Shared meaning in representational and abstract visual art: an empirical study

Astrid Schepman and Paul Rodway

University of Chester

Author Note

Department of Psychology, University of Chester.

The Javascript software that called the semantic similarity service and stored the scores was written by Brian Rodway (brian@affinitystudios.co.uk) of Affinity Studios, UK, (http://www.affinitystudios.co.uk/index.html). We are grateful to two anonymous reviewers for their helpful suggestions.

Correspondence concerning this article should be addressed to Dr Astrid Schepman, Department of Psychology, University of Chester, Parkgate Road, Chester, Cheshire, CH1 4BJ, United Kingdom. Email: a.schepman@chester.ac.uk

Abstract

A longstanding and important question is how meaning is generated by visual art. One view is that abstract art uses a universal language whereas representational art is tied to specific knowledge. This view predicts that meaning for abstract is shared across viewers to a greater extent than for representational art. This contrasts with a view of greater shared meaning for representational than abstract art, because of shared associations for the entities depicted in representational art, as supported by recent empirical findings. This study examined the contrasting predictions derived from these two views. 49 nonexpert adult participants wrote brief descriptions of meanings that they attributed to 20 abstract and 20 representational artworks, generating a corpus of 1918 texts. Computational analyses (semantic textual similarity, latent semantic analysis) and linguistic analysis (type-token ratio) provided triangulated quantitative data. Frequentist and Bayesian statistical analyses showed that meanings were shared to a somewhat greater extent for representational art, but that meanings for abstract artworks were also shared above baseline. Triangulated human and machine analyses of the texts showed core shared meanings for both art types, derived from literal and metaphoric interpretations of visual elements. The findings support the view that representational art elicits higher levels of shared meaning than abstract art. The empirical findings can be used to enhance theoretical and computational models of aesthetic evaluation, and the rigorous new methodologies developed can be deployed in many other contexts.

Key words: Natural Language Processing; Computational Linguistics; Empirical Aesthetics; Meaning; Art

Shared meaning in representational and abstract visual art: an empirical study

A longstanding problem in the psychology of aesthetics is our understanding of the viewer's attribution of meaning to visual artworks. In general, it is understood that viewers' interpretations of visual artworks derive from complex, multilayered psychological processes, including the perceptual interpretation of the artworks, activations of stored representations, as well as affective reactions (e.g. Leder, Belke, Oeberst, & Augustin, 2004). Whether the artwork is representational or abstract is thought to play a major role in how these processes are completed and ultimately in the semantic representations at which viewers arrive. However, there is some disagreement in the literature on the nature of these interpretative processes in relation to abstract vs. representational art, and more specifically how this relates to meaning, and how meaning is shared across viewers.

One school of thought suggests that abstract art follows a universal language. To illustrate, British artist Ben Nicholson (1941) likens the experience of looking at a representational painting of Greece as an effortful event, where the viewer has to concentrate, whereas, he argues, an abstract painting will provide "the actual quality of Greece itself, and this will become a part of the light and the space and the life in the room – there is no need to concentrate, it becomes a part of living" (p. 1). Thus, Nicholson (1941) argues "I think that so far from being limited expression, understood by few, abstract art is a powerful, unlimited and universal language" (p. 3). This universal language provides viewers with meanings grounded in "moments of human consciousness" (Museum of Modern Art, 1957, p. 136) that the artist is thought to express via psychological processes that translate these into the color, texture, movement and form visible in the artwork. These ideas can be followed through early

20th century artists / art theorists such as Wassily Kandinsky (e.g. Wünsche, 2016), František Kupka (e.g. Hume, 2016), Ernst Wilhelm Nay (see Müller, 2016), Adolph Gottlieb (Hirsch, n.d.), Mark Rothko (2006), and Piet Mondrian (1937), while also being advanced by art historians such as Haftmann (1954; see Müller, 2016, p. 191). Under this view, abstract art contains ideas or depictions of psychological states that are culture-free, not bounded by specific individual experiences or specific knowledge, but based in universal human psychological processes. The artist creating abstract art could convey a shared meaning to viewers, because the psychological reaction is elicited by a shared and universal language of abstract art. Representational artwork, on the other hand, may need specific knowledge to interpret the material depicted in the artwork, with meaning being bounded by culture and specific experiences, and therefore less likely to be shared across viewers. The view that abstract art draws on a universal language has been held and promoted over the years by artists, art historians, art theorists, curators (e.g. Arvidsson & Dahlström, 2018) and auctioneers (e.g. Schwartzman, 2018). Under this view, abstract art is thought to create a higher level of shared meaning than representational art.

The idea that abstract art leads to shared meaning via a universal language has been noted to have received little empirical testing (Brinkmann, Commare, Leder, & Rosenberg, 2014). When a recent test using eye movements and evaluations in response to artworks was conducted this did not support the notion of abstract art using a universal language (Brinkmann et al., 2014). The suggestion of a direct transfer of an idea from the artist to the viewer has also been critiqued on philosophical grounds. For example, Crowther (2017) suggests that while abstract artists provide meaning, at times augmented with titles, this meaning can only be allusive, and audiences bring their own ideas to create their own individual interpretations of abstract art. More recently, however, it has been found (Kuipers, Jones, & Thierry, 2018) that words that had been rated as “related” to an abstract artwork were primed by that artwork, when compared to words rated “unrelated”, a process also visible in the N400 component of the EEG/ERP signal, traditionally thought to be an index of semantic processing. This has been interpreted as evidence of shared meaning for abstract art, though it could be argued that the evidence is restricted, because the researchers generated the meanings, and the viewers engaged in a recognition of matching or mismatching semantic labels somewhat passively. Further, no comparison was made with representational art. Thus, there is contradictory and incomplete evidence on the notion of abstract art leading to shared meanings.

A contrasting way of conceptualizing semantic interpretations of images suggests that representational images may have more defined meanings than abstract images, because they depict real-life scenes and objects that relate to the viewer’s experiences (Vessel & Rubin, 2010). There is growing empirical support for this view. Vessel and Rubin (2010) found that aesthetic evaluations are shared to a greater extent for representational photographs than for abstract photographs. Their interpretation was that this was due to the shared associations that viewers have of the entities depicted in representational images, which are lacking in abstract images, making evaluations of abstract photographs more idiosyncratic. Further findings by Schepman, Rodway, Pullen and Kirkham (2015a) replicated the pattern observed by Vessel and Rubin (2010). They made use of abstract and representational artworks, eliciting liking and associations (thoughts brought to mind). Participants then rated the valence of these associations. Both the liking ratings and the valence of the associations in this study replicated the pattern observed by Vessel & Rubin (2010). The pattern of higher shared liking for representational than abstract art was also observed in children aged eight and ten, but not aged four and six (Rodway, Kirkham, Schepman, Lambert, & Locke, 2016). The extent of private and shared taste in abstract art was further quantified based on a variance components analysis by Leder, Goller, Rigotti, and Forster (2016), who observed 75% of private taste for abstract art, compared to 40% for images of faces, which form natural and representational stimuli. This was further extended to a demonstration of greater shared taste for natural stimuli than for human artifacts (Vessel, Maurer, Denker, & Starr, 2018). In a test of the role of shared emotion in shared taste, Schepman et al. (2015a) showed that the emotional valence of the associations triggered by artworks followed the same pattern, with greater shared valence for associations triggered by representational than abstract artworks. This was extended by Tinio and Gartus (2018), who used the Geneva Emotion Wheel (Scherer, 2005) to show that viewers share aesthetic emotions.

To explore the link between shared taste and shared associations Schepman, Rodway and Pullen, (2015b) used a computational semantic similarity analysis of the viewers' associations with abstract and representational artworks from Schepman et al.(2015a), and found that associations provided in response to representational artwork showed statistically significantly higher levels of semantic similarity to each other than those provided in response to abstract artwork. Interestingly, semantic similarity scores for both art types significantly exceeded a random baseline control condition, suggesting shared associations for abstract as well as for representational artworks. A computational semantic similarity analysis of meaning-based responses generated by the child participants in Rodway at al, (2016) suggested that shared meaning could be detected from age four onwards, with statistically significantly higher levels of semantic similarity for representational than abstract art (Schepman, Kirkham, Rodway, Lambert, & Locke, 2018).

Although the work on shared associations cited above provides evidence to support the idea that associations attributed to representational art are shared to a greater extent than those attributed to abstract art, there is a gap in the evidence when attempting to generalize this to shared meaning. In this context it is important to distinguish these two concepts. In Schepman et al. (2015a, 2015b), the participants had been instructed to "write a word or short description … of any thoughts that the work of art brought to mind", following a technique used to gauge associations that consumers had with commercial brands, as well as the valence of these associations (Spears, Brown, and Dacin, 2006). This elicited a variety of rapidly generated mostly brief verbal associations, which included the naming of unconnected visual elements, emotions, and connotations, but was not a direct measure of meaning. Thus, in prior work that has used semantic similarity analyses, the input to the analysis did not consist of viewers' meanings of artworks, but instead of viewers' associations (Schepman et al, 2015b) or their meaning-based justifications for aesthetic evaluations (Schepman et al. 2018). While they provide some evidence suggestive of shared semantic elements, empirical support based on meanings that viewers attribute to artworks via direct instructions to elicit meanings and interpretations would form stronger evidence of shared meaning. In this context, we draw on a distinction made by Parsons (2015) between "connotations", which may be generated in response to an artwork, with "interpretation" involving a more deliberate process of selection among the many possible connotations. The first aim of the current paper was to measure shared meaning in art directly to test whether meaning is more shared in abstract or representational art. To achieve this, we phrased the instructions with the aim of encouraging deeper thought and interpretation about the meaning of the artwork than had been the case in Schepman et al. (2015a, 2015b). The second aim was to further explore whether there was evidence of shared meaning in both abstract and representational art. We tested this using quantitative analyses of textual data, and explored this further with content-based analyses.

The operationalization of the quantitative data involved two independent variables, art type (abstract vs. representational) and randomization (experimental vs. baseline control, more fully explained in the Results section) in a fully crossed design. We took multiple dependent variables that measured the extent of shared meaning.

Our first hypothesis to test was whether there would be evidence of shared meaning. This would be supported by difference between the baseline control randomization and the experimental randomization (as in Schepman et al. 2015b).

Our second hypothesis was derived from the two opposing theoretical stances regarding shared meaning in abstract vs. representational art. Support for the view that abstract art uses a universal language and therefore has greater shared meaning than representational art would show abstract art leading to greater shared meaning. In contrast, greater shared meaning for representational art would provide support for the notion that cumulative experiences with real-life entities would lead to greater shared meaning (see e.g. Vessel & Rubin, 2010).

For the content-based data we used triangulated content analysis methods (human coding, word frequency analysis, and latent semantic analysis). We examined whether distinct, non-overlapping core meanings for each artwork could be distilled from the participants' responses, and whether this applied to both art types.

Method

Participants

There were 49 adult student participants (44 female, 5 male) recruited via the University of Chester Psychology Participant Pool. Data from one further participant were excluded due to a failure to comply with the instructions. The University of Chester student population matches the socio-economic strata in the UK population closely, based on its institutional Gini coefficient (Martin, 2018). The mean age was 22 years (SD 5.8 years; range 18-48 years).

The participants' art background was established by brief survey questions with multiple choice options. 43% never visited art galleries, 33% did so once per year, 22% between one and four times per year, and 2% more than four times per year. 41% had never produced an artwork, 31% had produced between one and four artworks, 12% between four and twenty, and 16% more than twenty. 10% reported no art education, 74% basic art lessons at school, 14% arts at A-level (UK pre-university high school qualification at age 18), 1% arts at degree level, and none at master's or doctoral level. Overall, the sample was best classified as consisting primarily of nonexperts, for whom art education, interests, and engagement were at a level comparable to the general population.

Ethical Approval

The research was reviewed and approved by the University of Chester Department of Psychology Ethics Committee, and was compliant with the ethical guidelines of the British Psychological Society.

Stimuli

The stimuli were 20 abstract and 20 representational painting, mostly from the 20th Century, mostly from Europe, North America, and Australia, with some exceptions. The artworks were chosen from online databases hosted by a range of art galleries, art museums or art collections, as well as other internet sites, so that the set included a variety of artists, styles, colors and subjects where applicable, i.e. for the representational artworks. Other factors such as popularity or quality did not form a selection criterion. The digital images were sized so that the longest dimension was 567 pixels (which equated to 15 cm when printed). Aspect ratios from the original source were retained. A full list of the artworks can be seen in Supplemental Materials 1.

Procedure

Following informed consent and brief demographic and art background questions (see ‘Participants’), the participants were given the following instructions:

“We are interested in examining people’s interpretations of artworks. We have produced a booklet with images of 40 artworks which incorporates work from a wide range of artists, and features a range of styles. Please look at each artwork in the booklet for as long as you would like and write a brief response (word, phrase or short sentence) below each artwork indicating what you think the artwork means. Because interpreting artwork is quite subjective and individual, there are no right or wrong answers. We are very interested in your own personal interpretation.”

Participants worked their way through a color-printed booklet in which each artwork was horizontally centered and vertically aligned near the top of an A4 white paper page, writing their response under each artwork in black pen. The artworks’ titles or artists’ names were not provided. The artworks appeared in one of three different random orders to mitigate against order effects. As part of the ethics arrangements, participants were free to withhold responses. Data were anonymous upon submission. The whole process took approximately 20 minutes.

The reason for our request for relatively short responses was to keep these amenable for our intended analysis techniques, enabling quantitative and statistical analysis. Additionally, keeping the overall task duration relatively short reduced the burden on the participants.

Results

Data treatment

Verbal data were transferred from handwritten to digital records (Excel, SPSS) for analysis. Any spelling errors were corrected during data entry. There were 42 (2%) missing responses, and 1918 valid responses.

Measures and statistical analysis: General

We used computational analyses that built on earlier work (Schepman et al. 2015b, 2018) which had used semantic textual similarity. To strengthen the evidence, this measure was triangulated against a further computational measure, latent semantic analysis, and a quantitative linguistic analysis, type-token ratio. More detailed justifications and explanations of these measures follow shortly.

We calculated frequentist statistics using IBM SPSS 24/25, and, where available, Bayesian statistics with JASP (JASP Team, 2018; versions 0.8.6.0, 0.9.0.0 and 0.9.1.0; the latter two versions were used primarily for more recently added non-parametric Bayesian statistics). For the Bayesian analyses, we were mainly interested in reporting BF_{10} values (Bayes Factor for the Experimental over the Null Hypothesis; see e.g. Schönbrodt & Wagenmakers, 2018) to express the strength of the evidence for the experimental hypothesis in relation to the null hypothesis. Where priors were under user control, we consistently chose the default priors as set in the software, because there was no specific prior information to determine the center or width of informed prior distributions. Finally, we used Jamovi's (2019) GAMLj 1.0.5 module (Galluci, 2019) to run Linear Mixed Models.

Experimental and baseline control pairings and matrices

This is a brief explanation of our random pairing protocols, an essential first step in the computational analysis. To enable the semantic textual similarity and pairwise latent semantic analysis we followed the previously developed protocol of creating random pairings between pairs of responses (Schepman et al., 2015b, 2018). This is because the semantic textual similarity analysis requires pairs of stimuli as its input, to allow it to provide score of the semantic similarity between the members of the pair as its output. Latent semantic analysis can be run on the same pairwise basis, and we included that version to maximize comparability between the two measures.

We created two main types of pairings. The first type, the experimental pairings, involved randomly pairing a participant’s response given to a specific artwork with another participant’s response given to that same artwork. The second type, the baseline control pairings, involved random pairings between a participant’s response to a specific art type (abstract, representational) and another response to that art type, pooling across all the artworks in that art type. Having a baseline control for each art type was necessary because responses could differ as a function of art type for reasons not related to shared semantics, and having a baseline control condition allowed us to anchor the quantitative output for the experimental pairings to these baselines to aid interpretation. Pairings were performed exhaustively, so that each response was paired with another response in each randomization. In total there were 3836 (2 x 1918) pairs across the two randomizations. Note that we supplemented the pairwise latent semantic analysis sampling with a further matrix measure based on a more comprehensive sampling method. This was because of concerns about a potential sampling bias inherent in using pairwise measures based on just one cycle of random pairings. For the experimental condition, we established how semantically similar, on average, a participant's response to an artwork was to all other participants' responses to that same item. To accomplish this, we generated latent semantic analysis scores for each participant response to an item paired with all other participants' responses to that item. We then averaged those scores across each participant responding to each item. For the baseline control condition, we created control matrices that were matched in size to the matrices described above, but the participants and items were randomly shuffled within each art type. We refer to this measure as matrix latent semantic analysis.

Quantitative measures and justifications

As stated, we used three measures (one with two sampling methods) to estimate the semantic similarity of the responses. This was to triangulate the data with multiple estimates, rather than relying on one estimate of semantic similarity. The three measures are briefly described and justified next.

Semantic Textual Similarity. We made use of UMBC eBiquity Semantic Textual Similarity software (Han et al., 2013; http://swoogle.umbc.edu/StsService/index.html; http://swoogle.umbc.edu/StsService/GetStsSim), which was also used in Schepman et al. (2015b, 2018). As explained in Han et al. (2013), the software operates in a multi-layered way, drawing on alignment algorithms, as well as lexical semantics (Wordnet; Miller, 1995) and latent semantic analysis (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990; Landauer & Dumais, 1997). It performs well against other similar software (see Han et al., 2013). While it is possible that better software may be developed in the future, the performance of Han et al.'s (2013) software was not bettered by software developed subsequently, as noted by Vo and Popescu (2016). This suggests it can still be considered state-of-the-art (see Schepman et al., 2018). The software produces scores between 0 and 1, with 0 being no semantic similarity between the pair of texts provided as input, and 1 being maximum similarity. To illustrate with some examples (not from our corpus), “a smiling woman” paired with “a cheerful person” gives a semantic similarity score of 0.53, the pair “the cat scratched the dog” and “the dog was scratched by the cat” gives a score of 1, while “the car will not start” paired with “a bouquet of flowers” receives a similarity score of 0.

Latent semantic analysis. Latent semantic analysis analyzes corpora for word co-occurrences in separate texts or passages. Unlike semantic textual similarity, latent semantic analysis treats texts as “bags of words”, without order or syntax. Latent semantic analysis works on the principle that words about particular topics may co-occur in specific texts alongside other words that are also relevant to those topics. As described in detail in Landauer, Foltz and Laham (1998) the co-occurrences of words across text are expressed using a matrix, and, with entropy weighting to correct for the undue impact of frequent function words (“the”, “a” etc.) over meaning-bearing words (Laundauer et al., 1998; see also Nakov, Popova, & Mateev, 200), then transformed into a vector space using singular value decomposition (a process akin to factor analysis, leading to dimension reduction). The cosine of the angle between vectors is a measure of semantic similarity. Latent semantic analysis has been widely used in informatics, for example to generate key words for indexing (Deerwester et al., 1990). Empirically, similarity scores produced by latent semantic analysis correspond well with human judgements (see discussion in Landauer et al., 1998) and show a good correspondence with priming data (Günther, Dudschig, & Kaup, 2016).

Latent semantic analysis is performed on a training database, which can then be queried. This allows users to investigate, for example, the coherence between two further pieces of text that were not in the training database. While it is possible to build one's own training database, a ready-made set of databases is available for use via Latent Semantic Analysis @ CU Boulder (2018). This was trained by its developers on a variety of semantic spaces, available for selection. It has good correspondence with the output from other databases (Heinen & Johnson, 2018). We chose "General reading up to first year of college" because that formed a good match with our needs, matching the lowest level of education present among our participants. We initially used the pairwise comparison querying tool, and we followed this up with the matrix querying tool in a further, separate analysis. Both were used with the "document to document" setting, which activated the entropy weighting. In theory, the cosine can vary between -1 and 1, but in practice, we tended to observe only some negative values just below 0, with the minimum being -.17. The higher the value of the cosine, the greater the association or coherence between the words or word strings in question. To illustrate with the same examples as before "a smiling woman" paired with "a cheerful person" gives a pairwise latent semantic analysis score of 0.03; "the cat scratched the dog" with "the dog was scratched by the cat" gives a score of 1, while "the car will not start" paired with "a bouquet of flowers" receives a score of 0.02. It can be seen that the scores differ from those yielded by the semantic textual similarity, suggesting this measure allows for an independent test of the hypotheses. The ready availability of the database and querying tools may serve to facilitate replication. A calibration exercise of the pairwise measure was run on separate picture-naming data, reported next, to complement an earlier calibration exercise reported in Schepman et al. (2018) for semantic textual similarity.

Pre-calibration of pairwise latent semantic analysis. In Schepman et al. (2018) a calibration analysis was reported in which picture naming data from 30 participants responding to 28 pictures from Cycowicz, Friedman, Rothstein, and Snodgrass (1997) were put through the same pairing and semantic textual similarity analysis protocol as the main data, using the randomization process outlined above. These pictures were unambiguous black-and-white line drawings of single objects (e.g. bottle, rabbit, carrot) and the participants were simply asked what the picture was. As previously reported in Schepman et al. (2018), this had shown a mean semantic similarity score for the experimental (within-items) pairings of .87 (SD = .24, SE = .01, 95% CI [.86, .89]), while the baseline control (across-items) pairings had shown a mean semantic textual similarity score of .08 (SD = .19, SE = .01, 95% CI [.07, .09]). There had been a significant difference between the experimental pairings and baseline control pairings on a Mann-Whitney test, Z = 33.63, W = 15557, $p < .001$, BF_{10} = 2.116e +19, r = .86 (note that r is used as an effect size).

Here, we report a new calibration analysis, using the same random pairings as for the semantic textual similarity measure reported previously, but using pairwise latent semantic analysis measures. The aim was to establish whether this measure would also show a significant difference between the sample of experimental and baseline control pairs taken from Cycowicz et al. (1997). We also checked whether the latent semantic analysis scores correlated with the semantic textual similarity scores.

Results showed that for the experimental pairings, the mean latent semantic analysis score was .82 (SD = .31, SE = .01, 95% CI [.80, .84]), while the baseline control pairings showed a mean latent semantic analysis score of .06 (SD = .17, SE = .01, 95% CI [.05, .07]). These sets of scores differed significantly on a Mann-Whitney test, Z = 31.67, W = 27355, $p < .001$, BF_{10} = 1.609e +18, r = .81. The two-tailed Kendall's tau-b correlation between the two measures, pooled across experimental and baseline control pairings, was very strongly positive, $\tau_b = .79$, $p < .001$, $BF_{10} = \infty$. These calibration analyses suggest that pairwise latent semantic analysis scores are a useful additional measure to provide converging computationally-derived quantitative evidence in our main study. The strong correlation may make it seem as if the measures might be almost identical, but the examples cited earlier showed this not to be the case. This matter will be revisited with reference to the main data. Please note that these calibration data are not directly comparable in the same inferential statistical analysis to the data from our study, because of the use of different participants, image-types and numerical parameters. The outcome from the calibrations simply serve as points of reference in the interpretation of the level of shared meaning observed in the main data.

Type-token ratio. The type-token ratio captures the number of unique words in a piece of text. It is calculated by counting the total number of words (tokens) and the number of unique words (type) in a corpus of text, and taking a ratio by dividing type by token. This can be converted to a percentage. If the text contains a high level of lexical diversity, then the type-token ratio is high, while a low ratio shows low lexical diversity. The type-token ratio has been put through a range of empirical tests for reliability (see e.g. Johansson, 2008) and it has been found that it is sensitive to the length of the text, with increased text length having a lowering effect on the type-token ratio. However, the effect of text length of the type-token ratio is more pronounced in certain ranges. In our analyses, mean text length of the combined participant responses per artwork overall was 163 words (SD = 37 words, range 104 – 262). By art type, the text length was 143 words (SD = 23 words) for abstract artworks, and 182 words (SD = 37 words) for representational artworks. Empirical studies (Koizumi, 2012; Šišková, 2012) suggest that the type-token ratio is sufficiently stable in this range of lengths, and as good as or better than alternative estimates of lexical diversity. We therefore used this measure as the best estimate of lexical diversity in our data.

Statistical analyses of computational measures

Statistics were calculated for the three computational measures, semantic textual similarity, pairwise latent semantic analysis scores, and matrix latent semantic analysis scores. Descriptive statistics are displayed in Table 1. In relation to the first hypothesis, that there would be shared meaning, we noted that the means for the experimental randomizations exceeded those for the control conditions of their art type, but the means for the experimental randomizations were relatively low, particularly compared to the calibration picture naming data used to aid interpretation. Thus, any significant sharing of meaning above baseline control would need to be considered against a background of a very high proportion of idiosyncratic semantic elements. In relation to the second hypothesis, that there would be different levels of shared meaning as a function of art type, we also noted that the discrepancy between the baseline control randomization and the experimental randomization was greater for the representational than the abstract artworks.

Table 1: Descriptive Statistics

		Semantic Similarity			Pairwise Latent Semantic Analysis			Matrix Latent Semantic Analysis		
Art Type	Random-ization	Mean	SEM	CI	Mean	SEM	CI	Mean	SEM	CI
A	BC	.051	.003	.045, .056	.018	.003	.012, .023	.022	.0006	.020, .023
	E	.097	.005	.086, .107	.059	.005	.050, 0.69	.058	.0018	.054, .061
R	BC	.068	.003	.062, .074	.030	.003	.024, .036	.032	.0008	.031, .034
	E	.149	.006	.137, .161	.107	.006	.095, .119	.104	.0038	.098 .112

Table 1 Note: Means, Standard Errors of the Mean (SEM) and 95% Confidence Intervals (CI) for the Semantic Similarity, Pairwise Latent Semantic Analysis, and Matrix Latent Semantic Analysis scores for the two different art types, Abstract (A) and Representational (R), in their two randomizations, namely Baseline Control (BC), randomized with any other response from the same art type, and Experimental (E), randomized within image.

Inferential statistical analysis. We took a mixed approach. As a first approximation, we analyzed disaggregated data for the key differences with nonparametric tests for differences (Mann-Whitney) due to nonnormality of the raw data and to maintain comparability with earlier work. We then also used Linear Mixed Models, whose raw data are not required to be normal, and the normality of residuals is argued to be of little importance (Gelman & Hill, 2006, p. 46). Linear mixed models were used to inspect the fixed and simple main effects and interaction while including the random intercepts of items (nested within art type, crossed with randomization) and participants (crossed with both factors). Attempted models with random slopes were rejected due to singular fits. Deviation coding was used, degrees of freedom were estimated using the Satterthwaite approximation. We report both pairwise measures before the matrix latent semantic analysis measure.

Pairwise measures. In a test of the first hypothesis, the effect of randomization (Baseline Control vs. Experimental) pooling across both art types was significant for both pairwise measures using Mann-Whitney tests, $Z = 9.70$, $W = 1.516e+6$, $p < .001$, $BF_{10} = 1.258e+7$, $r = .15$, for Semantic Textual Similarity, and $Z = 11.43$, $W = 1.448e+6$, $p < .001$, $BF_{10} = 1.415e+9$, $r = .18$, for Latent Semantic Analysis, showing a significant difference between the two randomizations, in evidence of shared meaning exceeding baseline.

The results from the baseline control randomizations were compared to those from the experimental randomizations for each art type separately using further Mann-Whitney tests. For abstract art, both measures showed a significant difference between baseline and experimental randomizations, $Z = 5.03$, $W = 395014$, $p < .001$, $BF_{10} = 7488$, $r = .11$, for the semantic similarity score, and $Z = 6.15$, $W = 378618$, $p < .001$, $BF_{10} = 1885$, $r = .14$, for the pairwise latent semantic analysis score. These differences were also significant for representational art, $Z = 9.06$, $W = 357642$, $p < .001$, $BF_{10} = 2.581e+7$, $r = .21$, for semantic similarity scores, and $Z = 10.10$, $W = 343676$, $p < .001$, $BF_{10} = 1.036e+8$, $r = .23$, for the pairwise latent semantic analysis scores. This showed that for each art type, the experimental randomization showed higher shared semantics scores than the baseline control, with larger effect sizes for representational than abstract art.

In the Linear Mixed Model for Semantic Textual Similarity, there were significant effects for art type, $F\ (1, 38.1) = 13.9$, $p < .001$, randomization, $F\ (1, 3746.3) = 202.8$, $p < .001$, and a significant interaction between the two factors, $F\ (1, 3746.3) = 15.1$, $p < .001$. Items showed an ICC of .03, participants .01. Both art types showed significantly higher semantic similarity in the experimental than the baseline control condition, both $p < .001$, and 95% CI [.02, .03] for abstract, [.03, .05] for representational art. The simple main effect of art type was significant for the experimental randomization, $p < .001$, 95% CI [.02, .04], but not the baseline control randomization, $p = .094$, 95% CI [-.001, .012]. The model with fixed-effects and interaction had a Marginal R^2 of .07, while the model with random factors had a Conditional R^2 of .11.

For pairwise Latent Semantic Analysis, the pattern was similar, with significant effects for art type, $F\ (1, 38) = 7.8$, $p = .008$, and randomization, $F\ (1, 3746.9) = 185.33$, $p < .001$, and a significant interaction between the two factors, $F\ (1, 3746.9) = 16.46$, $p < .001$. The ICC for items was .05, for participants .002. Simple main effects of art type were significant for the experimental randomization, $p < .001$, 95% CI [.01, .04], but not for the baseline control randomization, $p = .29$, 95% CI [-.005, .02]. Both art types showed significantly higher semantic similarity in the experimental than the baseline control condition, $p < .001$, 95% CI [.01 ,03] for abstract, [.03, .04] for representational art. The Marginal R^2 was .06, Conditional R^2 was .11.

Matrix measure. As above, we checked the key comparisons using Mann-Whitney tests. These showed a significant main effect of randomization $Z = 27.33$, $W = 236608$, $p <$

.001, BF_{10} = 1.465e +14, r = .44. Simple main effects of randomization were consistently present for each art type, Z = 18.00, W = 236608, p < .001, BF_{10} = 6.313e +12, r = .41, for abstract,

and Z = 21.67, W = 201419, p < .001, BF_{10} = 1.046e +15, r = .49, for representational art.

In the Linear Mixed Model the main effect of art type was significant, F (1, 38) = 6.95, p = .013, as was the main effect of randomization, F (1, 3746.5) = 846.35, p < .001, and the interaction between the two factors, F (1, 3746.5) = 93.41, p < .001. The ICC for items was .26, for participants .02. Simple main effects of randomization were significant for each art type (both p < .001, 95% CI [.016, .02] for abstract, [.033, .039] for representational art), while the simple main effect of art type was only significant in the experimental randomization (p < .001, 95% CI [.01, .03]) and not the baseline control randomization (p = .34, 95% CI [-.005, .016]). For this model, the Marginal R^2 was .18, Conditional R^2 was .40. Analysis of this additional measure confirms and strengthens the findings from the pairwise measures reported above. The notable difference is that variance partitioned to art type was lower than in the pairwise models, while variance attributed to items was higher. In addition, the R^2 measures were higher in the matrix measure than the pairwise measures. Overall, the effects were stable across all three computational measures, with all showing greater shared meaning for representational than abstract art, but with abstract art also showing shared meaning above baseline.

Statistical analysis of type-token ratios. Based on by-items analyses, in which all text for one item was treated as a single input to the algorithm, results revealed that representational artworks (M = 64.1%, SEM = .02%, 95% CI = 60.8%, 67.4%) showed a lower type-token ratio than abstract artworks (M = 70.4%, SEM = 1.1%, 95% CI = 67.9%, 73.0%). This difference was significant on an independent-samples t-test (used because the assumptions of a normal distribution and homogeneity of variance were met for these data), t (38) = 3.14, p = .003, d = 0.99, BF_{10} = 11.964, as well as a nonparametric Mann-Whitney test, used for comparability to previously reported analyses, Z = 2.79, W = 303, p = .005, BF_{10} = 7.024, r = .44. This analysis shows that participants used more divergent words to describe the meaning of abstract than representational artworks, which suggests that meaning was shared to a larger extent for representational than abstract artworks.

Correlational checks for independence of quantitative measures

Correlation coefficients were calculated to evaluate the possibility that the four measures were simply redundant transformations of each other. This was done by-items, to match the type-token ratio measure. The correlations showed that the measures corresponded in the predicted directions, with all semantic similarity measures correlating negatively with type-token ratio, τ_b = -0.459, p < .001, BF_{10} = 946.2 with semantic similarity, τ_b = -0.397, p < .001, BF_{10} = 115.1 with pairwise latent semantic analysis, and τ_b = -0.456, p < .001, BF_{10} = 861.5 with matrix latent semantic analysis. Semantic similarity correlated positively with both variants of latent semantic analysis, τ_b = 0.631, p < .001, BF_{10} = 1.628e +6 for pairwise, and τ_b = 0.582, p < .001, BF_{10} = 156031 for matrix, while the two latent semantic analysis measures also correlated positively with each other, τ_b = 0.633, p < .001, BF_{10} =

1.851e +6. The measures did not correlate so strongly that they made each other redundant as calibrating measures.

Discussion

The quantitative data indicated that both abstract and representational art showed shared meaning, and that the difference in shared meaning between the two art types was significant, but relatively modest. Correlation analyses showed correspondences in the predicted direction, and suggested that the measures were suitably independent to serve the purpose of triangulation. Fuller implications are discussed in the General discussion.

Although the statistics reported so far suggest that there was shared meaning for both types of art, confidence in these results could be further strengthened with evidence of the actual shared meaning. Providing this was the aim of the next set of analyses. These analyses were aimed at identifying the core shared meaning, while fully acknowledging the wider range of many idiosyncratic meanings that were also provided.

Content-related analysis methods

We used three techniques with complementary strengths to provide estimates of the contents of the shared meaning. The aim was to distil the shared content for each artwork. The focal question was whether a distinct core shared meaning could be found for each artwork that did not overlap with the shared meaning of other artworks. A key aim was to evaluate whether this applied to abstract and representational artworks. As stated before, this core shared meaning existed alongside the great variation in meaning across participants that was also observed in the quantitative data, but the primary aim was to identify the shared component of the meanings provided by the participants, to make this transparent by way of additional evidence of the existence of a shared core in the meanings.

Human coding. Two human coders (the authors) created summary descriptions for each artwork based on all participants' responses to that artwork. Initially, the coders read through the response set for each artwork independently, with reference to the texts but not the images, identifying themes for each artwork. Note that while AS had high levels of familiarity with the images and was therefore able to access mental representations of them during coding, PR had not seen the images for quite some time and was not able to recall them during coding. Separately, the two coders developed codes from the themes, and placed codes next to participants' responses in a spreadsheet. They then distilled a summary description for each artwork that they felt was representative of the views expressed by the majority of the participants.

Semantic similarity and pairwise latent semantic analysis were run on the summaries produced by the independent coders, to estimate the level of agreement between the two coders, taking their paired descriptions for each item as input ($N = 40$). Semantic textual similarity showed a mean of .46 ($SEM = .01$, 95% CI .43, .49) and pairwise latent semantic analysis a mean of .43 ($SEM = .03$, 95% CI = .38, .49). These values were much higher than the values on these measures observed for the participants' responses, and about half-way between the highest (experimental) and lowest (baseline control) scores in the picture naming calibration, suggesting good, but not trivially matched inter-coder similarity of the constructed shared meanings. This is compatible with the inherently subjective nature of this process. Statistically, based on Mann-Whitney tests, the coders showed no difference in levels of agreement between art types (abstract, representational), as measured through semantic similarity scores (Abstract: $M = .45$, $SEM = .02$, Representational: $M = .46$; $SEM = .02$; $Z = .08$, $W = 203$, $p = .94$, , $BF_{10} = 0.292$, $r = .01$,) and pairwise latent semantic analysis

(Abstract: M = .48, SEM = .04, Representational: M = .39, SEM = .05; Z = 1.48, W = 254.5, p = .14, BF_{10} = 0.783, r = .23).

Once the independent coding had been completed, the coders agreed a shared summary of the responses for each artwork. The independent summaries were considered and discrepancies were resolved with reference to the original responses and the independent coding. Discussions concerned the inclusion of terms, and the order in which elements occurred. Frequently present or important themes were listed earlier, and less frequent themes later. Responses mentioned by just one participant were not included (Augustin, Wagemans, & Carbon, 2012). An example item from the coding and agreement document can be seen in Supplemental Materials 2.

Frequency analysis. In this analysis, all words with a frequency of 2 or greater were listed in order of frequency by running all the words into of each set of pooled participant responses, separately for each artwork, through www.textalyzer.net. This software detected all words with a minimum of 1 character, and applied its default English "stop list", which filters common function words such as "a", "the", "of" etc. We checked for any meaningless output (e.g. "don" separated from "t") and removed these if they were function words. We additionally removed most function words that had not been filtered out by the stop word filter (e.g. "you", "too"), but, due to the potential for content being present, we retained prepositions that had passed through the stop word filter (e.g. "through"). We spot-checked the list of stop words filtered out to see if any content words were removed, but did not detect such problems.

Latent semantic analysis: nearest neighbors. We used latent semantic analysis (Latent Semantic Analysis @ CU Boulder, 2018) to select nearest neighbors of the set of descriptions given to each image, with an aim of having the system detect latent links between the responses provided by the participants. We set the software to produce the nearest five neighbors as trials showed larger neighbor sets yielded no greater insights. We chose the "pseudodoc" setting which uses entropy weighting in the semantic spaces to prioritize content over function words.

Content-based results and discussion

The output from the three methods of content-based analysis are shown in Supplemental Materials 2. A distinct description, word set, or set of nearest neighbors could be seen for each artwork. Mostly these did not overlap with those for other artworks. This applied similarly to abstract and representational art. This strengthens the findings from the quantitative analysis.

It was interesting to see high levels of shared metaphoric interpretations for the abstract artworks. For example, *Axl II* (by László Moholy-Nagy) depicts two pairs of parallel intersecting lines. A remarkable number of participants saw in this a crossroad (junction), and interpreted this as a metaphor for a decision or choice that needed to be made. Similarly, *Red and Black Composition No. 5* (by Burhan Dogancay) is an abstract painting which resembles a collage with multiple layers of paper, some peeling to reveal other layers. Participants showed a tendency to interpret this as "a metaphor for fragmentation or ruin; cover-up or layering that can be peeled away to reveal hidden depths" (portion of human coding, item 16A). Similar shared metaphoric interpretations can be seen in evidence for many of the abstract artworks.

One might think that interpretations for representational artworks may contain only surface descriptions of visual elements, with metaphoric interpretations being reserved for abstract artworks. However, this was not the case. For example *Lone wolf* (by Alfred Kowalski) shows a wolf in a snowy landscape, but participants interpreted beyond the visual image to suggest “representing sad isolation or ex-communication” (item 29R). Likewise, for *Cold Cream* (by Mary Pratt), a portrait of a woman with towel-wrapped hair and cream on her face, participants inferred that a deeper meaning was that the subject sought to “mitigate ageing or hide flaws or true identity, seeking perfection” (item 22R). Thus, the interpretations contained some references to visual elements, and some to interpretations beyond the visible elements, and this applied to both art types. This interpretative process differs qualitatively from more basic picture labelling for the purpose of capturing visual content, which both humans (Cycowicz et al, 1997) and machines (Farhadi et al., 2010) are able to do.

The nearest neighbor latent semantic analysis generally yielded useful results that mostly showed understandable correspondences to the outputs from the other two content-based analyses. Rarely, it yielded some words lists that seemed only tangentially related to the artwork and the output from the other analyses, but this is a well-documented by-product of latent semantic analysis (Landauer, Foltz & Laham, 1998).

At times, remarkably insightful latent nearest neighbors were produced. For example, for item 6A (*Untitled* by Ludwig Hirschfeld Mack), the human coder description included “A starry outer-space feel”, while the nearest neighbor analysis produced “Baade” (an astronomer). Further, *Cats (rayist percep.[tion] in rose, black, and yellow)*, a rayist abstract image of cats, characterized by multi-directional angular lines interpreted by some viewers as “triangles”, yielded the semantic neighbor “Sikyátki”, a Hopi village (now an archeological site) known for its pottery with geometric decorative markings, found in shards during excavations. This is a striking link. *Unexpected Visitors* (by Fred Bervoets), yielded

“Dubuffet”. Jean Dubuffet was a French artist whose style and preference for depicting ordinary life overlaps with Bervoets’s. It seems remarkable that latent semantic analysis of the verbal responses by nonexpert viewers can link these two artists. Overall, latent semantic analysis yielded useful nearest neighbors for most artworks, which served the purpose of triangulating the data, while at times providing interesting food for thought, suggesting that deeper explorations of this analysis technique may be fruitful in future work.

General discussion

We have provided quantitative evidence that there was significantly detectable shared meaning across viewers for both representational and abstract art, with greater shared meaning for representational art, but with a relatively modest difference between the two art types. We base this inference on the robust and triangulated evidence from the four non-overlapping quantitative measures, which each supported the same result. The finding of greater shared meaning for representational art conflicts with the view that, because abstract art uses a universal language (and representational art does not), meanings for abstract art should be shared to a greater extent than those for representational art. Instead, our findings support the notion that representational art generates meanings that are more shared across viewers than those for abstract art, extending recent related empirical work (Leder et al. 2016, Schepman et al, 2015a, 2015b, Vessel & Rubin, 2010). The extension of these earlier findings is significant and important, because this is the first study to look at viewer-generated shared meaning directly, rather than associations or justifications (Schepman et al. 2015b, 2018), researcher-generated meaning (Kuipers et al., 2018), or shared image associations as inferred from shared taste (Vessel & Rubin, 2010).

While there were differences in the extent to which meaning was shared across art types, there was evidence of shared meaning for both art types. However, it must also be noted that in both cases, the mean semantic similarity and latent semantic analysis scores were relatively low in comparison to calibration analyses of picture naming data. This shows that there were many idiosyncratic meanings for both types of artworks. However, despite this idiosyncrasy, there was also strong and statistically robust evidence for a shared “core” in these diverse meanings.

These shared core meanings were made transparent by the content-based analyses. The shared meanings tended to include reference to visual elements, but also metaphoric interpretations that linked the visual information to prior knowledge (see Leder et al. 2004). While on some level these data could mean that there may be a universal language that artists use when producing a painting, and viewers “understand” this language directly, many alternative explanations could be generated that are more compatible with the patterns in the data. As observed by Crowther (2017), the artist may intend to depict something, but viewers are at liberty to impose their own meanings on the artworks. For the artwork to act as a token of “language” that viewers simply “understand”, we would have expected much higher means for the shared meaning, close to those of the calibration analyses based on picture-naming data, but our means were much lower. Thus, the observation that the meanings that viewers generated were shared may not be due to the artists’ imposition of meanings being automatically received by the viewer. Instead, these may arise via interpretative processes in the viewers. We did not probe the exact nature of the interpretative processes in this research project, but an interesting possibility was put forward by Parsons (2015), who builds on prior philosophical work. He proposes that art interpretations, including visual metaphors, derive from general ideas that are mapped to the artwork. Each general idea has its own properties and connotations, and the process of interpretation involves selecting from these many connotations and properties in a specific context. This way of conceptualizing interpretation, as opposed to connotations (or associations) would provide an excellent framework for studying interpretative processes in more detail in future empirical research.

It is important to evaluate whether there may be alternative interpretations of the findings. First of all, it may seem trivially obvious that meanings for representational artworks are more shared, because viewers may have engaged in something akin to picture naming (i.e. naming the objects in the image), which is something that cannot be done for abstract artworks. However, the data show that this was not the case. As just stated, the means for computational measures capturing shared meaning were much lower for the interpretations of artworks than for the calibration with picture naming data, while the content-based analysis clearly showed metaphoric alongside perceptual (and other) interpretative elements for both art types. Further, it might be considered that viewers may have simply named the visual features of abstract artworks, such as color and form, which may have led to an above baseline level of shared meaning for abstract art. However, the analysis of the contents shows that this cannot be an explanation for the shared core meaning, because viewers provided many interpretive elements in the responses to abstract art. In addition, such a situation would have yielded a very high baseline mean for abstract artworks, because each artwork would have been given similar verbal responses, but this was not seen in the data.

It is possible to consider the brief responses that our participants provided not to be representative of ordinary art interpretation. First of all, it may be that viewers do not routinely generate meanings when viewing artworks. This would mean that the results can only generalize to situations in which this occurs spontaneously or as a result of some request. Such a request may therefore trigger more meaning than would otherwise arise. The findings of Kuipers et al. (2018) suggest that meanings may arise spontaneously, but we acknowledge that this may not always be the case. Secondly, it may be that viewers felt restricted in their time to generate meaning compared to the longer meaning-making experience described by e.g. White (2011) and Specker, Tinio, and van Elk (2018). However, the duration of the experiment overall allowed for a similar length of viewing and interpretation time as the average time taken by naïve viewers visiting art galleries, estimated as 27 seconds (Smith & Smith, 2001) or 33 seconds on first pass (Carbon, 2017, note 17 seconds longer with re-visits factored in). This is much more time than more time-restricted experimental approaches that infer the presence of semantic interpretations, e.g. 1 second displays in Vessel and Rubin (2010), and 6.3 seconds for the trial sequence in Kuipers et al, (2018). Thus, it seems that the time available to our participants was sufficient to generate a meaningful semantic response. It may also be thought that our instruction to provide a word, phrase or short sentence restricted the participants' freedom of expression. However, if there had been an urge to express more text than the level provided, we would have seen a preponderance of sentence-length responses, but we often saw relatively short responses consisting of a few words, suggesting that there was no artificial ceiling on participants' freedom to express perceived meanings. Instead, it seems that participants performed naturally within the response length restrictions.

It is important to consider the validity of the content-based analysis. Human coders showed a good level of correspondence in their interpretations prior to agreement. The agreed human coding statements corresponded well with the word frequency analysis. The two formed useful complementary techniques with balanced strengths and weaknesses. The word frequency technique was fully objective but lacked any form of interpretation or integration, while the human coding technique was more interpretative, which had the disadvantage of being potentially more subjective (despite the protections afforded by using two independent coders), but benefited from human interpretation during the summarization process. Latent semantic analysis using nearest neighbors most often further corroborated general themes identified by the other two methods. The fact that the results from these three fully independent analysis processes showed a strong correspondence suggests that they provided useful data triangulation. Thus, the insights provided by these combined methods can be relied upon as being valid.

A shortcoming of our research, shared with the research of many others, is that the paintings were of reduced size, lacked any three-dimensionality that may come from textures, or contexts that may come from e.g. gallery environments. Two aspects are important to consider. First, it is possible to think that this shortcoming may have affected one art type more than the other, making this a confounding factor. For example, some abstract artworks may be more likely to evoke psychological reactions in their full size, rather than as small visual objects, because only then, perhaps, does the artwork convey the experience fully. However, a similar argument could apply to representational art. Examples for artworks of both types, in authentic settings and full size, triggering overwhelming emotions are documented in Elkins (2001), suggesting that overwhelming emotions do not only occur in response to either abstract or representational artwork, but in response to both. Thus, it seems difficult to sustain an argument that the small size of the images had a confounding effect by affecting one art type more than the other. Second, it may be that findings based on small images do not generalize to images in their original size. The issue is discussed extensively in Pelowski, Forster, Tinio, Scholl, and Leder (2017), who argue that lab studies are valuable, but that psychological processes in response to artwork are likely to differ in museum or gallery settings. Specker et al. (2017) support this argument with data. This is a possibility that we cannot address based on the current data, but an issue shared with many lab studies. Future research on this issue would be very interesting.

Using our methods, meaning becomes a reliably extractable dimension of art experience that may play an important role in larger multidimensional models of art appreciation (e.g. Leder et al., 2004; Pelowski, Markey, Forster, Gerger, & Leder, 2017; Pelowski, Markey, Lauring, & Leder, 2016). In the related field of computer vision, it has proven fruitful to model semantic information into complex neural nets, where this has served to enhance the accuracy of machine-generated aesthetic evaluations (Kao, He, & Huang, 2017). It may be that models of human art evaluation could be enriched by including this important dimension with more precision and detail than has thus far been possible. Having rigorous methods to analyse and extract data, such as the ones used in the current paper, may help build and test these multidimensional theoretical models.

We have addressed the constrained questions we set out to answer, but this research generates many questions that would benefit from further research. For example, it would be very interesting to explore the mapping between visual elements, color, form and different elements of semantic content. In addition, it would be valuable to probe more deeply into the processes that lead to shared meaning. Exploring the links between shared meaning and shared liking would also be valuable.

Conclusion

The research provides strong quantitative and content-based evidence for the idea that there is a core of shared meaning in the interpretation of artworks across multiple viewers, alongside varied idiosyncratic meanings. This is the first study that has directly measured shared meaning in response to artworks. Contradicting the predictions of those who see abstract art as conveying meaning that is more universal than that conveyed by representational art, we found greater shared meaning for representational art, in line with other recent empirical evidence. In addition, both art types show evidence of a core shared meaning, replicating and extending prior work. The new methodologies used in this article has made the complex topic of shared semantic interpretations of artworks more amenable to future research, aided by the use of clearly documented protocols. The techniques developed are applicable in many other settings, and serve to operationalize the traditionally difficult-to-study area of subjective meaning with high levels of scientific rigor. Modeling meaning into future theoretical and computational models of image processing is likely to enhance the performance of these models.

References

Arvidson, K, & Dahlström, P. (Curators) (2018). The universal language. Post-war abstract art. Gothenburg Art Museum. June 13-11 November 2018. [Accessed 15 June 2018 via http://goteborgskonstmuseum.se/en/exhibitions/the-universal-language/]

Augustin, M. D., Wagemans, J., & Carbon, C. C. (2012). All is beautiful? Generality vs. specificity of word usage in visual aesthetics. *Acta Psychologica*, *139*(1), 187-201. https://doi.org/10.1016/j.actpsy.2011.10.004

Brinkmann, H., Commare, L., Leder, H., & Rosenberg, R. (2014). Abstract art as a universal language?. *Leonardo*, *47*(3), 256-257.

Carbon, C.-C. (2017). Carbon, C. C. (2017). Art perception in the museum: How we spend time and space in art exhibitions. *i-Perception*, *8*(1), 2041669517694184.

Crowther, P. (2017). *What Drawing and Painting Really Mean*. New York, NY: Routledge.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science*, 41(6), 391-407.

Elkins, J. (2001). *Pictures and Tears: a history of people who have cried in front of paintings*. New York, NY: Routledge.

Farhadi A., Hejrati, M., Sadeghi, M.A., Young, P., Rashtchian, C., Hockenmaier, J., & Forsyth, D. (2010) Every Picture Tells a Story: Generating Sentences from Images. In: Daniilidis K., Maragos P., Paragios N. (eds) *Computer Vision – European Conference on Computer Vision ECCV 2010 Lecture Notes in Computer Science, Vol. 6314, pp 15-29*. Berlin, Heidelberg: Springer, https://doi.org/10.1007/978-3-642-15561-1_2

Galluci, M. (2019). GAMLj: General Analyses for the Linear Model in Jamovi. https://mcfanda.github.io/gamlj_docs/index.html

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge university press.

Günther, F., Dudschig, C., & Kaup, B. (2016). Latent semantic analysis cosines as a cognitive similarity measure: Evidence from priming studies. *The Quarterly Journal of Experimental Psychology, 69*(4), 626-653.

Haftmann, W. (1954). *Malerei im 20. Jahrhundert 1: Eine Entwicklungsgeschichte.* Munich: Prestel.

Han, L., Kashyap, A., Finin, T., Mayfield, J., & Weese, J. (2013). UMBC EBIQUITY-CORE: Semantic textual similarity systems. Retrieved from http://ebiquity.umbc.edu/paper/html/id/621

Heinen, D. J., & Johnson, D. R. (2018). Semantic distance: An automated measure of creativity that is novel and appropriate. *Psychology of Aesthetics, Creativity, and the Arts*, 12(2), 144-156. http://dx.doi.org/10.1037/aca0000125

Hirsch, S. (n.d.). Adolph Gottlieb and Art in New York in the 1930s. The Adolph & Esther Gottlieb Foundation. [Accessed 14 June 2018, via cached website archive of https://www.gottliebfoundation.org/s/Hirsch_Essay.pdf]

Hume, N. (2016). Creating the Ideal: František Kupka's Social Reform and Anarchist Abstraction. In Wünsche, I. & Wiebke Gronemeyer, W. (Eds.) (2016). *Practices of Abstract Art: Between Anarchism and Appropriation*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 31-52.

Jamovi project (2018). jamovi (Version 0.9.6.0) [Computer Software]. Retrieved from https://www.jamovi.org

JASP Team (2018). JASP (Version 0.8.6 and version 0.9.0) [Computer software]. Retrieved from https://jasp-stats.org/).

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: a developmental perspective. *Working Papers in Linguistics*, *53*, 61-79.

Kao, Y., He, R., & Huang, K. (2017). Deep aesthetic quality assessment with semantic information. *IEEE Transactions on Image Processing*, *26*(3), 1482-1495. https://doi.org/10.1109/TIP.2017.2651399

Koizumi, R. (2012). Relationships between text length and lexical diversity measures: Can we use short texts of less than 100 tokens? *Vocabulary Learning and Instruction*, *1*(1), 60-69. http://dx.doi.org/10.7820/vli.v01.1.koizumi

Kuipers, J. R., Jones, M. W., & Thierry, G. (2018). Abstract images and words can convey the same meaning. *Scientific reports*, *8*(1), 7190. https://doi.org/10.1038/s41598-018-25441-5

Museum of Modern Art (1957). German art of the twentieth century. The Museum of Modern Art in collaboration with the City Art Museum of St. Louis, Missouri: Distributed by Simon & Schuster. [Haftmann, W., Hentzen, A. & Lieberman, W.S., with Ritchie, A.C. (Ed.)]. Exhibition URL: www.moma.org/calendar/exhibitions/2815

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, *25*(2-3), 259-284.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*(2), 211.

Latent Semantic Analysis @ CU Boulder (2018). [Accessed 15 June 2018 via http://lsa.colorado.edu/]

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*, 489-508.

Leder, H., Goller, J., Rigotti, T., & Forster, M. (2016). Private and shared taste in art and face appreciation. *Frontiers in human neuroscience*, *10*, 155. https://dx.doi.org/10.3389%2Ffnhum.2016.00155

Martin, I. (2018). Benchmarking widening participation: how should we measure and report progress? Higher Education Policy Institute Policy Note 6 [http://www.hepi.ac.uk/wp-content/uploads/2018/04/HEPI-Policy-Note-6-Benchmarking-widening-participation-FINAL.pdf retrieved April 27th 2018].

Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*, 38(11), 39-41. http://dx.doi.org/10.1145/219717.219748

Mondrian, P. (1937). Plastic art and pure plastic art. From Mondrian, P. (1945). *Plastic Art and Pure Plastic Art 1937 and Other Essays*, 1941-1943. [accessed 14 June 2018 via https://monoskop.org/File:Mondrian_Piet_1937_2000_Plastic_Art_and_Pure_Plastic_Art.pdf]

Müller, F. (2016). Ernst Wilhelm Nay: *Vom Gestaltwert der Farbe*—Artistic Concepts and Cultural Policy in Postwar Germany. In Wünsche, I. & Wiebke Gronemeyer, W. (Eds.) (2016). *Practices of Abstract Art: Between Anarchism and Appropriation*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 179-194.

Nakov, P., Popova, A., & Mateev, P. (2001). Weight functions impact on LSA performance. in: *Proceedings of the EuroConference Recent Advances in Natural Language Processing* (RANLP 2001) pp. 187–193. [Accessed 15 June 2018 via http://people.ischool.berkeley.edu/~nakov/selected_papers_list/nakov_RANLP01.pdf]

Nicholson, B. (1941). Typescript entitled, 'Notes on "Abstract" Art' by Ben Nicholson. Tate Gallery Archives, TGA 8717/3/1/7. [Accessed 14 June 2018 via http://www.tate.org.uk/art/archive/items/tga-8717-3-1-7/typescript-entitled-notes-on-abstract-art-by-ben-nicholson/1].

Parsons, M. (2015). Visual metaphors: Meaning, interpretation and culture. In: Fleming, M., Bresler, L., & O'Toole, J. (Eds.) (2015). *The Routledge International Handbook of the Arts and Education*. London: Routledge, pp. 68 – 75.

Pelowski, M., Forster, M., Tinio, P. P. L., Scholl, M., & Leder, H. (2017). Beyond the lab: An examination of key factors influencing interaction with 'real' and museum-based art. *Psychology of Aesthetics, Creativity, and the Arts, 11*(3), 245-264. http://dx.doi.org/10.1037/aca0000141

Pelowski, M., Markey, P. S., Forster, M., Gerger, G., & Leder, H. (2017). Move me, astonish me… delight my eyes and brain: The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Physics of Life Reviews*. http://doi.org/10.1016/j.plrev.2017.02.003

Pelowski, M., Markey, P. S., Lauring, J. O., & Leder, H. (2016). Visualizing the impact of art: An update and comparison of current psychological models of art experience. *Frontiers in Human Neuroscience*, 10. http://doi:10.3389/fnhum.2016.00160

Rodway, P., Kirkham J., Schepman, A., Lambert, J., Locke, A. (2016). The development of shared liking of representational but not abstract art in primary school children and their justifications for liking. *Frontiers in Human Neuroscience,* 10 (21).

Rothko, M. (2006). *The artist's reality: Philosophies of art*. New Haven: Yale University Press.

Schepman, A., Kirkham, J. A., Rodway, P., Lambert, J., & Locke, A. (2018). Shared meaning in children's evaluations of art: A computational analysis. *Psychology of Aesthetics, Creativity, and the Arts.* Advance online publication. http://dx.doi.org/10.1037/aca0000159

Schepman, A., Rodway, P., Pullen, S., & Kirkham, J. (2015a). Shared liking and association valence for representational art but not abstract art. *Journal of Vision*, 15, 11. doi:10.1167/15.5.11

Schepman, A. Rodway, P. & Pullen, S. J. (2015b). Greater cross-viewer similarity of semantic associations for representational than for abstract artworks. *Journal of Vision*, 15(14):12, 1–6.

Scherer, K. R. (2005). What are emotions? And how can they be measured?. *Social science information*, *44*(4), 695-729. https://doi.org/10.1177%2F0539018405058216

Schönbrodt, F. D., & Wagenmakers, E. J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic bulletin & review*, *25*(1), 128-142. https://doi.org/10.3758/s13423-017-1230-y

Schwartzmann, A. (2018). Universal abstraction throughout the twentieth century. Sotheby's at large. [Accessed 15 June 2018 via http://www.sothebys.com/en/news-video/blogs/all-blogs/sotheby-s-at-large/2017/05/universal-abstraction-throughout-the-twentieth-century.html]

Šišková, Z. (2012). Lexical Richness in EFL Students' Narratives. In Ciarlo C. & Giannoni D.S. (Eds.) *Language Studies Working Papers, Vol. 4,* pp. 26-36, Reading, UK: University of Reading. ISSN 2040-3461.

Smith J. K., & Smith L. F. (2001). Spending time on art. *Empirical Studies of the Arts*, 19 (2), 229–236, doi:10.2190/5MQM-59JH-X21R-JN5J.

Spears N., Brown T. J., Dacin P. A. (2006). Assessing the corporate brand: The unique corporate association valence (UCAV) approach. *Journal of Brand Management*, 14 (1), 5–19. https://doi.org/10.1057/palgrave.bm.2550051

Specker, E., Tinio, P. P., & van Elk, M. (2017). Do you see what I see? An investigation of the aesthetic experience in the laboratory and museum. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(3), 265- 275. DOI: 10.1037/aca0000107

Tinio, P. P., & Gartus, A. (2018). Characterizing the emotional response to art beyond pleasure: Correspondence between the emotional characteristics of artworks and viewers' emotional responses. In Christensen, J. F. & Gomila, A. (Eds.) *Progress in brain research*, 237, 319-342. Amsterdam: Elsevier.

Vessel, E. A., Maurer, N., Denker, A. H., & Starr, G. G. (2018). Stronger shared taste for natural aesthetic domains than for artifacts of human culture. *Cognition*, *179*, 121-131. https://doi.org/10.1016/j.cognition.2018.06.009

Vessel E. A., Rubin N. (2010). Beauty and the beholder: Highly individual taste for abstract, but not real-world images. Journal of Vision, 10 (2): 18, 1–14, https://doi.org/10.1167/10.2.18

Vo, N., & Popescu, O. (2016). A multi-layer system for semantic textual similarity. In A. Fred, J. Dietz, D. Aveiro, K. Liu, J. Bernardino, & J. Filipe (Eds.), Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2016) (pp. 56–67). Setúbal, Portugal: Scitepress, Science and Technology Publications, Lda. http://dx.doi.org/10.5220/0006045800560067

White, B. (2011). Private perceptions, public reflections: Aesthetic encounters as vehicles for shared meaning making. *International Journal of Education & the Arts*, *12*(LAI 2). [Retrieved 3 July 2018 from http://www.ijea.org/v12lai2/]

Wünsche, I. (2016). Wassily Kandinsky and František Kupka: Between Metaphysics and Psychophysics. In Wünsche, I. & Wiebke Gronemeyer, W. (Eds.) (2016). *Practices of Abstract Art: Between Anarchism and Appropriation*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp 11-30.

Supplemental Materials (also deposited at http://dx.doi.org/10.1037/aca0000279.supp)

Supplemental materials 1: List of artworks used in the study

List of artworks shown: **Artist, title (translation of title), (year), URL.** Note that artworks 1-20 are abstract (A), and 21-40 are representational (R). Artworks were standardised for size by resizing the work so that the longest dimension was 567 pixels, and15 cm in length when printed.

1A. Giacomo Balla: *Velocitàastratta + rumore* (Abstract Speed + Sound) (1958)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/300

2A. Mark Rothko: Untitled (1947)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/3532

3A. Mark Tobey: Advance of History (1964)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/4057

4A. James Cant: Abstract with Aboriginal Motif (1947)

http://www.ngv.vic.gov.au/explore/collection/work/5340/

5A. Helen Frankenthaler: A Green Thought in a Green Shade (1981) http://www.wikiart.org/en/helen-frankenthaler/a-green-thought-in-a-green-shade-1981

6A. Ludwig Hirschfeld Mack: Untitled (c. 1964)

http://www.ngv.vic.gov.au/explore/collection/work/56367/

7A. Grahame King: Blue Poem (1975)

http://www.ngv.vic.gov.au/explore/collection/work/57040/

8A. Jan Andriesse: Ocean in Motion (1994)

https://www.stedelijk.nl/en/collection/95523-jan-andriesse-ocean-in-motion

9A. Jean Hélion: Composition (1934)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/1638

10A. Maria Helena Vieira da Silva: Danse (nd)

http://www.wikiart.org/en/maria-helena-vieira-da-silva/danse

11A. László Moholy-Nagy: AXL II (1927)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/2983

12A. František Kupka: Arabesque II (1925-26)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/2393

13A. Pieter Engels: Blind Gold Embrace (1984)

https://www.stedelijk.nl/en/collection/89628-pieter-engels-blind-golded-embrace-posthumous-to-p.-%28blind-golded-painted%29

14A. Yves Klein: La grande Anthropométrie bleue (ANT 105) (Large Blue Anthropometry (ANT 105)) (c. 1960)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/27

15A. Natalia Goncharova: Koshki [luchistoe vospr.{iiatie} rozovoe, chernoe i zheltoe] (Cats (rayist percep.[tion] in rose, black, and yellow) (1913) http://www.guggenheim.org/new-york/collections/collection-online/artwork/1500

16A. Burhan Dogancay: Red and Black Composition No. 5 (1974)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/1081

17A. Mark Bradford: Daddy, Daddy, Daddy (2001)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/10230

18A. Rudolf Bauer: Invention (Composition 31) (1933)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/462

19A. Barnett Newman: Concord (1949)

http://www.metmuseum.org/toah/works-of-art/68.178

20A. Clyfford Still: 1953 (1953)

http://www.tate.org.uk/art/artworks/still-1953-t01498

21R. Stanley Spencer: Wisteria, Cookham (1942)

http://www.wikiart.org/en/stanley-spencer/wisteria-cookham-1942

22R. Mary Pratt: Cold Cream (1983)

http://www.wikiart.org/en/mary-pratt/cold-cream-1983

23R. Alejandro Cabeza: Garden Joaquín Sorolla

http://www.wikiart.org/en/alejandro-cabeza/garden-joaquin-sorolla

24R. Eric Fischl: The Empress of Sorrow (1992)

http://www.wikiart.org/en/eric-fischl/the-empress-of-sorrow

25R. Andrew Wyeth: Wind from the Sea (1947)

http://www.nga.gov/content/ngaweb/press/exh/3456.html

26R. Felix Vallotton: Coucher de soleil à Villerville (Sunset at Villerville) (1917)

http://www.wikiart.org/en/felix-vallotton/to-lay-down-sun-with-villerville-1917

27R. Ernst Ludwig Kirchner: Selbstbildnis als Kranker (Self-portrait as a sick man) (1918)

https://commons.wikimedia.org/wiki/File:Ernst_Ludwig_Kirchner_Selbstbildnis_als_Kranker_1918-1.jpg

28R. Maurice Prendergast: On the Beach, St. Malo (c. 1907)

http://www.wikiart.org/nl/maurice-prendergast/on-the-beach-st-malo

29R. Alfred Kowalski (1849-1915): Lone Wolf (Date Unknown)

https://en.wikipedia.org/wiki/Alfred_Kowalski#/media/File:Alfred_Wierusz-Kowalski_-_Samotny_wilk.jpg

30R. Michael Bell: Never look back (2009)

http://www.wikiart.org/en/michael-bell/never-look-back-2009

31R. Xul Solar: Barreras Melódicas (Melodious Barriers) (1948)

http://www.wikiart.org/en/xul-solar/barreras-mel-dicas-1948

32R. Milan Kunc: Liebeserklärung (Declaration of Love) (1977)

http://milan-kunc.com/images/paintings/70-embarrassing-realism/1977%20-Liebeserklarung-Lovedaclration.200x240cm.jpg

33R. Jeffrey Smart: Civitella (2008)

http://www.wikiart.org/en/jeffrey-smart/civitella

34R. Boris Kustodiev: Village Holiday (Autumn holiday in the village) (1914)

http://www.wikiart.org/en/boris-kustodiev/village-holiday-autumn-holiday-in-the-village-1914

35R. Lucian Freud: Box of applies in Wales (1939)

http://www.wikiart.org/nl/lucian-freud/box-of-apples-in-wales-1939

36R. Rémi Blanchard: Fruit and Irises (1990)

https://www.mutualart.com/Artwork/FRUIT-AND-IRISES/53AA58CFBAA4FB6F

37R. Fred Bervoets: Onverwachts bezoek (Unexpected visitors) (2010)

http://cobra.canvas.be/cm/cobra/kunst/1.1003246

38R. David Salle: Comedy (1995)

http://www.guggenheim.org/new-york/collections/collection-online/artwork/3767

39R. L. C. Armstrong: Bow Bridge Boaters (2012)

http://www.marlboroughgallery.com/exhibitions/lc-armstrong

40R. Carel Willink: De Zeppelin (The Zeppelin) (1933)

http://www.museummore.nl/carel-willink

Supplemental Materials 2: Output from content-related analysis and coding process for sample item

Shared meaning for all artworks extracted by agreement by two human coders, word frequency counts, and latent semantic analysis (LSA) near neighbors. Please note the British spellings of some words reflect the raw data (e.g. colour). Item numbers include A for Abstract or R for Representational artworks. At the end of this document, there is an example of the raw responses, the two coders' codes, the two coders' summaries, the agreed summary and a note on the agreeing process, to serve as an illustration of the coding process, to aid transparency and replicability.

Item	English Title	Human coding	Words with frequency > 2	LSA Near Neighours
1A	**Abstract Speed and Sound**	Natural landscape with water or hills, range of mixed emotions; Energy, confusion, complexity, colours.	sky, colour, trees, different, blue, mixed, shapes, natural, nature, hills, harsh, beauty, Sydney, colours	crest, wave, teepees, waves, crests
2A	**Untitled**	Blurred, heat and burning, desert, desolation, death.	warmth, desert, life, lonely, brown, monster, burnt, fuzzy, heat, anger, blurry, death	heat, hot, death, radiators, groundhogs
3A	**Advance of History**	Chaotic, messy, hectic; perhaps a city, with people, possibly bikes. Feeling lost, physically or mentally, or entrapped, too close, connected. Entanglement of elements like wires. Viewing perspective like a map on a bird's eye view.	busy, city, messy, thoughts, chaos, houses, view, entanglement, winding, mind, chaotic, random, much, people, find, maze, autumn, layer, bird, eye, lost, map	city, streets, sparrows, expressways, buildings

4A	**Abstract with Aboriginal Motif**	Feelings of confusion from atypical forms, seeing faces, usually two, that are atypical; cartoonish or childlike style; Elements of sadness and negative states. Possibly altered state of consciousness.	confused, face, confusion, abstract, person, dazed, sadness	mental, psychoses, person, psychosomatic, artist
5A	**A Green Thought in a Green Shade**	Seeing an expanse of water (lake, pond) that is muddy suggesting pollution, dirt, triggering associations with illness, but also birds, leaves and nature, with green colours and animals that live in lake (e.g. frogs).	bird, green, pollution, fish, nature, water, colours, sea, lake, like, leaf, looks, murky, sad, through, peace, environment, dirty, life, birds, pond, being	perch, bird, herons, beak, sewage
6A	**Untitled**	A bright, colourful, light-filled city scape with a mystical, mythical 'other world', starry outer-space feel, looking beautiful, reminiscent of maze or map.	city, cities, stars, lights, like, candles, light, bright, colour, reflection, night, globalisation, world, space, different	city, cities, Baade, stars, brighter
7A	**Blue Poem**	Mountains and sky, possibly depicting falling, floating, flying, breaking, eroding, which may in turn mean journeys, freedom or release, goals or religious pathways.	away, sky, mountains, floating, breaking, freedom, sadness, piece, cliff, top, interesting, freely, mountain, falling	peaks, mountain, mountains, slopes, towering
8A	**Ocean in Motion**	Landscape with sand (beach / desert), wind or waves; wispy hair; peaceful and calm; simple, blank or nothingness; possible metaphor for a problematic state.	waves, simple, calm, blank, wind, natural, sea, hair, desert, flow, plain, sand	longshore, waves, crest, wave, suntans

9A	**Composition**	Connections represented by doors opening or closing providing new opportunities or choices. Colour, order, structural elements, tools.	doors, shapes, different, connections, places, lots, doorways, looks, order, tools, like, art, door, windows, structure, connection, quite, blocks, another, connected	shapes, doors, door, windows, open
10A	**Dance**	Busy, overcrowded city or slums, dark but with lightness, enclosed, feeling closed in; blocks.	dark, colours, busy, dull, city, life, light, blurred, buildings, abstract, darkness, little, squares, slums, areas, being, together	streets, dim, dark, suburbs, derailing
11A	**AXL II**	Seeing crossroads lead to several related concepts, paths to a particular goal, making choices / decisions, noughts and crosses, religion (from cross) and maths.	crossroads, cross, paths, religion, pathways, lead, decision, crossing, same, light, choices, different, space, choosing, right, roads, crosses, noughts, target	oncoming, decision, decisions, flashers, choice
12A	**Arabesque II**	Confusing maze-like lines, arrows, paths, or patterns, modern, futuristic, technology, masculinity, complexity; Egypt, harsh struggles.	lines, maze, pathways, confused, arrow, abstract, life, complex, modern, shapes	lines, contour, topographic, overlays, drawn
13A	**Blind Gold Embrace**	Mistakes, hiding them from a messy, tangled confusion, falling apart, illness, injury, with several negative emotions (sadness, anger, disgust).	sadness, chaos, mess, mistakes, covering, apart, thoughts, anger, person, hiding, injured, falling, destruction	anger, feelings, feeling, sadness, emotions
14A	**Large Blue Anthropometry**	Bird, butterfly or animal, freedom; possibly uncontrollable nature or messy chaos;	bird, freedom, nature, anger, running, ink, stuck, butterfly, being, blue, like, sky, free	bird, birds, eagle, beak, feathers

		waves, mountains, sky; mixed emotions.		
15A	**Cats**	Perception of darkness and related animals (black crows, horses, wolves) and concepts (death), eliciting feelings of evil, sharpness, jaggedness, pain; possibly contrasting with light, good, hope or life.	sharp, evil, life, light, darkness, black, dark, like, jagged, glass, pain, modern, bird, happy, good, wolf, nature, through, scary, man, parts, red	light, dark, incised, unslipped, Sikyatki
16A	**Red and Black Composition No. 5**	Ripped paper, possibly a metaphor for fragmentation or ruin; cover-up or layering that can be peeled away to reveal hidden depths.	paper, ripped, tearing, torn, apart, new, something, through, rips, life, anger, away, mind, together, layers, destroyed, pages, pieces, revealing, broken	paper, cardboard, piece, pieces, endsheets
17A	**Daddy, Daddy, Daddy**	Yellow colour, e.g. sunshine, corn, bees, sand, lights, with regularity, building blocks, wall, bricks, squares.	bright, happy, building, yellow, overlapping, blocks, lights, wall	brick, bright, tooled, bricks, masonry
18A	**Invention (Composition 31)**	Geometric shapes, possibly featuring a black hole. May represent dark gloom and disorganisation, or fun and hope. Diversity and difference.	shapes, together, abstract, fun, black, sharp, different, games, modern, maths, art, hole, darkness, funky, objects	shapes, artistic, cubists, artists, art
19A	**Concord**	Two lines that could be a metaphor for pathways or roads, division or duality, equality or opportunities, but may reflect simple nothingness.	lines, road, simple, door, straight, two, equal, parallel, calm, boring, never, duality, closed, double, window, yellow	lines, door, drawn, parallel, opened

20A	**1953**	Darkness, night time, ocean, abyss; the unknown, uncertainty, damage but elements of calmness.	darkness, blue, ocean, sky, night, sea, deep, damaged, wallpaper, calmness, lightning, unknown, abyss, down, light, peeling, first	ocean, dark, horizon, shine, dim
21R	**Wisteria, Cookham**	Summer or spring flowers and plants growing in an urban or suburban English garden, vibrant colours, beautiful.	flowers, summer, spring, growth, clouds, garden, nature, day, life, dark, bloom, beauty, bright, beautiful, happy, world, country, sky, calm, England, home, contrast, growing, natural, light, polluted	bloom, flowers, blossoms, buds, autumn
22R	**Cold Cream**	Woman preparing face with morning routine for the day to look beautiful, using treatments, cleansing, to mitigate ageing or hide flaws or true identity, seeking perfection, natural beauty; androgynous.	beauty, morning, ready, perfection, self, getting, woman, something, pamper, cover, age, strange, routines, hidden, cleansing, beautiful, cream, covered, face, body, covering, female, being, routine	sexism, rejoices, sex, orgasms, androgynous
23R	**Garden Joaquín Sorolla**	A garden abroad (Spanish, Moroccan, Tibetan) evoking peacefulness, tranquility, calmness and serenity; meditation.	garden, peaceful, plants, nature, nice, emptiness, serenity, peace, tranquil, summer, meditation, quiet, calm, relaxation, life, growth, gardens, urban	garden, flowers, roses, gardens, plants

24R	**The Empress of Sorrow**	Set in Oriental culture, possibly royal or wealthy; discord between a female, possibly an entertainer, and a man: negative emotions linked to relationship, jealousy, betrayal, hostility, power.	culture, wealth, cultural, authority, jealousy, carousel, fun, anger, royalty, life, like, woman, tension, art, power, argument	culture, shoguns, Manchuria, cultures, Vladivostock
25R	**Wind from the Sea**	Open window, representing bleak, scary, haunted, lonely abandonment or, in contrast, freedom, escape, tranquility.	loneliness, freedom, window, day, breeze, abandoned, house, cold, windy, ghost, haunted, loss, isolation, open, air, escape, alone, quiet	shutters, windows, windy, blew, house
26R	**Sunset at Villerville**	Sunset at a beach, peaceful, tranquil and warm, may be beautiful or polluted; for some a metaphor for an ending.	sunset, beauty, peaceful, beach, beautiful, peace, nature, calm, calming, something, contrast, warm, colour, end, tranquility	beach, beaches, breezes, gulls, shining
27R	**Self-portrait as a sick man**	Person who is worried, anxious, frightened, sad or ill; thoughtful contemplation; insomnia.	anxiety, sad, thinking, illness, fear, worry, man, thoughts, thought, lonely, anxious, sadness, contemplation, nightmares, night, scared, bad, looks	psychosomatic, neuroses, anxiety, feelings, mental
28R	**On the Beach, St. Malo**	Busy beach in summer, possibly in Britain, bringing enjoyment, fun, happiness, togetherness to families and communities taking a holiday from daily life.	beach, holiday, busy, fun, day, summer, family, happiness, scene, seaside, colour, bringing, hectic, Blackpool, pleasure, community, happy, enjoyment, people, together, life	beach, surf, teepees, beaches, family

29R	**Lone Wolf**	Lone wolf in cold snowy winter landscape, representing sad isolation or ex-communication; hints of protection; village.	wolf, lone, lonely, loneliness, winter, cold, nature, solitude, dark, dog, isolation, looking, lack, sad, wolves, lost, snow, hope	sled, Eskimo, huskies, wolf, howling
30R	**Never look back**	Woman rushing away, escaping, or fleeing from something potentially dangerous, possibly in fear or panic, possible association with crime, in modern urban setting.	running, lady, woman, away, back, scared, danger, someone, modern, panic, fear, day, life, looking, rushing, rushed	woman, prostitution, Picassos, feminists, legitimately
31R	**Melodious Barriers**	A journey through life, with choices and obstacles on the way to one's destination; one might be lost or lonely, or bridges may connect people. Human achievement in conquering nature's challenges.	life, way, bridges, lost, goals, obstacles, everyone, path, ladders, building, people, always, travelling, many, place, hills, new, take, put, go, sand, human, paths, great, journey, own, through, where, dunes, roads, different, desert	life, roads, people, way, Gobi
32R	**Declaration of Love**	War, possibly historic; Ignorance, destruction, bad times, conflict; Image of hope and some positive emotions despite destruction and carnage of war; Love and romance in war; possible propaganda.	war, love, military, male, hope, ignorance, danger, bad, looks, place, close, being, times, things, army, conflict, propaganda, destruction	war, wartime, defeat, ii, peace
33R	**Civitella**	Road or motorway, with roadworks or construction; workers; reminiscent of transport, journey, and normal everyday life; quiet.	road, life, working, journey, day, entry, off, being, quiet, stop, motorway, going, signs, construction	road, highway, Bonello, roads, roadbed

34R	**Autumn holiday in the village**	Festival, joyous celebration, happiness, dancing, feeling a sense of community, village in the countryside, colourful autumnal festival.	community, festival, happiness, party, happy, family, dance, village, celebration, joyful, autumn, farm, music, country, dancing, old, colours, joy, social, togetherness, history, unity, good, celebrations, together	dance, music, dances, dancing, musicians
35R	**Box of apples in Wales**	Apples, fruit box, food wasting; isolation, desolation, loneliness and abandonment.	fruit, food, apples, life, isolated, away, lonely, waste, lost, place, crate, nutrition, apple, deserted, desolate	food, life, bleak, fruit, starving
36R	**Fruit and Irises**	Woman-centered couple, love, relationship, imagining event, Paris, male giving to woman, reversal of gender roles, home, comfort.	life, woman, love, home, man, family, comfort, class, Paris, imagining, what, couple, book, romance, power, representation, trying, reading, lovers, marriage, relationships	marriage, couples, promiscuity, marriages, marital
37R	**Unexpected visitors**	Creative, eccentric artist in busy, chaotic, cluttered, complex setting; drugs or drink; crazy; puppets.	man, artist, chaotic, busy, crazy, very, life, old, puppets, weird, creativity, eccentric, painter, while, mind, puppet, creative, madness, strange, cartoon, chaos, art	artist, painting, Dubuffet, artists, painter
38R	**Comedy**	Contrast between dark and light, two entities; creepy and crazy; institution or hospital, fake positive emotion.	light, vs, dark, two, creepy, mental, room, family, very, what, times, evil, look, combined, images, happiness, forced, contrast, people, smile,	light, mental, illness, dark, psychoses

			between, colour, false, sanity, do, being	
39R	**Bow Bridge Boaters**	Peaceful, tranquil, relaxed day in summer / autumn with natural beauty and happiness; lake or river, sunflowers.	autumn, summer, happiness, peace, beauty, nature, warmth, tranquil, spring, colourful, peaceful, cat, natural, sunflowers, peacefulness, river, tranquility, happy, holiday, day, relaxing, afternoon, park, flowers	summer, spring, bloom, autumn, winter
40R	**The Zeppelin**	Goodbye, waving, set in wartime or end of war, Zeppelin, bleak, gloomy, dark, dull, with some hope; technology or invention.	goodbye, war, saying, hope, end, blimp, dull, dark, waving, image, light, old, new, century, greeting, bleak, men	war, ii, wartime, developments, world

Participant responses for Item 1A, with codes for the two coders. During the process, each coder generated his or her own set of codes, as a first step in summarizing the shared meaning. This involved some abstraction and generalization. Codes for each participant response can be seen under code AS and code PR. Summaries and notes can be seen in the table that follows below.

Participant response	**Code AS**	**Code PR**
A beach, a wave, trees.	watery landscape	water association
colour	colour, shape and art	appearance/colour
fear	mixed emotion and high energy	emotions
vibrance, mixture of cold and warm colours	mixed emotion and high energy	appearance/colour
busy, vibrant	mixed emotion and high energy	busy elements
Soft and harsh textures combined	mixed emotion and high energy	appearance/elements
Sydney	watery landscape	water association
Kaleidoscope of feelings	mixed emotion and high energy	emotions
hill tops	land-based landscape	shapes
colourful shapes / lines	colour, shape and art	appearance/colour/shapes
Global warming. Negative effect.	mixed emotion and high energy	associations abtract

hills. Twisting pathways.	land-based landscape	shapes
a lake	watery landscape	water association
framing abstract art to fit in society	colour, shape and art	association abstract
Chinese / Japanese, competitive vibe.	mixed emotion and high energy	oriental, competition
show piece in a gallery	colour, shape and art	association abstract
hidden natural beauty	land-based landscape	nature
overload of information	confusion	busy elements
landscape	land-based landscape	nature
colour	colour, shape and art	appearance/colour
mixed emotions	mixed emotion and high energy	emotions
Sydney Harbour Bridge	watery landscape	water association
different times of day	land-based landscape	change
chaos	confusion	busy elements
fading together / water related	watery landscape	water association
creative	colour, shape and art	association abstract
different levels of intensity	mixed emotion and high energy	appearance/elements
too much, confusing	confusion	busy elements
distortion	confusion	association abstract

mess. Even the frame is messy. Red shows waves of violence. Whirlpool	confusion	busy elements
anger	mixed emotion and high energy	emotions
a ship in a storm	watery landscape	water association
Looking for an escape.	mixed emotion and high energy	emotions
speed. Energy. Life.	mixed emotion and high energy	change
natural beauty. Seen + viewed differently.	land-based landscape	nature
nature through a window	land-based landscape	nature
fifty shades of colours	colour, shape and art	appearance/colour
nature	land-based landscape	nature
harsh waters - boat sinking	watery landscape	water association
colour expression	colour, shape and art	appearance/colour
Scenery entwined, blue sky, ocean, sunset, desert, grass, trees.	land-based landscape	nature, coloour
frantic and quite angry	mixed emotion and high energy	emotions
different moods	mixed emotion and high energy	emotions

Sky and trees	land-based landscape	nature
a scene where shapes mixed	colour, shape and art	shapes
movement	mixed emotion and high energy	change
green hills and blue sky	land-based landscape	nature

Independently, each coder wrote a summary that he or she felt captured the essence of the shared meaning. During an agreeing meeting the two coders agreed a joint summary for each artwork, keeping notes on key decisions. The independent summaries, joint summary and note for item 1A can be seen below.

Coder AS summary	Coder PR summary	Joint agreed summary	Note from agreeing process
Landscape featuring water or hills; Depiction of mixed or negative emotions; High energy and confusion; Colours, shapes and art.	natural landscapes, range of mixed emotions, associations with water, emphasis on colour, complexity and busyness, aspects of change	Natural landscape with water or hills, range of mixed emotions; Energy, confusion, complexity, colours.	We checked energy, and decided it needed to be included.