

Ethical aspects

The research project was approved by the Ethics Committee of the Nürtingen-Geislingen University. All participants and all raters gave a written informed declaration of consent prior to admission.

Results

Participants

A total of 15 raters took part in the picture ranking test. Each rater assigned themselves to one of three professions: a) psychology/psychotherapy, b) art therapy/theater education, and c) horticulture/landscaping. Each group consisted out of five members.

The raters of group a) (5 male) on average were 31.4 years old ($SD=14.89$) and had an average of 8 years of professional experience ($SD=15.34$). Three of the raters were studying for a master’s degree, one rater was training to be a psychotherapist and one rater was self-employed as a psychotherapist. Both personal gardening skills ($\bar{x}=2.4$, $\min=2$, $\max=3$) and personal artistic skills ($\bar{x}=4.6$, $\min=3$, $\max=6$) were rated by the raters in the lower average range.

The raters of group b) (5 women) on average were 53.4 years old ($SD=3.507$) and had an average of 19.0 years of professional experience ($SD=8.19$). Two of the raters were self-employed, and three of them were employed. All raters had completed a university degree in art therapy at least at bachelor level. The raters rated their personal gardening skills ($\bar{x}=6.0$, $\min=4$, $\max=9$) in the average range and their personal artistic abilities ($\bar{x}=7.6$, $\min=6$, $\max=9$) in the upper average range.

The raters of group c) (1 female, 4 male) on average were 23.4 years old ($SD=.548$) and had an average of 5 years of professional experience ($SD=1$). All raters completed one year of further education after completing practical vocational training. Both the personal gardening skills ($\bar{x}=6.6$, $\min=5$, $\max=9$) and the personal artistic skills ($\bar{x}=5.0$, $\min=3$, $\max=9$) were assessed by the raters in the average range.

Rater-measure agreements

Ranking of garden artwork, Table 2, shows the rater’s individual measures of agreement with regard to the three image categories, sorted by profession. Results indicate high inconsistencies in all raters from the three profes-

sions. However, while no rater of group c) showed significant correlations, high to very high correlation was observed in three raters of group b) – ranking the garden artwork. One rater (SH14) showed high correlations in the ranking of all three picture categories while another rater (U025) did not correlate in any of the ranking with the true ranking. In group a), three raters showed high correlations with the tree drawings, and one rater (SH09) showed high correlations in the assignment of the other two picture categories.

Two raters (SH09 and SH14) show high levels of agreement both within their own rating and with regard to the true ranking list. They are therefore described in more details below in order to identify similarities.

Rater SH09 is a 26 years old male psychology student in a master’s course. He has no previous professional experience and rates his personal horticultural skills with 3/10 and his personal artistic skills with 4/10.

Rater SH14 is a 49 years old female art therapist with 20 years of professional experience. She rates her personal horticultural skills a 4/10 and her personal artistic skills a 7/10.

For reference, the complete ranking results of the most successful rater SH09 and the least successful rater SE14 with regard to the true ranking (GSI_{corr}) divided by the categories is provided in the supplementary material (see Attachment 1).

Table 3 shows intra-class coefficients for testing the rater-measure agreement. A cut-off according to [38] was set to $r>.5$, which corresponds to an average agreement. A total of two raters each from group a) psychology/psychotherapy (LB07, SH09) and group b) art therapy/theater (SH14, U025) achieve this value (see Table 3). The low level of agreement both on individual measurement as well as on average measurement raises the question of the difficulty of the task.

Difficulty level of the ranking task

Table 4 shows the rater-measure agreement according to Cohen’s Kappa. Following Landis and Koch [39], results show moderate and substantial correlations with respect to the ranking of garden artwork by raters of group b) art therapy and c) landscape design. Furthermore, results show moderate agreement with respect to the ranking of tree drawings by raters of group a) psychology/psychotherapy. There are no significant agreements with respect to the ranking of movement trajectories.

Possible predictive models

Figure 4, Figure 5, and Figure 6 show all ratings for the pictorial material according to the true ranking list (GSI_{corr}), in addition to gender-specific regression lines. For the garden artworks, the regressions show an explained variance of $R^2=.164$ for women and $R^2=.505$ for men. For the tree drawings, the explained variance is $R^2=.162$ for women and $R^2=.161$ for men. For the movement tra-

Table 2: Spearman correlations between raters' judgments and GSI_{corr}.

Profession	Code _{Rater}	<i>r</i> _{garden artwork}	<i>r</i> _{tree drawings}	<i>r</i> _{movement trajectories}
Psychology/psychotherapy	MEAN	.336	.678**	.535*
	BI21	-.257	.207	.343
	GM21	-.268	.582*	.239
	KH03	.147	.443	.061
	LB07	.479	.632*	.007
	SH09	.761**	.572*	.754**
Art therapy/theater education	MEAN	.818**	.317	.381
	BC08	.711**	-.071	.136
	BK03	.200	-.004	.029
	EO02	.729**	.107	.314
	SH14	.746**	.554*	.607*
	UO25	-.264	.096	-.104
Horticulture/landscaping	MEAN	.590*	.257	.052
	MA05	.225	.171	.200
	MB11	.164	-.345	.236
	MF28	.236	-.043	-.275
	PK17	.311	.217	.021
	SE14	.114	.000	-.289

* $p < 0.05$.; ** $p < 0.01$.**Table 3: Rater-measure agreement**

Intra-class coefficients of each rater's judgment sorted by profession				
Profession	Code _{Rater}	ICC	CI _{lower bound}	CI _{upper bound}
Psychology/psychotherapy	BI21	-.018	-1.417	.630
	GM21	.182	-.942	.702
	KH03	.290	-.686	.742
	LB07	.643*	.152	.870
	SH09	.818*	.568	.934
Art therapy/theater education	BC08	.288	-.689	.741
	BK03	-1.364	-4.612	.140
	EO02	.329	-.593	.756
	SH14	.682*	.245	.884
	UO25	.560*	-.045	.840
Horticulture/landscaping	MA05	-.286	-.695	-.740
	MB11	-.109	-1.632	.597
	MF28	-1.227	-4.288	.190
	PK17	.410	-.401	.785
	SE14	-1.731	-5.484	.007

*Cutt-off by $r > .5$ for average agreement**Table 4: Kappa coefficients**

Difficulty level of the ranking task						
Profession	K _{garden artwork}	Sign.	K _{tree drawings}	Sign.	K _{movement trajectories}	Sign.
Psychology/psychotherapy	.054	.833	.595	.020	.211	.298
Art therapy/theater education	.732	.005	.054	.833	.336	.189
Horticulture/landscaping	.595	.020	.211	.398	.071	.782

jectories the explained variance is $R^2=.040$ for women and $R^2=.169$ for men.

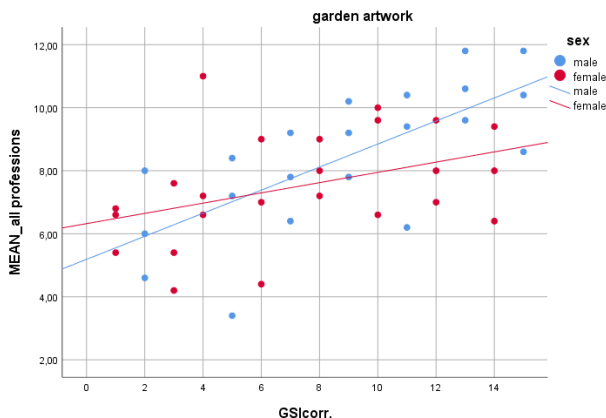


Figure 4: All experts rating of the garden artwork supplemented by gender-specific regression lines

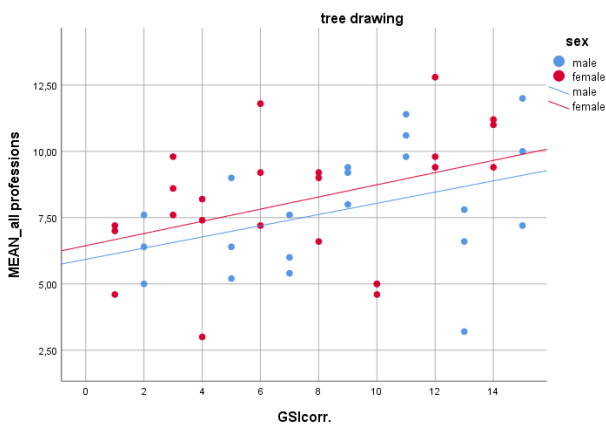


Figure 5: All experts rating of the tree drawings supplemented by gender-specific regression lines

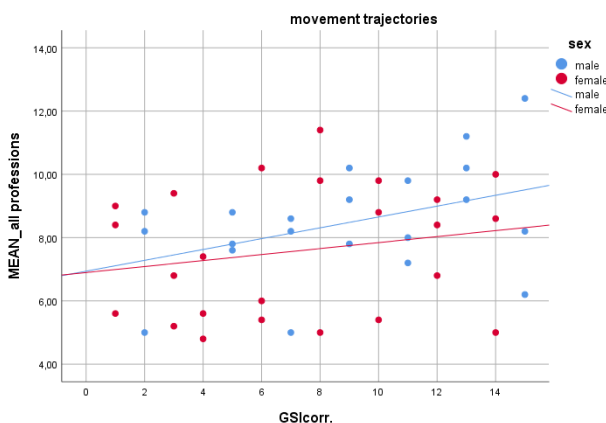


Figure 6: All experts rating of the movement trajectories supplemented by gender-specific regression lines

Table 5 shows three different multiple linear regressions with the original ranking list (GSI_{corr}) as the independent variable, and the average rating by each profession ($MEAN_{profession}$) and sex of the pilot study participants as the predictors. Regressions were calculated according to the categories of the pictorial material. With respect to the ranking of the garden artwork, results show a significant

impact of group b) art therapy. Regarding the ranking of the tree drawings, results show a significant impact of group a) psychology/psychotherapy. Regarding the ranking of the movement trajectories, results show a significant impact of group a) and group b). Neither predictor group c) horticulture/landscaping, nor gender, showed significant impact of any ranking task.

Discussion

The aim of this study was to examine the possibility of using the picture ranking test at hand to detect psychological stress levels in terms of feasibility and reliability. The results indicate that a picture ranking test might be of use in the field of clinical-psychological diagnostics. However, data at present is too heterogeneous to draw final conclusions and further research is needed to strengthen the results and the value of the ranking test. With respect to the first hypothesis, this study shows that raters of psychology/psychotherapy or art therapy/theater education or horticulture/landscaping can make gradual assessments of a person's condition based on a) photographs of the garden artwork, b) photographs of tree drawings and c) movement trajectories. For all surveyed raters the picture ranking task was a manageable task. However, not all expert groups came to valid results.

In support of the second hypothesis, this study found some evidence that raters of art therapy/theater education can use a) photographs of garden artwork, b) photographs of tree drawings and c) movement trajectories to make more precise statements about the condition of a person than raters of horticulture/landscaping or psychology/psychotherapy. At the same time, raters of psychology/psychotherapy did the tasks similarly well while raters of horticulture/landscaping performed poorly in comparison, even when the analysis was simplified using the median split.

Considering the third hypothesis, a gender effect could not be replicated at the level of the outcome evaluation. This is somehow contradictory to another analysis with data from this pilot study where spatial movement process was analyzed [40], [41]. In these analyses, results indicated a negative correlation between spatial movement and mental state for female participants and a positive correlation for male participants. In conclusion, women express mental stress with increased spatial movement activity whilst men restrict movement activity. However, this behavior was not visible in the final pictures of the ranking test applied here. Nevertheless, the results of this study contribute to the scientific call [42] to examine the gender differences in gardening behavior.

Impact of the rater's profession

Two raters (one from art therapy/theater education and one from psychology/psychotherapy) reached an acceptable rater-measure agreement. Results revealed they showed both high correlation with all three categories of

Table 5: Multiple linear regressions according to categories with $GSI_{corr.}$ and multiple predictors

Categories	Covariates	$R^2_{corr.}$	F	Sign.	B	SD	Sign.
Garden artwork	model	.671	8.133	.003			
	MEAN _{psychology/psychotherapy}				.323	.344	.307
	MEAN _{art therapy}				1.201	.398	.013
	MEAN _{horticulture/landscaping}				1.099	.557	.077
	Sex				1.397	1.461	.362
Tree drawing	model	.486	4.307	.028			
	MEAN _{psychology/psychotherapy}				1.439	.433	.008
	MEAN _{art therapy}				-.294	.702	.684
	MEAN _{horticulture/landscaping}				-.723	.444	.135
	sex				.387	2.042	.854
Movement trajectories	model	.486	4.081	.032			
	MEAN _{psychology/psychotherapy}				1.376	.466	.015
	MEAN _{art therapy}				1.634	.685	.038
	MEAN _{horticulture/landscaping}				.514	.484	.314
	sex				2.026	2.225	.384

pictorial material as well as high ICCs. Both showed a heterogeneous background according to demographic data such as gender or professional experiences. Thus and in accordance to [43] it remains unclear, if and under which conditions, reliable ranking results can be expected. At this pilot stage, it would be unreasonable to pose indications especially for practical clinical implementation. Furthermore, regression analysis showed an impact of the correct ranking by the profession psychology/psychotherapy in the tree drawings, while in the garden artwork, regressions show a positive impact of the profession art therapy/theater education. Two professions (art therapy/theater education and psychology/psychotherapy) do have an impact on the assessment of movement trajectories. In comparison, raters from horticulture/landscaping did not show significant performance in any of the three pictorial categories. None of the raters showed a significant match with the ranking even within their own rating. Statistical values did improve by simplifying the analysis (as it is displayed in the Kappa coefficients), but only in the category garden artwork, which is well-known material for these raters.

Results indicate that raters predicted the picture owners' mental state by assessing pictorial material best that belongs to their profession: Experts of psychology/psychotherapy performed well with projective tests such as the TDT as did experts of art therapy/theater education with drawn landscapes. This indicated an ability to transfer their disorder-specific and therapeutic knowledge. Likewise, experts of art therapy/theater education are familiar with artwork analysis. These raters have effectively developed a *diagnostic view*, even at a students' level.

Given these findings, experienced garden therapists might show better performance in the ranking task due to their specific therapeutic knowledge in the field of landscape and garden design. Furthermore, experts in the field of movement therapy, dance therapy or eurhythmmy may have a high degree of agreement in the ranking of photo-

graphs of movement trajectories. Semiotic, and perceptual aspects, (e.g. the aesthetic response), delight the interpretation of a picture and should therefore be considered in future. However, it would be beneficial for students of art therapy/theater education to come into contact with projective tests during their studies.

In a further evaluation, video recordings of the gardening task process were analyzed by experts from the three professions above. Subsequent qualitative data analysis showed four preliminary categories for assessing a person's mental stress level: movement patterns, attachment with the task, selection of material and hesitation/interruptions [44]. Once again, the difference between work analysis and process analysis should be explicitly pointed out. These considerations emphasize the multimedia aspect of this pilot study. In the clinical context, it also makes sense in terms of the multi-professions of a clinical-therapeutic team to include multimodal tests and evaluation instruments in the diagnostic process to produce a more holistic assessment of a person.

Photographs of the garden artwork

Interestingly, photographs of the garden artwork were best ranked by raters of art therapy/theater education. Being a rater of this group was the only significant predictor in the regression model. Once more, affiliation towards a certain pictorial material can be assumed. At the individual level, there were three raters of art therapy/theater education and one rater of psychology/psychotherapy who completed the ranking close to the original ranking list. Raters of art therapy/theater education and raters of horticulture/landscaping might benefit from making the task easier in terms of sorting the photographs into two categories as it was shown in the analysis of the kappa coefficients.

Though, being a rater of horticulture/landscaping is no predictor for ranking pictorial material not even for photographs of garden artwork.

In Figure 4 a gender effect could be assumed but in the multiple regression analysis the effect was not present. In the evaluation of the artwork as an end product of a gardening task the gender effect found in literature [45] could not be verified. At this point process evaluation and outcome evaluation clearly differ. In literature, women are more likely to plant ornamentals than men, who prefer to grow crop plants [44]. However, as a limitation, it is difficult to compare plant richness in real-world gardens with an experimental garden setting. In particular, material, time and attitude towards designing a garden for a small period of time (i.e. 30 min) instead of creating a garden for at least one season, is a reasonable bias to recognize when interpreting these results.

Photographs of the tree drawings

Raters of psychology/psychotherapy achieved the best results in assessing tree drawings. Average rating level, as well as individual level, showed high correlations to the mental state of the painters. Especially rater SH09 who fits the original ranking best with this pictorial material. Regression for the tree drawing ranking task showed an impact only when being a rater of psychology/psychotherapy. This result was stable when the difficulty level was reduced notionally to a binary outcome. As the TDT is well known in psychology, it is appropriate to assume an affinity towards this kind of drawing or test in the rater of psychology/psychotherapy.

Photographs of the movement trajectories

Movement trajectories do have a weak impact on the ranking task. Although the raters were able to complete this task, it seems that the pictorial material of the movement trajectories is quite abstract and hard to access. Even the simplification of the ranking task towards a binary outcome did not improve the raters ranking skills. Yet, the regression shows a significant impact of raters' profession regarding art therapy/theater education and psychology/psychotherapy.

In contrast to results here, digital assessed movement trajectories over time are more informative for clinical diagnostic as was shown in [40] Presumably, the higher precision of generated movement trajectories and the underlying numerical data could show that on the process-level men had less spatial movement activity in the case of psychological stress than women [40]. Other covariates, (e.g. age), showed no significant effects in multivariate analysis in this study. Compared to the approach in [40], [41] the current design does not involve a direct and objective measurements of the movement patterns, and thus may lead to inconclusive results.

Limitations

Although the picture ranking task shows clear results there are some limitations. First of all, picture ranking relies on the background and expertise of the raters [46]. To get a clearer impression on the impact of expertise, it would be desirable to include more than five raters for each profession. This could increase the variance of demographic aspects and work experience.

From a technical point of view, the pictorial material (especially the garden artwork) lacks high quality image format and resolution. Thus, important details may have been lost, which might have led to different rating results. Future research should also investigate more advanced approaches to create movement trajectories [40], [41]. Detailed information about how the test material was constructed, could improve rater's judgement. This also represents an approach closer to a clinical situation with more background knowledge about a person.

From the methodological part, the complexity of ranking 15 pictures simultaneously was quite high with a total of $15! = 1,307,674,368,000$ possible combinations. Thus and in accordance to the recommendations of [47] stepwise binary ranking or a free ranking task into i.e. healthy vs. non healthy might produce more robust results in future research. Although this study attempted to simplify the ranking task to a binary outcome, the underlying ranking task of the complete pictures might be the reason, why no improvements were observed in this part of the analysis.

Independent of the type of ranking task, the reasons why participants decided for a specific ranking should be obtained to get a clearer picture of the underlying decision processes and this is also recommended by [27]. In addition, the time needed for ranking each of the fifteen pictures should be considered. In a similar task in [48] it took a "striking" longer time for untrained participants to place the first picture compared to the time of trained participants.

Conclusions

Although the potential for bias in this study has to be taken into account as our observations could only cover certain aspects, our study adds to the existing literature on approaches to assess human behavior in nature and landscape, and builds bridges between research in art therapy and landscape design research [17], [49]. In particular, the combination of qualitative and quantitative aspects by means of a picture-ranking approach and a quantitative analysis of rater-measure agreement, might serve as a promising basis for further research [50]. Future research should, however, focus on improving the ranking test and determining the characteristics of successful raters. If applied to patients with mental disorders, comprehensive diagnostics should also be carried out alongside the process and outcome evaluation.

Notes

Competing interests

The authors declare that they have no competing interests.

Funding

Christina Niedermann was funded by the Andreas Tobias Kind Foundation (funding number: 2017/08) for the implementation of the pilot study and received a graduate scholarship from the University of Applied Sciences and Arts, Ottersberg. The Scholarship was part of the research focus *Artistic Interventions in Health Promotion and Prevention* funded by the Volkswagen Foundation and by the Lower Saxony Ministry for Science and Culture.

Acknowledgements

A big thank you goes to Linda Lösser and Marvin Beringer for actively supporting this pilot study at the Witten/Herdecke University as part of their studies in psychology, to Bethan Fox for proofreading the article, and to all the raters who contributed to the results of this study.

Attachments

Available from

<https://www.egms.de/en/journals/jat/2021-3/jat000011.shtml>

1. Attachment1_jat000011.pdf (4395 KB)
Complete ranking results

References

1. Pike AW, Hoffmann DL, García-Díez M, Pettitt PB, Alcolea J, De Balbín R, González-Sainz C, de las Heras C, Lasheras JA, Montes R, Zilhão J. U-series dating of Paleolithic art in 11 caves in Spain. *Science*. 2012 Jun;336(6087):1409-13. DOI: 10.1126/science.1219957
2. Brewin CR. Episodic memory, perceptual memory, and their interaction: foundations for a theory of posttraumatic stress disorder. *Psychol Bull*. 2014 Jan;140(1):69-97. DOI: 10.1037/a0033722
3. Ramponi C, Murphy FC, Calder AJ, Barnard PJ. Recognition memory for pictorial material in subclinical depression. *Acta Psychol (Amst)*. 2010 Nov;135(3):293-301. DOI: 10.1016/j.actpsy.2010.07.015
4. Wirtz MA. Bildbegriffe und Bildtheorien. In: Gretsche P, Holzäpfel L, editors. *Lernen mit Visualisierungen: Erkenntnisse aus der Forschung und deren Implikationen für die Fachdidaktik*. Münster: Waxmann Verlag; 2016.
5. Mohr J, Seyfarth J, Lueschow A, Weber JE, Wichmann FA, Obermayer K. BOIS-Berlin Object in Scene Database: Controlled Photographic Images for Visual Search Experiments with Quantified Contextual Priors. *Front Psychol*. 2016;7:749. DOI: 10.3389/fpsyg.2016.00749
6. Messerschmidt J. Professionell coachen mit Bildmaterialien: Die Sprache des Unbewussten verstehen und nutzen. essentials. Wiesbaden: Springer Fachmedien Wiesbaden; 2019. p 55. DOI: 10.1007/978-3-658-23692-2
7. Çöltekin A, Lokka IE, Zahner M. On the Usability and Usefulness of 3-D (Geo)visualizations – A Focus on Virtual Reality Environments. *Int Arch Photogramm Remote Sens Spatial Inf Sci*. 2016;XLII2:387-92. DOI: 10.5194/isprs-archives-XLII-2-387-2016
8. Tuleya LG. *Thesaurus of psychological index terms*. 11th ed. Washington, DC: American Psychological Association; 2007. xxvi, p. 463.
9. Schoch K, Gruber H, Ostermann T. Measuring art: Methodical development of a quantitative rating instrument measuring pictorial expression (RizbA). *The Arts in Psychotherapy*. 2017;55:73-9. DOI: 10.1016/j.aip.2017.04.014
10. Oster GD, Crone PG. *Using drawings in assessment and therapy: A guide for mental health professionals*. 2nd ed. New York, NY: Brunner-Routledge; 2004. p. 328.
11. Dyer FJ. Prediction of Psychiatric Hospitalization, Diagnoses, Arrests, and Violent Behavior Through Scored Drawings and Associations. *Psychol Rep*. 2018 Feb;121(1):4-25. DOI: 10.1177/0033294117718556
12. Blain GH, Bergner RM, Lewis ML, Goldstein MA. The use of objectively scorable house-tree-person indicators to establish child abuse. *J Clin Psychol*. 1981;37(3):667-73. DOI: 10.1002/1097-4679(198107)37:3<667::aid-jclp2270370339>3.0.co;2-p
13. Safran DS. An art therapy approach to attention-deficit/hyperactivity disorder. In: Case C, Dalley T, editors. *Handbook of art therapy // The handbook of art therapy*. Howe, East Sussex, New York, NY: Routledge, Taylor and Francis Group; 2014. pp. 181-92.
14. Schneider S, Niedermann C, Troche S, Ostermann T. Der Baumtest nach Karl Koch – seine diagnostische Relevanz im klinisch-therapeutischen Kontext. Ein systematischer Review. In: *Wissenschaftliche Fachgesellschaft für Künstlerische Therapien e.V., editor. Forschungstagung der Wissenschaftlichen Fachgesellschaft für Künstlerische Therapien (WFKT)*. Heidelberg, 20.09.2019. Düsseldorf: German Medical Science GMS Publishing House; 2020. Doc19wfk10. DOI: 10.3205/19wfk10
15. Maurer MH. Draw, Write, Speak, Play: The Role of Projection in Diagnosis and Therapy of Children and Adolescents. In: Maurer MH, editor. *Child and Adolescent Mental Health*. IntechOpen; 2017. DOI: 10.5772/67578
16. Wienand F. *Projektive Diagnostik bei Kindern, Jugendlichen und Familien: Grundlagen und Praxis – ein Handbuch*. 2nd ed. Stuttgart: Kohlhammer Verlag; 2018. S. 414.
17. Brady E. Aesthetic Regard for Nature in Environmental and Land Art. *Ethics Place Envir*. 2007;10(3):287-300. DOI: 10.1080/13668790701578538
18. Lorberg F. *Wahrnehmungspsychologie und Landschaft*. 2010. URN: urn:nbn:de:hebis:34-2010093034668
19. Herklots GA, Perrott R, Synge PM. Gardening. *Encyclopedia Britannica* [cited 2021 Aug 10]. Available from: <https://www.britannica.com/science/gardening>
20. Ganter-Argast C. Land Art als kunsttherapeutische Methode in der Psychosomatik – ein Bildungsprozess. In: Duncker H, Hampe R, Wigger M, editors. *Kreative Lernfelder: Künstlerische Therapien in Kultur- und Bildungskontexten*. Verlag Karl Alber; 2019. pp. 144-56.
21. Pfeifer E. *Natur in Psychotherapie und Künstlerischer Therapie*. In: *Therapie & Beratung*. Gießen: Psychosozial-Verlag; 2019. p. 454.

22. Petzold HG, Ellerbrock B, Hömberg R, editors. Die Neuen Naturtherapien: Handbuch der Garten-, Landschafts-, Wald- und Tiergestützten Therapie. Band 1: Grundlagen Garten- und Landschaftstherapie. Bielefeld: Aisthetis; 2019. p. 1006.
23. Dee C. To Design Landscape: Art, Nature & Utility. Hoboken: Taylor and Francis; 2012. p. 281.
24. Nadal M, Skov M. Introduction to the special issue: Toward an interdisciplinary neuroaesthetics. *Psych Aesth Crea Arts*. 2013;7(1):1-12. DOI: 10.1037/a0031842
25. de Thier-Patscher A. Kunsttherapeutische Diagnostik in der Psychiatrie und Psychotherapie mit Kindern und Jugendlichen: Entwicklung und Evaluation eines psychodynamischen Konzepts. Theorie und Forschung. Psychologie, 267. Regensburg: Roderer; 2014. p. 250.
26. Fincher S, Tenenbergs J. Making sense of card sorting data. *Expert Systems*. 2005;22(3):89-93.
27. Lobinger K, Brantner C. Picture-Sorting Techniques: Card-Sorting and Q-Sort as Alternative and Complementary Approaches in Visual Social Research. In: Pauwels L, Mannay D, editors. The SAGE handbook of visual research methods. Sage reference. London: SAGE Publications; 2020. pp. 309-21.
28. Hulin WS, Katz D. The Frois-Wittmann pictures of facial expression. *J Exp Psych*. 1935;18(4):482-98. DOI: 10.1037/h0056770
29. Lobinger K, Mengis J. Visuelle Methoden. In: Lobinger K, editor. Handbuch Visuelle Kommunikationsforschung. Springer Reference Sozialwissenschaften. Wiesbaden: Springer VS; 2018. pp. 1-24.
30. Valentin D, Cholet S, Nestrud M, Abdi H. Projective Mapping and Sorting Tasks. In: Hort J, Kemp S, Hollowood T, editors. Descriptive Analysis in Sensory Evaluation. London: Wiley-Blackwell; 2016.
31. Raman R, Kriegsman MA, Abdi H, Tillmann B, Dowling WJ. Bach, Mozart, and Beethoven: Sorting piano excerpts based on perceived similarity using DISTATIS. *New Ideas Psych*. 2020;57:100757. DOI: 10.1016/j.newideapsych.2019.100757
32. Djalal FM, Ameal E, Storms G. The Typicality Ranking Task: A New Method to Derive Typicality Judgments from Children. *PLoS One*. 2016;11(6):e0157936. DOI: 10.1371/journal.pone.0157936
33. Franke GH, Jaeger S, Glaesmer H, Barkmann C, Petrowski K, Braehler E. Psychometric analysis of the brief symptom inventory 18 (BSI-18) in a representative German sample. *BMC Med Res Methodol*. 2017 Jan;17(1):14. DOI: 10.1186/s12874-016-0283-3
34. Spitzer C, Hammer S, Löwe B, Grabe HJ, Barnow S, Rose M, Wingenfeld K, Freyberger HJ, Franke GH. Die Kurzform des Brief Symptom Inventory (BSI -18): erste Befunde zu den psychometrischen Kennwerten der deutschen Version [The short version of the Brief Symptom Inventory (BSI -18): preliminary psychometric properties of the German translation]. *Fortschr Neurol Psychiatr*. 2011 Sep;79(9):517-23. DOI: 10.1055/s-0031-1281602
35. Laux L, Glanzmann P, Schaffner P, Spielberger CD. Das State-Trait-Angstinventar: Theoretische Grundlagen und Handanweisung. Weinheim: Beltz Test GmbH; 1981.
36. Knight RG, Waal-Manning HJ, Spears GF. Some norms and reliability data for the State-Trait Anxiety Inventory and the Zung Self-Rating Depression scale. *Br J Clin Psychol*. 1983 Nov;22 (Pt 4):245-9. DOI: 10.1111/j.2044-8260.1983.tb00610.x
37. Koch K. Der Baumtest: der Baumzeichenversuch als psychodiagnostisches Hilfsmittel. 12th ed. Bern, Göttingen: Huber; 2008.
38. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016 Jun;15(2):155-63. DOI: 10.1016/j.jcm.2016.02.012
39. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977 Mar;33(1):159-74.
40. Unger S, Appelbaum S, Ostermann T, Niedermann C. Movement Entropy in a Gardening Design Task as a Diagnostic Marker for Mental Disorders: Results of a Pilot Study. In: Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technology. February 11-13, 2021. Vol. 5 – HEALTHINF. SciTePress (Science and Technology Publications, Lda); 2021. pp. 337-43. DOI: 10.5220/0010227203370343
41. Unger S, Ostermann T, Appelbaum S, Niedermann C. Tracking human spatial motion in the design of gardens as a potential method to detect Mental Disorders – A pilot study using a Homography Approach. *Acta Hort*. 2021 [in press].
42. Zypchyn K. Getting Back to the Garden – Reflections on Gendered Behaviours in Home Gardening. *Earth Common J*. 2011;2(1). DOI: 10.31542/j.ecj.60
43. Ryan GW, Bernard HR. Techniques to Identify Themes. *Field Methods*. 2003;15(1):85-109. DOI: 10.1177/1525822X02239569
44. Niedermann C, Ostermann T. Landscape modelling as a prognostic tool to detect mental disorders – First results of a pilot study. In: International Union of Psychological Science (IUPsyS), editor. Open Minds, Societies and the World. 32nd International Conference of Psychology; 2021 Jul 18.-23; Prague.
45. Philpott SM, Egerer MH, Bichier P, Cohen H, Cohen R, Liere H, Jha S, Lin BB. Gardener demographics, experience, and motivations drive differences in plant species richness and composition in urban gardens. *Ecol Soc*. 2020;25(4). DOI: 10.5751/ES-11666-250408
46. Weibel N, Hwang SO, Rick S, Sayyari E, Lenzen D, Hillan J. Hands that Speak: An Integrated Approach to Studying Complex Human Communicative Body Movements. In: 2016 49th Hawaii International Conference on System Sciences (HICSS), 5-8 Jan 2016. Koloa, HI, USA: IEEE; 2016. pp. 610-9. DOI: 10.1109/HICSS.2016.82
47. Blanchard SJ, Banerji I. Evidence-based recommendations for designing free-sorting experiments. *Behav Res Methods*. 2016 Dec;48(4):1318-36. DOI: 10.3758/s13428-015-0644-6
48. Hansen M, Spada H. Designing Instructional Support for Individual and Collaborative Demands on Net-Based Problem-Solving in Dyads. In: Proceedings of the 7th International Conference on Learning Sciences. ICLS '06: International Society of the Learning Sciences; 2006. pp. 229-35.
49. Whitaker P. Groundswell: The Nature and Landscape of Art Therapy. In: Moon C, editor. Materials and Media in Art Therapie: Critical understandings of diverse artistic vocabularies. New York: Routledge; 2010. pp. 119-36.
50. Bergmann T, Sappok T, Diefenbacher A, Dziobek I. Music in diagnostics: using musical interactional settings for diagnosing autism in adults with intellectual developmental disabilities. *Nordic J Music Ther*. 2016;25(4):319-51. DOI: 10.1080/08098131.2015.1039567

Corresponding author:

Christina Niedermann

Department of Psychology and Psychotherapy,

Witten/Herdecke University, Alfred-Herrhausen-Str. 50,

58448 Witten, Germany

christina.niedermann@uni-wh.de

Please cite as

Niedermann C, Ostermann T. Ranking pictorial material of a landscape modeling task to detect mental stress levels – a pilot study. *GMS J Art Ther.* 2021;3:Doc02.

DOI: 10.3205/jat000011, URN: urn:nbn:de:0183-jat0000117

This article is freely available from

<https://www.egms.de/en/journals/jat/2021-3/jat000011.shtml>

Published: 2021-11-23

Copyright

©2021 Niedermann et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.