JOURNAL OF
Creative Behavior
Publication of the Creative Education Foundation™

SOMA CHAUDHURI
ALAN PICKERING
JOYDEEP BHATTACHARYA

Evaluating Poetry: Navigating the Divide between Aesthetical and Creativity Judgments

ABSTRACT

The comprehension and appreciation of poetry are inherently subjective, involving both creativity and aesthetic appeal. However, do these assessments of aesthetics and creativity rely on identical criteria, or do they vary depending on underlying factors? We addressed this question in this study. Participants ($N = 96$) evaluated 25 English poems across nine subjective characteristics: reading fluency, vivid imagery, perceived emotions (valence and arousal), surprise, originality, usefulness, aesthetic appeal, and creativity. Linear mixed-model analysis revealed that a poem's creativity was primarily predicted by its originality, followed by its usefulness and vivid imagery. Conversely, the evaluation of a poem's aesthetic appeal followed a different route; it was mainly predicted by a poem's reading fluency, followed by arousal, valence, and vivid imagery. Additionally, the association between creativity and originality was significantly moderated by participants' personality traits, specifically, openness, vividness of visual imagery, and curiosity. The relationship between aesthetic appeal and reading fluency was moderated by the vividness of auditory imagery trait. These findings suggest that a poem's creativity evaluation closely aligns with the standard definition of creativity, relying on its originality and usefulness. The study provides novel insights into the implicit internal models in the evaluation of poetry.

Keywords: poetry, evaluation, creativity, aesthetic appeal.

Poetry is one of the finest forms of verbal art, specifically excelling in the art of diction or word choice. It embodies emotive as well as interpretative art material (Whitcomb-Hess, 1944). The appreciation of any art is inherently subjective, involving a complex interaction between stimuli, individuals, and contextual factors (Leder, Gerger, Dressler, & Schabmann, 2012). In poetry, a poet's words pierce the reader, leaving a lasting impression (Robinson, 2002). Thus, poetry is the expression of intense personal experience perceived from a unique perspective (Furniss & Bath, 2013). Different readers might interpret the same poem differently, drawing from their own subjective experiences, knowledge, and perceptual skills. This idiosyncrasy in poetry perception results in distinctiveness in the assessment of both the creativity and aesthetic appeal of poems. When individuals contemplate a poem and evaluate its creativity and aesthetic appeal based on their subjective perceptions and personal definitions, a critical question arises: What factors contribute to the assessment of a poem's creativity and its aesthetic appeal, and do these assessments align or differ in terms of their underlying predictors? Also, how are these assessments influenced by the individual differences in readers' personality traits? We addressed these questions in the current study.

The standard definition of creativity asserts that the creativity of any product or idea requires both originality and usefulness (Runco & Jaeger, 2012). Given the highly subjective nature of creativity assessments of poems (Amabile, 1982), interpretations of nebulous concepts such as the originality and usefulness of a poem may vary among different readers. The impact of these two fundamental ingredients—originality and usefulness—on the evaluation of a poem remains largely uncharacterized. On the other hand, earlier studies have indicated that the perceived beauty and subsequent aesthetic appreciation of a poem predominantly rely on its structural elements, such as phonological constructs (Aryani, Kraxenberger, Ullrich, Jacobs, & Conrad, 2016), rhyme, meter, rhythm, prosodic fluency (Greene, Bodrumlu, & Knight, 2010; Lau, Cohn, Baldwin, Brooke, & Hammond, 2018; Obermeier et al., 2013; Reber, Schwarz, & Winkielman, 2004), metaphors (Rasse, Onysko, & Citron, 2020), as well as various subjective attributes such as ambiguity (Margulis,

Levine, Simchy-Gross, & Kroger, 2017), vivid imagery, perceived emotions (Belfi, Vessel, & Starr, 2018; Mehl, Gugliano, & Belfi, 2023), readers' expertise, psychological states and traits of readers (Hitsuwari & Nomura, 2022), and affective responses and feelings (Lüdtke, Meyer-Sickendieck, & Jacobs, 2014). Literature suggests that although the fields of creativity and aesthetics are often viewed separately, there is a notable correspondence between the aspects of art creation and a perceiver's aesthetic experience of that artwork (Tinio, 2013). This highlights the bridge between creative ideation and the aesthetical evaluation of art. However, the link between how individuals evaluate a poem's creativity and aesthetic appeal is not properly investigated. Therefore, in the present study, rather than focusing on creation or idea generation, we investigated the process of evaluation of perceived creativity and aesthetic appeal of poems.

We conducted a study with 96 participants who read and evaluated 25 contemporary English poems across nine dimensions: reading fluency, vividness in imagery, surprise, perceived emotions (valence and arousal), originality, usefulness, aesthetic appeal, and creativity. Using separate linear mixed models, the ratings of creativity and aesthetic appeal were predicted based on the ratings of the remaining seven dimensions. Additionally, we explored whether the relationships between these specific predictors and creativity or aesthetic appeal were moderated by readers' specific personality traits: openness, intellect, curiosity, vividness of visual imagery, and vividness of auditory imagery. The potential predictors and the moderators were selected after prior empirical research (Amabile, 1982; Belfi et al., 2018; Chamorro-Premuzic, Burke, Hsu, & Swami, 2010; Frame, Mehl, Head, & Belfi, 2023; Furnham & Chamorro-Premuzic, 2004; Furnham & Walker, 2001; Hitsuwari & Nomura, 2022; Johnson-Laird & Oatley, 2022; Kraxenberger & Menninghaus, 2017; Lloyd-Cox, Pickering, & Bhattacharya, 2022; Miall & Kuiken, 1994; Mussel, 2010; Reisenzein, 2013; Silvia & Christensen, 2020; Wassiliwizky, Koelsch, Wagner, Jacobsen, & Menninghaus, 2017).

MATERIALS AND METHODS

PARTICIPANTS

The experiment was designed using Qualtrics® software, and participants were recruited through Sona Systems®, receiving 2.5 course credits as compensation. Using the G*Power software (v. 3.1.9.4) (Faul, Erdfelder, Lang, & Buchner, 2007), we found that a minimum sample size of 103 was required for multiple linear regression, to detect a medium effect ($f^2 = 0.15$) at a significance level of 5% and a statistical power of 80%. Furthermore, by considering a multilevel model with 103 cluster groups, assuming a small-to-medium effect size (Cohen's d) of 0.3, and considering 25 observations per cluster group, the "samplesize_mixed" function in R (https://strengejacke.github.io/sjstats/) calculated that a total of 772 observations were required, equating to a minimum of 31 participants (772/25). We recruited 96 participants (12 males, 79 females, 4 non-binary, and 1 preferred not to disclose gender) with a mean (SD) age of 20.54 (4.97) years, resulting in a total of 2400 observations. This ensured sufficient statistical power for our study.

STIMULI

Twenty-five contemporary English poems from the early 20th Century to 2018 were selected in their entirety from esteemed online poetry repositories, such as Poetry.org (http://www.poetry.org/), Poetry Foundation (https://www.poetryfoundation.org/), and the Academy of American Poets (https://poets.org/). We intentionally selected short poems for their structured brevity, completeness, and diverse themes. All poems were original English compositions (see Table 1), each consisting of eight lines (mean word count = 49.4, $SD = 14.12$), exhibiting semantic and lexical diversity. Lexical diversity (LD) was assessed using the vocabulary-to-text ratio method (McCarthy & Jarvis, 2010), which calculates the ratio of the vocabulary size to the total number of words. Semantic diversity was measured using divergent semantic integration (DSI) scores, employing the semantic distance measuring approach (Johnson et al., 2023). The mean ($\pm$ SD) LD and DSI scores of the selected poems were 0.78 (0.03) and 0.77 (0.09), respectively. Additionally, we performed sentiment analysis using the "sentimentr" package (v 2.9.1) in R with the NRC sentiment lexicon (Rinker, 2021). The analysis revealed that 14 poems had an overall positive tone, 7 had a negative, and 4 exhibited a neutral tone. Table 1 provides detailed information about the stimuli. R scripts for sample size calculation, lexical diversity computation, and sentiment analysis are available on the OSF platform: https://osf.io/8m5y9/?view_only=be1b4a5aeda04d979b548c92c6462a23.

PROCEDURE

Each poem was displayed for 2 minutes for reading and contemplation. The titles of the poems were provided for potential anchoring, while the names of the poets were deliberately omitted to mitigate any

TABLE 1. Comprehensive Details of the Poems Utilized in the Experiment, Including Titles, Authors, Styles, Themes, Overall Tone, and Semantic and Lexical Diversity

Poem No.	Poem Name	Poet	Period	1st Line	No. of Lines	Word_ Count	Style/ Genre	Theme	Overall tone	Tone Coded (1 = P, 2 = N and 3 = Neu)	Semantic Diversity (DSI Scores)	Lexical Diversity (VoC-D)
1	Dust of Snow	Robert Frost	1874–1963 (1920)	The way a crow	8	34	Lyric	Nature	Positive	1	0.782467665	0.823529412
2	Biscuit	Jane Kenyon	1947–1995 (1993)	The dog has cleaned his bowl	8	47	Narrative	Trust and betrayal	Positive	1	0.794353712	0.787234043
3	Corner Seat	Louis MacNeice	1907–1963	Suspended in a moving night	8	51	Lyric	Loneliness and isolation	Negative	2	0.772157333	0.705882353
4	I Didn't Go To Church Today	Ogden Nash	1902–1971	I Didn't Go To Church Today	8	54	Lyric	Nature and spirituality	Positive	1	0.787875925	0.722222222
5	An Eastern Ballad	Allen Ginsberg	1926–1997	I speak of love that comes to mind	8	55	Lyric	Love	Positive	1	0.778244852	0.8
6	Delay	Elizabeth Jennings	1926–2001	The radiance of the star that leans on me	8	67	Lyric	Beauty and love	Positive	1	0.781102833	0.791044776
7	Be Frugal	Richard Church	1893–1972	Be frugal in the gift of love	8	55	Lyric	Love and relationship	Positive	1	0.785171663	0.818181818
8	In Innocence	J.V. Cunningham	1911–1985	In innocence I said	8	32	Lyric	Uncertainty	Neutral	3	0.787764194	0.875
9	Snow	Edward Thomas	1878–1917 (1913)	In the gloom of whiteness,	8	55	Lyric	Tragedy	Neutral	3	0.778845359	0.727272727
10	Hedges Freaked With Snow	Robert Graves	1895–1985	No argument, no anger, no remorse	8	51	Lyric	Acceptance and detachment	Negative	2	0.81801053	0.862745098
11	Mentor	Timothy Murphy	1951–2018	Had I known, only known	8	35	Lyric	Regret	Negative	2	0.758952997	0.771428571
12	The Traveler	Maya Angelou	1928–2014	Byways and bygone	8	31	Lyric	Human loneliness	Positive	1	0.818923027	0.677419355
13	I shall imagine life	E E Cummings	1894–1962	I shall imagine life	8	32	Lyric	Joy of life	Negative	2	0.803178012	0.971428571
14	Antimatter	Russell Edson	1935–2014	On the other side of a mirror there's an inverse world	8	69	Narrative	Duality of joy and sorrow	Positive	1	0.778287501	0.739130435
15	A Birthday Poem	James Simmons	1933–2001	For every year of life we light	8	48	Narrative	Inevitability of aging and death	Positive	1	0.784434966	0.791666667
16	Kyrie	Tomas Tranströmer	1931–2015	At times my life suddenly opens its eyes in the dark	8	72	Dramatic	Solitude	Negative	2	0.789966585	0.805555556
17	Love Comes Quietly	Robert Creeley	1926–2005	Love comes quietly	8	26	Lyric	Subtlety of love	Positive	1	0.743559988	0.923076923
18	Passing Remark	William Stafford	1914–1993	In scenery I like flat country	8	61	Narrative	Relationship	Positive	1	0.783156895	0.783333333

TABLE 1. (Continued)

Poem No.	Poem Name	Poet	Period	1st Line	No. of Lines	Word_ Count	Style/ Genre	Theme	Overall tone	Tone Coded (1 = P, 2 = N and 3 = Neu)	Semantic Diversity (DSI Scores)	Lexical Diversity (VoC-D)
19	All You Who Sleep Tonight	Vikram Seth	1952-	All you who sleep tonight	8	44	Lyric	Loneliness and longing for loved ones	Positive	1	0.764025784	0.813953488
20	Solitude	A.A. Milne	1882–1956 (1927)	I have a house where I go	8	49	Narrative	Solitude and solace	Neutral	3	0.671688145	0.5
21	I Remember You Because Of A Grassy Hill	Muna Lee	1895–1965	I remember you because of a grassy hill	8	71	Lyric	Recalling memories	Negative	2	0.787899686	0.712328767
22	Imagination	James Baldwin	1924–1987	Imagination	8	26	Lyric	Imagination and reality	Positive	1	0.777905425	0.807692308
23	The Night Will Never Stay	Eleanor Farjeon	1881–1965 (Early 20th)	The night will never stay	8	46	Lyric	Transient life	Negative	2	0.756266715	0.652173913
24	Running Water	Alfonsina Storni	1892–1938	Yes, I move, I live, I wander astray	8	60	Lyric	Exploration of life	Neutral	3	0.804632277	0.733333333
25	Song in Space	Adrian Mitchell	1932–2008	When man first flew beyond the sky	8	64	Lyric+Narrative	Humanity and the Earth	Positive	1	0.782554494	0.71875

Note. P = positive tone, N = negative tone, Neu = neutral tone; DSI = divergent semantic integration; VoC-D = vocabulary-to-text ratio method for measuring lexical diversity.

potential bias toward specific poets. On the next page, participants were asked to evaluate the poem across nine dimensions in the following order: reading fluency ("How easy is it to read this poem?"), aesthetic appeal ("How aesthetically appealing is this poem?"), perceived valence ("How positive (higher scores) or negative (lower scores) is the content of the poem?"), arousal ("How stimulating (higher scores) or relaxing (lower scores) is this poem?"), surprise ("How surprising is this poem?" By Surprise "we mean a contrast to expectation in the concept of the poem."), vividness in imagery ("How vivid is the imagery evoked from this poem?"), originality ("How original do you find this poem?"), usefulness ("How useful to you do you find this poem?"), and overall creativity ("How creative is this poem?"). Of note, participants were not provided with explicit definitions for originality, usefulness, creativity, and aesthetic appeal. Instead, they were instructed to rely on their intuitive and subjective understanding of these constructs. This approach is similar to consensual assessment technique (Amabile, 1982), recognized as the "gold standard" of creativity assessment (Kaufman, Plucker, & Baer, 2008), where the judges are not provided with predefined definitions of creativity and other constructs including aesthetic appeal, novelty, and appropriateness, but are instead instructed to apply their own subjective interpretations of the constructs. Prior research on the perception of poetry also adopts this technique of employing non-restrictive definitions for such constructs (Belfi et al., 2018). In this study, we aimed to identify which of the selected constructs predicted judgments of creativity and aesthetic appeal, despite differences in decontextualized dimensions like originality and usefulness of poems. Finally, participants completed questionnaires on demographic information and five personality traits, namely openness, intellect (DeYoung, Quilty, & Peterson, 2007), Vividness of Visual Imagery Questionnaire (VVIQ) (Marks, 1973), The Bucknell Auditory Imagery Scale—Vividness (BAIS-V, termed here as AVIQ, Auditory Vividness Imagery Questionnaire, for clarity) (Halpern, 2015), and Curiosity and Exploration Inventory-II (Kashdan et al., 2009).

DATA ANALYSIS

No noteworthy multicollinearity was observed among the independent variables: the variance inflation factor (VIF) < 3. The VIF assesses multicollinearity by indicating if a predictor exhibits a strong linear relationship with other predictors and is defined as $1/(1\text{-}R^2)$, where R^2 represents the coefficient of determination for the model predicting the variable from all the other predictor variables. VIF values greater than 10 indicate potential multicollinearity concerns (Bowerman & O'connell, 1990). Additionally, the measurement's reliability was affirmed by evaluating the internal consistency across items (Cronbach's alpha = 0.76; McDonald's omega total = 0.82; omega H asymptotic = 0.8; and omega hierarchical = 0.66) (Cronbach, 1951; McDonald, 2014). We used linear mixed-effects models with predictors (group mean centered) as fixed effects, participants as the grouping variable, and the intercepts for participants as random effects. The order of predictor inclusion was determined through a forward selection approach based on their strength of correlations, in descending order, with the respective outcome variables—creativity and aesthetic appeal. For predicting creativity, the hierarchical order of fixed effects was as follows: originality, vividness in imagery, usefulness, surprise, arousal, reading fluency, and valence. For predicting aesthetic appeal, the order was as follows: reading fluency, arousal, vividness in imagery, originality, valence, usefulness, and surprise. While predicting each outcome variable—creativity and aesthetic appeal—we compared seven linear mixed models using various criteria, including the Akaike information criterion (AIC) (Akaike, 1974), Schwarz Bayesian information criterion (BIC) (Schwarz, 1978), the proportion of variance explained by fixed effects (R^2), and the likelihood ratio test statistic ($\Delta?^2$). The best model fit results identified the potential predictors of creativity and aesthetic appeal.

Finally, we investigated how the five personality traits moderated the potential predictors of poetic creativity and aesthetic appeal. To achieve this, we established five distinct linear mixed models for each personality trait, treating both creativity and aesthetic appeal as outcome variables, to examine the interactions between the predictors and the corresponding personality trait.

RESULTS

Table 2 presents descriptive statistics for all nine variables related to poem ratings provided by the readers. The data exhibit slightly negative skewness and mild negative kurtosis, indicating a distribution that approaches normality. Table 3 presents descriptive statistics for the reader's five personality traits. Table 4 shows the bivariate correlations, where the means and standard deviations are over $N = 96$, but the ratings being averaged for the predictor variables (variables 1–9 in Table 4) are first each averaged over the 25 poems before being averaged over the participants. Creativity was significantly and positively correlated with

TABLE 2. Descriptive Statistics of the Poem-Related Variables Including Mean, Standard Deviation (*SD*), Skewness, Kurtosis, Standard Error (*SE*), and Variance Inflation Factor (VIF)

Variable	*N*	Mean	*SD*	Median	Min	Max	Skewness	Kurtosis	*SE*	VIF
Reading Fluency	2400	4.9	1.56	5	1	7	−0.42	−0.56	0.03	1.46
Aesthetic Appeal	2400	4.6	1.45	5	1	7	−0.18	−0.5	0.03	1.79
Valence	2400	3.96	1.45	4	1	7	0.08	−0.41	0.03	1.26
Arousal	2400	4	1.49	4	1	7	−0.1	−0.39	0.03	1.77
Surprise	2400	3.95	1.52	4	1	7	−0.06	−0.42	0.03	1.6
Vividness in Imagery	2400	4.72	1.57	5	1	7	−0.46	−0.42	0.03	1.57
Originality	2400	4.79	1.36	5	1	7	−0.36	−0.1	0.03	1.62
Usefulness	2400	3.92	1.36	4	1	7	0.04	−0.03	0.03	1.41
Creativity	2400	4.8	1.41	5	1	7	−0.37	−0.21	0.03	

Note. AVIQ = vividness of auditory imagery scores; *N* = No. of participants; *SD* = standard deviation; *SE* = standard error; VVIQ = vividness of visual imagery scores.

TABLE 3. Descriptive Statistics of Readers' Personality Trait Variables Including Mean, Standard Deviation (*SD*), Skewness, Kurtosis, and Standard Error (*SE*)

Variable	*N*	Mean	*SD*	Median	Min	Max	Skewness	Kurtosis	*SE*
Openness	96	4.78	0.74	4.6	3.1	6.4	0.25	−0.44	0.02
Intellect	96	4.4	0.81	4.5	2.1	6.1	−0.27	−0.36	0.02
Curiosity	96	4.55	1.06	4.4	2.4	7	0.25	−0.61	0.02
VVIQ	96	5.16	1	5.31	1	6.81	−1.23	2.88	0.02
AVIQ	96	4.49	1.16	4.46	1	6.64	−0.75	0.71	0.02

Note. AVIQ = vividness of auditory imagery scores; *N* = No. of participants; SD = standard deviation; *SE* = standard error; VVIQ = vividness of visual imagery scores.

all predictors ($p < .01$), except valence ($r = 0.19$, $p = .07$). Aesthetic appeal showed significant and positive correlations with all predictors ($p < .01$). Of note, creativity and aesthetic appeal were significantly correlated as well ($r = 0.62$, $p < .01$). Utilizing the qgraph package in R (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012), the network diagram (Figure 1) displays variables as nodes and partial correlations as edges highlighting robust connections among creativity, originality, and usefulness, alongside strong associations between aesthetic appeal and reading fluency.

Table 5 shows the linear mixed-model results for the best-fitting models predicting the creativity and aesthetic appeal of the poems. For predicting creativity, the most parsimonious model fit was achieved with a model comprising originality, appeal, usefulness, and vividness in imagery ($\Delta?^2 = 211.11$, BIC = 6462.1, $R^2 = 0.40$, $p < .001$). Specifically, originality was the best predictor ($b = 0.49$, $SE = 0.02$, $t = 27.01$, $p < .001$), followed by usefulness ($b = 0.24$, $SE = 0.02$, $t = 14.87$, $p < .001$) and vividness in imagery ($b = 0.15$, $SE = 0.01$, $t = 10.40$, $p < .001$). For predicting aesthetic appeal, we found that the model comprising reading fluency, arousal, valence, vividness imagery, and originality was the most parsimonious fit ($\Delta?^2 = 79.3$, BIC = 7185.7, $R^2 = 0.34$, $p < .001$). Specifically, reading fluency was the best predictor ($b = 0.32$, $SE = 0.02$, $t = 19.08$, $p < .001$), followed by arousal ($b = 0.20$, $SE = 0.02$, $t = 9.90$, $p < .001$), valence ($b = 0.16$, $SE = 0.01$, $t = 8.98$, $p < .001$), vividness in imagery ($b = 0.13$, $SE = 0.02$, $t = 7.16$, $p < .001$), and originality ($b = 0.09$, $SE = 0.02$, $t = 4.24$, $p < .001$). Figures 2a and 2b display the network diagrams for the best-fit models predicting creativity and aesthetic appeal, respectively. For model comparison results, see Table S1 in the Supplementary Materials.

Finally, we explored the moderating influence of the five personality traits—openness, intellect, epistemic curiosity, vividness of visual imagery, and vividness of auditory imagery—separately on the predictors of creativity and aesthetic appeal. The significant moderating effects are shown in Table 6 and are visually

TABLE 4. Bivariate Correlation Coefficients for Poetry-Related Variables and the Personality Measures of the Readers

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Reading Fluency	4.9	0.72													
2. Aesthetic Appeal	4.6	0.62	0.58**												
3. Valence	3.96	0.62	0.14	0.30**											
4. Arousal	4	0.79	0.16	0.46**	0.37**										
5. Surprise	3.95	0.68	0.11	0.40**	0.24*	0.62**									
6. Vividness Imagery	4.72	0.59	0.49**	0.53**	0.22*	0.26*	0.39**								
7. Originality	4.79	0.65	0.44**	0.59**	0.15	0.27**	0.41**	0.65**							
8. Usefulness	3.92	0.56	−0.01	0.23*	0.08	0.36**	0.64**	0.25*	0.36**						
9. Creativity	4.8	0.68	0.43**	0.62**	0.19	0.30**	0.48**	0.70**	0.90**	0.43**					
10. Openness	4.78	0.75	0.17	0.38**	−0.03	0.21*	0.19	0.37**	0.27**	0.29**	0.30**				
11. Intellect	4.4	0.82	0.21*	0.08	0.09	0.11	0.03	0.22*	0.21*	0.03	0.17	0.28**			
12. Curiosity	4.55	1.06	0.09	0.12	0.08	0.21*	0.13	0.12	0.18	0.15	0.22*	0.05	0.39**		
13. VVIQ	5.16	1	0.18	−0.04	−0.02	−0.15	−0.14	0.14	0.11	−0.02	0.09	0.06	0.27**	0.29**	
14. AVIQ	4.49	1.17	0.14	−0.1	0	−0.14	−0.1	0.04	−0.01	0.01	−0.01	−0.01	0.18	0.1	0.54**

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. The means and *SD* are over $N = 96$ but the ratings being averaged for variables 1–9 are first each averaged over the 25 poems before being averaged over the participants; AVIQ = Vividness of auditory imagery trait scores; VVIQ = Vividness of visual imagery trait scores. *Indicates $p < .05$. **Indicates $p < .01$.

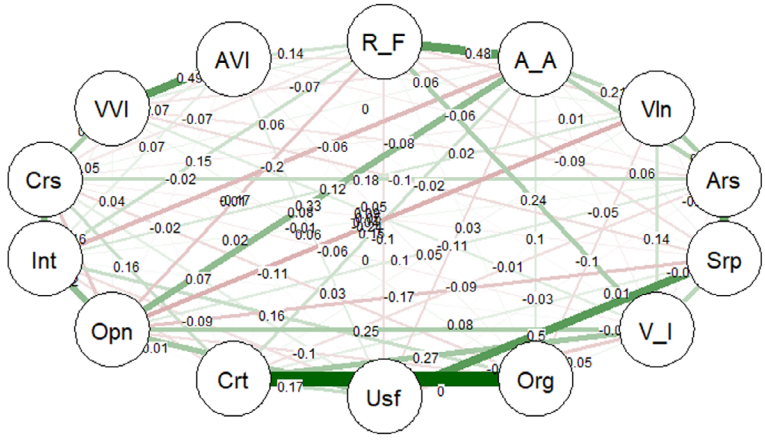


FIGURE 1. Network diagram representing the partial correlations among the variables. The nodes represent the different variables studied, with edges indicating the partial correlations between pairs of variables. The values on the edges represent the magnitude of these partial correlations. Positive correlations are shown in green, while negative correlations are shown in red. The thickness of the edges is proportional to the strength of the correlations. *Note*: R_F = Reading fluency; A_A = Aesthetic appeal; Vln = Valence; Ars = Arousal; Srp = Surprise; V_I=Vivid imagery; Org = Originality; Usf = Usefulness; Crt = Creativity; Opn = Openness; Int = Intellect; Crs = Curiosity; VVI=Vividness of visual imagery (VVIQ); AVI = Vividness of auditory imagery (AVIQ).

depicted in Figure 3a–j. Originality, the strongest predictor of creativity, was significantly influenced by openness ($b = -0.11$, $SE = 0.03$, $t = -4.37$, $p < .001$), VVIQ ($b = 0.06$, $SE = 0.02$, $t = 2.98$, $p < .001$), curiosity ($b = -0.04$, $SE = 0.02$, $t = -2.39$, $p = .02$), and marginally by intellect ($b = -0.05$, $SE = 0.02$, $t = -2.0$, $p = .05$). Usefulness was marginally influenced by VVIQ ($b = -0.03$, $SE = 0.02$, $t = -1.95$, $p = .05$). Simple slopes analysis (Table 7) indicated that the readers scoring lower in openness, intellect, and curiosity had a stronger influence on originality while predicting creativity relative to their higher-scoring counterparts (See Figures 3a to 3c). Readers with higher visual imagery ability (VVIQ scores) showed a stronger influence on originality (See Figure 3d) but a weaker influence on usefulness in predicting creativity (See Figure 3e). Notably, auditory imagery ability (AVIQ scores) did not significantly interact with any of the predictors of creativity, indicating that the reader's vivid auditory imagery ability did not influence the prediction of a poem's creativity.

Repeating the analysis for aesthetic appeal as the response variable, we found that reading fluency, the best predictor of aesthetic appeal, was significantly influenced only by the AVIQ ($b = -0.03$, $SE = 0.01$, $t = -2.28$, $p = .02$). Arousal was influenced by VVIQ ($b = -0.05$, $SE = 0.02$, $t = -2.25$, $p = .02$), vividness in imagery by openness ($b = 0.05$, $SE = 0.02$, $t = 2.11$, $p = .03$), originality by AVIQ ($b = 0.05$, $SE = 0.02$, $t = 2.84$, $p < .001$), and valence by intellect ($b = 0.07$, $SE = 0.02$, $t = 3.10$, $p = <.001$). Simple slopes analysis (Table 7) revealed that the relationship between vivid imagery and aesthetic appeal was more prominent in readers with higher openness (Figure 3f), and the valence–aesthetic appeal relationship was more pronounced in individuals with higher intellect (Figure 3g). The arousal–aesthetic appeal connection was stronger in readers with lower VVIQ scores (Figure 3h). Interestingly, participants with higher AVIQ scores exhibited a stronger connection between originality and aesthetic appeal (Figure 3i), while the association between reading fluency and aesthetic appeal was more pronounced in individuals with lower AVIQ scores (Figure 3j). For the detailed moderation results, see Table S2 in the Supplementary Materials.

DISCUSSION

The study unveils two crucial facets of poetry evaluation. First, it differentiates between the evaluation of a poem's creativity and its aesthetic appeal, demonstrating that these are distinct processes with minimal

TABLE 5. The Linear Mixed-model Results for the Best Model Fit in Predicting Creativity and Aesthetic Appeal: Creativity Prediction Model Shows Originality as the Best Predictor Followed by Usefulness, Vividness in Imagery, Arousal, and Surprise; Aesthetic Appeal Prediction Model Shows Reading Fluency as the Best Predictor Followed by Arousal, Valence, Vividness in Imagery, and Originality

The linear mixed-model results for the best model fit for predicting creativity

MODEL INFO

Observations: 2400

Dependent Variable: Creativity

Type: Mixed effects linear regression

MODEL FIT

AIC = 6427.4, BIC = 6462.1

Pseudo-R^2 (fixed effects) = 0.40

Pseudo-R^2 (total) = 0.62

Random effects

Groups	Name	Variance	*SD*
Participants	(Intercept)	0.43	0.65
Residual		0.76	0.87

Number of observations: 2400, grouping variable: participants, number of groups: 96, and ICC: 0.36

Fixed effects

	Estimate	*SE*	*df*	*t*-Value	*p*-Value
(Intercept)	4.80	0.07	95.99	69.54	<.001
Originality	0.49	0.02	2303.99	27.01	<.001
Vividness Imagery	0.15	0.01	2303.99	10.40	<.001
Usefulness	0.24	0.02	2303.99	14.87	<.001

The linear mixed-model results for the best model fit for predicting aesthetic appeal

MODEL INFO

Observations: 2400

Dependent variable: Aesthetic appeal

Type: Mixed effects linear regression

MODEL FIT

AIC = 7139.4, BIC = 7185.7

Pseudo-R^2 (fixed effects) = 0.34

Pseudo-R^2 (total) = 0.50

Random effects

Groups	Name	Variance	*SD*
Participants	*(Intercept)*	*0.34*	0.59
Residual		1.04	1.02

Number of observations: 2400, grouping variable: participants, number of groups: 96, and ICC: 0.25

Fixed effects

	Estimate	*SE*	*df*	*t*-Value	*p*-Value
(Intercept)	4.60	0.06	95.99	72.53	<.001

TABLE 5. (Continued)

Fixed effects	Estimate	*SE*	*df*	*t*-Value	*p*-Value
Reading Fluency	0.32	0.02	2303.99	19.08	<.001
Arousal	0.20	0.02	2303.99	9.90	<.001
Vividness–Imagery	0.13	0.02	2303.99	7.16	<.001
Originality	0.09	0.02	2303.99	4.24	<.001
Valence	0.16	0.02	2303.99	8.98	<.001

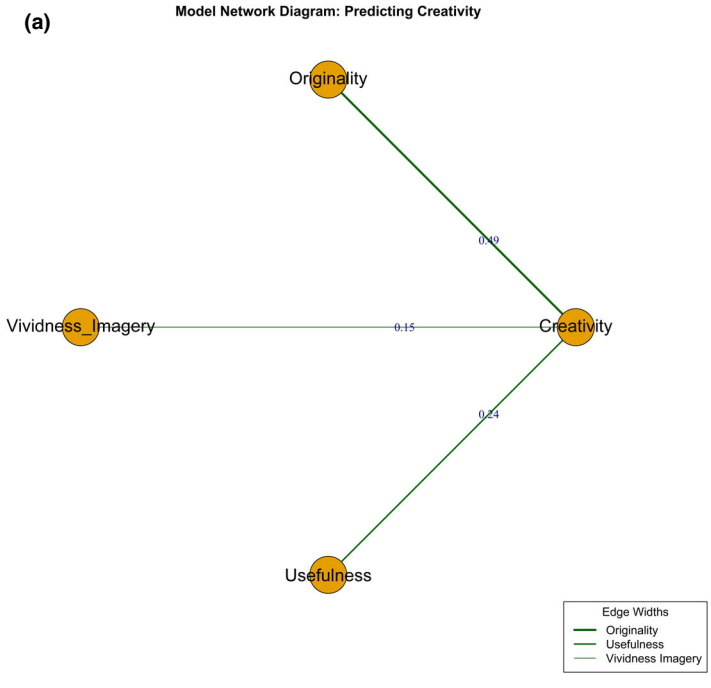
Note: AIC = Akaike information criterion (Akaike, 1974); BIC = Bayesian information criterion (Schwarz, 1978); *df* = Degrees of freedom; ICC = Intra-class correlation coefficient; *SE* = Standard error.

overlap. The assessment of a poem's creativity is based on its originality, usefulness, and vivid imagery. In contrast, the evaluation of its aesthetic appeal relies on its reading fluency, perceived arousal, perceived valence, and vivid imagery. Second, the study aligns the evaluation of poetic creativity with the standard definition of creativity (Runco & Jaeger, 2012), which emphasizes both originality and usefulness.

Consistent with traditional criteria for evaluating creative products, our findings underscore originality as the foremost determinant of a poem's creativity. Nevertheless, originality alone is insufficient; usefulness emerges as the second-most important predictor of creativity, reinforcing the notion that creative poems must be both original and useful (Runco, 1988). This indicates that, despite idiosyncratic subjective interpretations, the judgment of a poem's creativity fundamentally depends on these two necessary components: originality and usefulness. Additionally, the study shows that vivid imagery in poems significantly enhances perceived creativity; poems with rich and evocative imagery were judged as more creative.

Conversely, the assessment of a poem's aesthetic appeal follows a different route, with reading fluency emerging as the primary predictor. This corresponds with the notion that faster reading speed, indicative of enhanced processing fluency (Lea, Rapp, Elfenbein, Mitchel, & Romine, 2008), is pivotal for the aesthetic appreciation of a poem. Furthermore, this finding is consistent with prior research suggesting that aesthetic experience is positively influenced by the processing dynamics of the perceiver (Reber et al., 2004). Therefore, we propose that the more fluently the reader can comprehend the poem, the higher their aesthetic evaluation of it. Following reading fluency, arousal emerges as the next strongest predictor, followed by perceived valence and vivid imagery. This finding somewhat contrasts with earlier research that highlighted the predictive role of vivid imagery over emotional valence in specific poetry forms, such as haiku and sonnets (Belfi et al., 2018). Haiku is a genre of poetry commonly associated with seasons, often emphasizing nature imagery as its most important feature. Similarly, Petrarchan sonnets prominently feature the "volta" or "turn," which often leads to visual imagery (Whissell, 2018). In contrast to these structurally constrained genres of poetry, the poems selected for this study are from diverse styles and themes. We observed perceived arousal and valence to be more influential than vivid imagery in predicting a poem's aesthetic appeal. Hence, despite individual variations in responses, poems that evoke positive emotions and are stimulating are generally perceived as more aesthetically appealing, highlighting the strong connection between emotional valence and aesthetic appeal (Leder et al., 2012). In this context, the perceived emotions in our study can be referred to as "aesthetic emotions," associated with a special type of perceived aesthetic appeal that predicts the subjectively felt pleasure or displeasure and the liking or disliking connected with this type of appeal (Menninghaus et al., 2019). This finding aligns with earlier research suggesting that, unlike the negativity bias in classical emotions, the emotion terms used for the appraisal of intrinsic pleasantness predominantly include more positive than negative emotions (Menninghaus et al., 2019; Scherer, 2005). Moreover, the higher predictive power of arousal suggests that aesthetic emotions are typically pursued and enjoyed intrinsically, where the subjectively perceived intensity and/or emotional arousal serve as rewards in themselves (Menninghaus et al., 2019).

This study indicates that vivid imagery in poems predicts their aesthetic appeal. Poetic imagery being "the sensory and figurative language used in poetry" (https://www.britannica.com/art/poetic-imagery) is a universally central dimension in poetic meaning production (Brandt & Brandt, 2005). Our results corroborate earlier research suggesting that figurative languages evoke aesthetic experiences at the phonological and prosodic levels eliciting pleasurable feelings associated with the perception of beauty (Citron & Zervos, 2018;



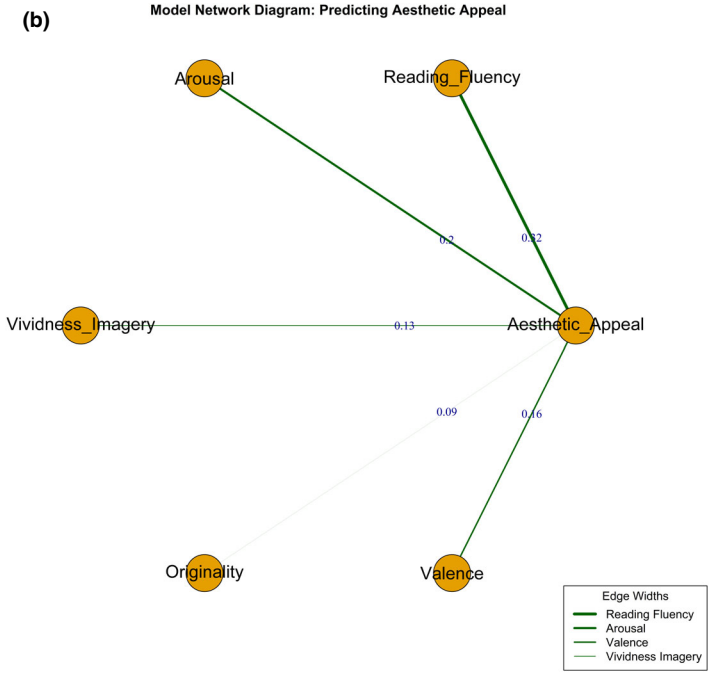


FIGURE 2. (a) Network diagram illustrating the model for predicting the creativity of a poem. Specifically, the diagram shows how originality, usefulness, and vivid imagery contribute to predicting creativity. The numeric values on the edges indicate the estimated coefficients of the predictor variables in the linear mixed model predicting creativity. (b) Network diagram illustrating the model for predicting the aesthetic appeal of a poem. Specifically, the diagram shows how reading fluency, arousal, valence, and vivid imagery contribute to predicting aesthetic appeal. The numeric values on the edges indicate the estimated coefficients of the predictor variables in the linear mixed model predicting aesthetic appeal.

TABLE 6. Results of the Significant Moderation Effects of the Personality Traits (Openness, Intellect, Curiosity, Vividness of Visual Imagery (VVIQ), and Vividness of Auditory Imagery (AVIQ)) in Predicting Creativity and Aesthetic Appeal

Interaction (Moderator*Predictor)	Estimate	*SE*	*t*	*p*	Fit (R^2)
Predicting creativity					
Openness*Originality	−0.11	0.03	−4.37	<.001	0.43
Intellect*Originality	−0.05	0.02	−2.49	.05	0.41
Curiosity*Originality	−0.04	0.02	−2.39	.02	0.41
VVIQ*Originality	0.06	0.02	2.98	<.001	0.41
VVIQ*Usefulness	−0.03	0.02	−1.95	.05	0.41
Predicting aesthetic appeal					
Openness*Vividness in Imagery	0.05	0.02	2.11	.03	0.37
Intellect*Valence	0.07	0.02	3.1	<.001	0.34
VVIQ*Arousal	−0.05	0.02	−2.25	.02	0.34
AVIQ*Originality	0.05	0.02	2.84	<.001	0.34
AVIQ*Reading Fluency	−0.03	0.01	−2.28	.02	0.34

Menninghaus et al., 2015; Van Peer, 1990). Additionally, our findings are consistent with prior studies that vivid imagery tends to enhance the aesthetic appeal of poems, like haiku and sonnets (Belfi et al., 2018). Notably, we observed that vivid imagery significantly predicts the assessment of poems' creativity. This further aligns with the notion that the creative interplay of language and thought is particularly evident in figurative language which helps construct a high-order linkage between the entities referred to (Cacciari, 1998; Katz, Cacciari, Gibbs, & Turner, 1998), manipulating implicit meaning in poems (Miall & Kuiken, 1994). Enhanced vivid imagery evoked by the figures of speech likely enhances readers' engagement, comprehension, and interpretation, fostering deeper critical thought and appreciation of a poem's creativity alongside its aesthetics. Hence, our results imply that poems evoking vivid imagery not only enhance aesthetic appreciation but also stimulate creative contemplation during poetry evaluation. Of note, we found that perceived emotions played a more effective influence on the aesthetic appeal compared to vivid imagery.

We also observed a significant distinction between the two models for evaluating creativity and aesthetic appeal, particularly concerning individual differences in readers' personality traits. Our findings revealed that only the visual imagery trait, as indicated by VVIQ scores, positively influences the link between originality and creativity. Specifically, readers with a greater capacity to visualize vivid mental images (higher VVIQ scores) tend to prioritize a poem's originality more during the assessment of its creativity compared to those with lower VVIQ scores. Conversely, readers with higher levels of openness, intellect, and curiosity tend to assign less importance to a poem's originality in their creativity evaluation. This finding somewhat challenges our conventional understanding of these traits, suggesting that individuals who are less open and less intellectually curious may be more judgmental regarding originality and its effect on creativity. Additionally, even a slight increase in the assessment of originality appears to have a more significant impact on the perception of poetry creativity for individuals who are less open and less intellectually curious. On the contrary, individuals with higher scores in these traits seem to be less biased toward originality when assessing poetry

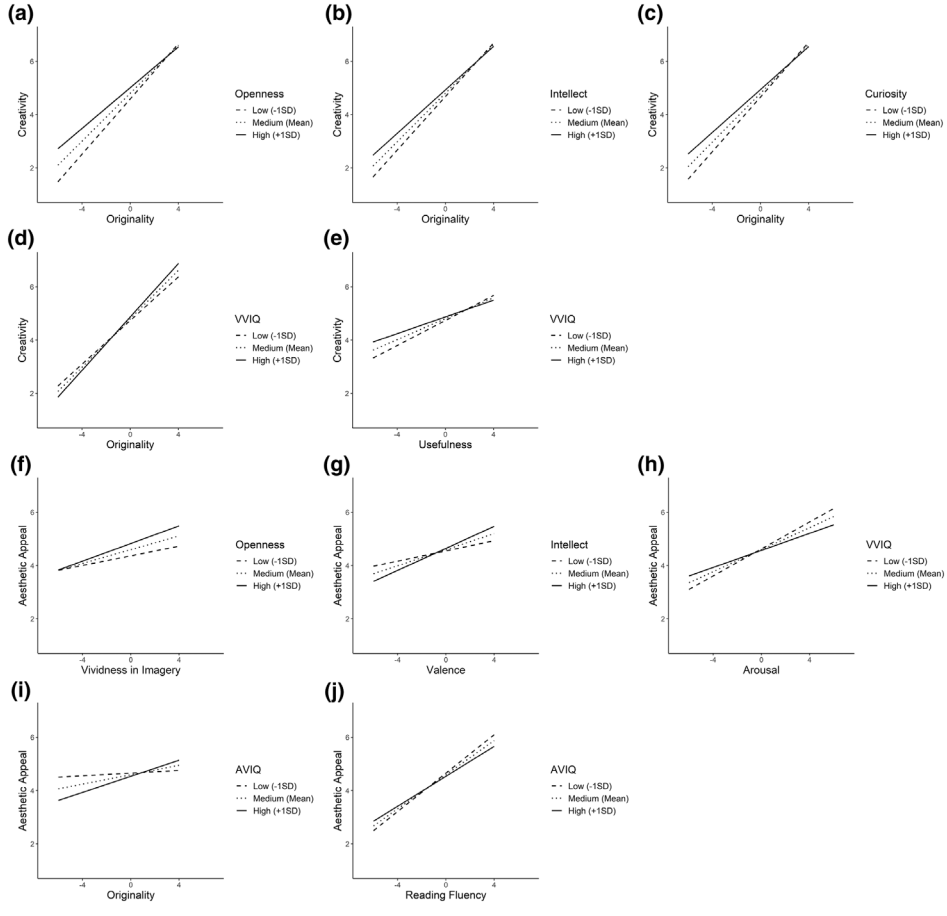


FIGURE 3. Simple slopes illustrating the moderation effects of five personality traits (openness, intellect, curiosity, vividness of visual imagery (VVIQ), and vividness of auditory imagery (AVIQ)) on the prediction of creativity and aesthetic appeal of poems. Each subplot, (a) to (j), represents how these personality traits moderate the relationship between a predictor (originality, usefulness, vividness in imagery, valence, arousal, or reading fluency) and the outcomes (creativity or aesthetic appeal). The slopes are displayed for three levels of the moderator: low (−1 *SD*), medium (mean), and high (+1 *SD*).

creativity. Further research is needed to explore this aspect in greater detail. While predicting the aesthetic appeal of a poem, we found that openness, intellect, and auditory imagery (AVIQ) positively influence the predictive roles of vividness in imagery, valence, and originality, respectively. Conversely, both auditory imagery and visual imagery abilities in readers exert negative influences on the predictive roles of reading fluency and arousal, respectively. Readers with heightened auditory imagery are less influenced by reading fluency, and those with higher visual imagery abilities are less influenced by arousal in their aesthetic judgments. An interesting finding is that reading fluency, the primary predictor of aesthetic appeal, is not positively influenced by any of the chosen personality traits among the readers. Our assessment of the ease of reading focused on perceptual fluency rather than conceptual fluency or the poem's meaningfulness. Since the poems we selected were brief, they were inherently easy to read. Prior research suggests that conceptual fluency is a better predictor of aesthetic appeal (Martindale, Moore, & Borkum, 1990). If participants were asked about how easily they conceptualized the poems, personality traits might have shown positive

TABLE 7. Results of the Simple Slopes Analyses for Significant Moderation Effects on Predictors of Creativity and Aesthetic Appeal at High and Low Levels of Five Personality Traits, such as Openness, Intellect, Curiosity, Vividness of Visual Imagery (VVIQ), and Vividness of Auditory Imagery (AVIQ)

Predicting creativity		**High (+1 *SD*)**				**Low (−1 *SD*)**				**Slope Difference (High-Low)**			
Moderator	**Predictor**	**Estimate**	***SE***	***t*-Value**	***p*-Value**	**Estimate**	***SE***	***t*-Value**	***p*-Value**	**Estimate**	***SE***	***t*-Value**	***p*-Value**
Openness	Originality	0.4	0.03	15.3	<.001	0.57	0.03	22.26	<.001	−0.17	0.04	−4.36	<.0001
Intellect	Originality	0.46	0.03	18.5	<.001	0.53	0.03	19.6	<.001	−0.07	0.04	−2	.05
Curiosity	Originality	0.45	0.02	18.91	<.001	0.54	0.03	20.02	<.001	−0.08	0.04	−2.39	.02
VVIQ	Originality	0.55	0.03	20.87	<.001	0.44	0.03	16.7	<.001	0.11	0.04	−2.98	.0029
VVIQ	Usefulness	0.21	0.02	9.03	<.001	0.27	0.02	11.8	<.001	−0.06	0.03	−1.95	.05
Predicting aesthetic appeal													
Openness	Vividness in Imagery	0.16	0.03	6.49	<.001	0.09	0.02	3.75	<.001	0.08	0.04	2.1	.04
Intellect	Valence	0.21	0.02	8.73	<.001	0.1	0.03	3.58	<.001	0.11	0.04	3.1	.002
VVIQ	Arousal	0.16	0.03	5.96	<.001	0.25	0.03	8.15	<.001	−0.09	0.04	−2.24	.02
AVIQ	Originality	0.15	0.03	4.99	<.001	0.02	0.03	0.84	0.4	0.12	0.04	−2.84	.005
AVIQ	Reading Fluency	0.28	0.02	11.63	<.001	0.36	0.02	14.93	<.001	−0.08	0.03	−2.28	.02

influences on conceptual fluency while judging a poem's aesthetical appeal. This would support the notion that "beauty is in the processing experiences of the beholder" (Reber et al., 2004, p. 378).

LIMITATIONS

It is important to acknowledge certain limitations of the present study. Firstly, the diversity of the poems might complicate the comparison of disparate aspects of creativity. Despite the thematic, periodic, and stylistic diversity of the selected poems, they exhibit structural uniformity, each being 8 lines long and containing, on average, 50 words. Furthermore, both the lexical and semantic diversity analysis suggested small standard deviations across the poems, implying a narrower range of vocabulary and semantic variation. This uniformity facilitated a more focused comparison of creativity and aesthetic appeal by providing some homogeneity among the selected poems. However, this also might limit the ability to draw generalized inferences, which would require a much larger set of poems with diverse content. Thus, we acknowledge a nuanced trade-off between subjective diversity and objective uniformity in assessing poetic creativity in our study. Secondly, our participants were not given explicit definitions for constructs such as originality, usefulness, and creativity. Instead, in line with earlier research (Amabile, 1982; Belfi et al., 2018), they were instructed to rely on their own subjective understanding of these constructs. Providing explicit context and definitions for each dimension might enhance the interpretation of findings. Alternatively, employing a semi-structured grounded theory approach could offer a more nuanced understanding of these constituent nebulous predictors (e.g., by exploring how and in what context a poem becomes useful).

CONCLUSION

In conclusion, our study addresses an important question in poetry evaluation: Are the judgments of a poem's creativity and aesthetic appeal aligned? Our findings indicate distinct evaluation paths: creativity is assessed based on originality, usefulness, and vivid imagery, while aesthetic appeal depends on reading fluency, perceived emotions, and vivid imagery. This distinction underscores that internal models of creativity adhere to the standard bipartite definition of creativity, emphasizing both originality and usefulness. Personality traits, notably openness and curiosity, along with vividness in visual imagery, influence creativity judgments. In contrast, aesthetic appeal judgments are shaped by personality traits, namely openness, intellect, curiosity, and vividness of both auditory and visual imagery. These results altogether offer valuable insights into the complex and varied intrapersonal models involved in the multifaceted nature of art evaluation in the context of poetry.

FUNDING INFORMATION

This research received no external funding.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ETHICS STATEMENT

The study protocol was approved by the local Ethics Committee of the Department of Psychology, Goldsmiths, University of London.

DATA AVAILABILITY STATEMENT

All data and code are available in the OSF at: https://osf.io/8m5y9/?view_only=be1b4a5aeda04d979b548c92c6462a23.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723.

Amabile, T.M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, *43*, 997–1013; doi: 10.1037/0022-3514.43.5.997.

Aryani, A., Kraxenberger, M., Ullrich, S., Jacobs, A.M., & Conrad, M. (2016). Measuring the basic affective tone of poems via phonological saliency and iconicity. *Psychology of Aesthetics, Creativity, and the Arts*, *10*, 191–204; doi: 10.1037/aca0000033.

Belfi, A.M., Vessel, E.A., & Starr, G.G. (2018). Individual ratings of vividness predict aesthetic appeal in poetry. *Psychology of Aesthetics, Creativity, and the Arts*, *12*, 341–350; doi: 10.1037/aca0000153.

Bowerman, B.L., & O'connell, R.T. (1990). *Linear statistical models: An applied approach* (2nd edn). Pacific Groove, California: Duxbury Press. https://lccn.loc.gov/89016367.

Brandt, L., & Brandt, P.A. (2005). Cognitive poetics and imagery. *European Journal of English Studies*, *9*, 117–130; doi: 10.1080/13825570500171861.

Cacciari, C. (1998). *Why do we speak metaphorically? Reflections on the functions of metaphor in discourse and reasoning. Figurative language and thought* (pp. 119–157). Oxford University Press.

Chamorro-Premuzic, T., Burke, C., Hsu, A., & Swami, V. (2010). Personality predictors of artistic preferences as a function of the emotional valence and perceived complexity of paintings. *Psychology of Aesthetics, Creativity, and the Arts*, *4*, 196.

Citron, F., & Zervos, E.A. (2018). A neuroimaging investigation into figurative language and aesthetic perception. In A. Baicchi, R. Digonnet, & J. Sandford (Eds.), *Sensory perceptions in language, embodiment and epistemology. Studies in Applied Philosophy, Epistemology and Rational Ethics* (vol. *42*, pp. 77–94). Cham: Springer; doi: 10.1007/978-3-319-91277-6_5.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334; doi: 10.1007/BF02310555.

DeYoung, C.G., Quilty, L.C., & Peterson, J.B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*, 880–896; doi: 10.1037/0022-3514.93.5.880.

Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D., & Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, *48*, 1–18; doi: 10.18637/jss.v048.i04.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191; doi: 10.3758/BF03193146.

Frame, J., Mehl, K., Head, K., & Belfi, A. (2023). The influence of sensory modality on aesthetic judgments of poetry. *Psychology of Aesthetics, Creativity, and the Arts*, Advance online publication. doi: 10.31234/osf.io/cv6px.

Furnham, A., & Chamorro-Premuzic, T. (2004). Personality, intelligence, and art. *Personality and Individual Differences*, *36*, 705–715; https://doi.org/10.1016/S0191-8869(03)00128-4.

Furnham, A., & Walker, J. (2001). The influence of personality traits, previous experience of art, and demographic variables on artistic preference. *Personality and Individual Differences*, *31*, 997–1017; doi: 10.1016/S0191-8869(00)00202-6.

Furniss, T., & Bath, M. (2013). *Reading poetry: an introduction*. London, UK: Routledge. https://doi.org/10.4324/9781315835303.

Greene, E., Bodrumlu, T., & Knight, K. (2010). Automatic analysis of rhythmic poetry with applications to generation and translation. Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, 524–533.

Halpern, A.R. (2015). Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain*, *25*, 37–47; doi: 10.1037/pmu0000081.

Hitsuwari, J., & Nomura, M. (2022). How individual states and traits predict aesthetic appreciation of haiku poetry. *Empirical Studies of the Arts*, *40*, 81–99; doi: 10.1177/0276237420986420.

Johnson, D.R., Kaufman, J.C., Baker, B.S., Patterson, J.D., Barbot, B., Green, A.E., ... & Beaty, R.E. (2023). Divergent semantic integration (DSI): Extracting creativity from narratives with distributional semantic modeling. *Behavior Research and Methods*, *55*, 3726–3759. doi: 10.3758/s13428-022-01986-2.

Johnson-Laird, P.N., & Oatley, K. (2022). How poetry evokes emotions. *Acta Psychologica*, *224*, 103506; doi: 10.1016/j.actpsy.2022.103506.

Kashdan, T.B., Gallagher, M.W., Silvia, P.J., Winterstein, B.P., Breen, W.E., Terhar, D., & Steger, M.F. (2009). The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality*, *43*, 987–998; doi: 10.1016/j.jrp.2009.04.011.

Katz, A.N., Cacciari, C., Gibbs, R.W., & Turner, M. (1998). *Figurative language and thought*. New York, USA: Oxford University Press. https://books.google.co.uk/books?id=lSxhb_fZLx0C.

Kaufman, J.C., Plucker, J.A., & Baer, J. (2008). *Essentials of creativity assessment*. New York, USA: John Wiley & Sons.

Kraxenberger, M., & Menninghaus, W. (2017). Affinity for poetry and aesthetic appreciation of joyful and sad poems. *Frontiers in Psychology*, *7*, 2051. doi: 10.3389/fpsyg.2016.02051.

Lau, J.H., Cohn, T., Baldwin, T., Brooke, J., & Hammond, A. (2018). Deep-speare: A joint neural model of poetic language, meter and rhyme. doi: 10.48550/arxiv.1807.03491.

Lea, R.B., Rapp, D.N., Elfenbein, A., Mitchel, A.D., & Romine, R.S. (2008). Sweet silent thought: Alliteration and resonance in poetry comprehension. *Psychological Science*, *19*, 709–716; doi: 10.1111/j.1467-9280.2008.02146.x.

Leder, H., Gerger, G., Dressler, S.G., & Schabmann, A. (2012). How art is appreciated. *Psychology of Aesthetics, Creativity, and the Arts*, *6*, 2–10; doi: 10.1037/a0026396.

Lloyd-Cox, J., Pickering, A., & Bhattacharya, J. (2022). Evaluating creativity: How idea context and rater personality affect considerations of novelty and usefulness. *Creativity Research Journal*, *34*, 373–390; doi: 10.1080/10400419.2022.2125721.

Lüdtke, J., Meyer-Sickendieck, B., & Jacobs, A.M. (2014). Immersing in the stillness of an early morning: Testing the mood empathy hypothesis of poetry reception. *Psychology of Aesthetics, Creativity, and the Arts*, *8*, 363.

Margulis, E.H., Levine, W.H., Simchy-Gross, R., & Kroger, C. (2017). Expressive intent, ambiguity, and aesthetic experiences of music and poetry. *PLoS One*, *12*, e0179145; doi: 10.1371/journal.pone.0179145.

Marks, D.F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*, 17–24; doi: 10.1111/j.2044-8295.1973.tb01322.x.

Martindale, C., Moore, K., & Borkum, J. (1990). Aesthetic preference: Anomalous findings for Berlyne's psychobiological theory. *The American Journal of Psychology*, *103*, 53–80; doi: 10.2307/1423259.

McCarthy, P.M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, *42*, 381–392; doi: 10.3758/BRM.42.2.381.

McDonald, R.P. (2014). *Factor analysis and related methods* (1st ed.). New York: Psychology Press; doi: 10.4324/9781315802510.

Mehl, K., Gugliano, M., & Belfi, A.M. (2023). The role of imagery and emotion in the aesthetic appeal of music, poetry, and paintings. *Psychology of Aesthetics, Creativity, and the Arts*; Advance online publication. doi: 10.1037/aca0000623.

Menninghaus, W., Bohrn, I.C., Knoop, C.A., Kotz, S.A., Schlotz, W., & Jacobs, A.M. (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*, *143*, 48–60; doi: 10.1016/j.cognition.2015.05.026.

Menninghaus, W., Wagner, V., Wassiliwizky, E., Schindler, I., Hanich, J., Jacobsen, T., & Koelsch, S. (2019). What are aesthetic emotions? *Psychological Review*, *126*, 171–195; doi: 10.1037/rev0000135.

Miall, D.S., & Kuiken, D. (1994). Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics*, *22*, 389–407; doi: 10.1016/0304-422X(94)00011-5.

Mussel, P. (2010). Epistemic curiosity and related constructs: Lacking evidence of discriminant validity. *Personality and Individual Differences*, *49*, 506–510; doi: 10.1016/j.paid.2010.05.014.

Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., & Kotz, S. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Frontiers in Psychology*, *4*, 10. doi: 10.3389/fpsyg.2013.00010.

Rasse, C., Onysko, A., & Citron, F.M.M. (2020). Conceptual metaphors in poetry interpretation: A psycholinguistic approach. *Language and Cognition*, *12*, 310–342; doi: 10.1017/langcog.2019.47.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*, 364–382; doi: 10.1207/s15327957pspr0804_3.

Reisenzein, R. (2013). The subjective experience of surprise. In: H. Bless & J. P. Forgas (Eds.). *The message within* (pp. 262–279). New York, USA: Psychology Press.

Rinker, T.W. (2021). sentimentr: Calculate Text Polarity Sentiment.version 2.9.1. Buffalo, New York, USA. https://github.com/trinker/sentimentr.

Robinson, P. (2002). *Poetry, poets, readers: Making things happen*. New York, USA: Oxford University Press.

Runco, M.A. (1988). Creativity research: Originality, utility, and integration. *Creativity Research Journal*, *1*, 1–7; doi: 10.1080/10400418809534283.

Runco, M.A., & Jaeger, G.J. (2012). The standard definition of creativity. *Creativity Research Journal*, *24*, 92–96; doi: 10.1080/10400419.2012.650092.

Scherer, K.R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*, 695–729; doi: 10.1177/0539018405058216.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.

Silvia, P.J., & Christensen, A.P. (2020). Looking up at the curious personality: individual differences in curiosity and openness to experience. *Current Opinion in Behavioral Sciences*, *35*, 1–6; doi: 10.1016/j.cobeha.2020.05.013.

Tinio, P.P.L. (2013). From artistic creation to aesthetic reception: The mirror model of art. *Psychology of Aesthetics, Creativity, and the Arts*, *7*, 265–275; doi: 10.1037/a0030872.

Van Peer, W. (1990). The measurement of metre: Its cognitive and affective functions. *Poetics*, *19*, 259–275; doi: 10.1016/0304-422X(90)90023-X.

Wassiliwizky, E., Koelsch, S., Wagner, V., Jacobsen, T., & Menninghaus, W. (2017). The emotional power of poetry: neural circuitry, psychophysiology and compositional principles. *Social Cognitive and Affective Neuroscience*, *12*, 1229–1240; doi: 10.1093/scan/nsx069.

Whissell, C. (2018). Emotional sound symbolism and the Volta in Shakespearean and Petrarchan sonnets. *English Language and Literature Studies*, *8*, 1–10; doi: 10.5539/ELLS.V8N1P1.

Whitcomb-Hess, M. (1944). The Language of Poetry. *The Philosophical Review*, *53*, 484–492; doi: 10.2307/2181358.

Soma Chaudhuri, Alan Pickering, Joydeep Bhattacharya, University of London

*Correspondence concerning this article should be addressed to Soma Chaudhuri, Department of Psychology, Goldsmiths, University of London, New Cross, London SE14 6NW, UK. E-mail: schau002@gold.ac.uk

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Comparison of models (M1 to M7) for predicting creativity and aesthetic appeal, showing Akaike information criteria (AIC), Bayesian information criteria (BIC), R^2 (marginal), and change in chi-square values ($\Delta\ \chi^2$).

Table S2. Moderation results of regression models predicting creativity and aesthetic appeal, including estimates, standard errors (SE), *t*-Values, *p*-Values, and model fit (R^2), for each personality trait-based model (openness, intellect, curiosity, VVIQ, and AVIQ).