Psychology of Aesthetics, Creativity, and the Arts
2018, Vol. 12, No. 1, 24–33

1931-3896/18/$12.00 http://dx.doi.org/10.1037/aca0000099

A Meta-Analysis of the Relationship Between Intelligence and Visual "Taste" Measures

Nils Myszkowski
Pace University

Pinar Çelik and Martin Storme
Université Paris Descartes - Sorbonne Paris Cité

What makes individuals experts in judging aesthetic value is actively researched in a variety of ways. In the visual domain, one classical paradigm—used in "T" (for *Taste*) tests (Eysenck, 1983)—consists in comparing one's evaluative judgments of beauty with a standard judgment—provided by consensual or expert agreement. The association between general intelligence (g) and performance in "T" tests has been investigated since over 70 years (Eysenck, 1940; Myszkowski, Storme, Zenasni, & Lubart, 2014), but has led to a variety results, from negative weak to positive strong correlations. We aimed at clearing the resulting confusion through a meta-analysis of the correlations observed in the literature (k = 23, N = 1,531). We found a significant positive weak to moderate correlation between g and "T" (ρ = .30, 95% CI (confidence interval) = [.23, .36], z = 9.00, p < .001), suggesting that common cognitive processes are involved in both g and "T". Reinforcing this conclusion, no publication bias was found through the regression test, and none of the tested moderators—year of publication, gender, age, "T" measure, and g measure—had a significant effect on the correlation.

Keywords: aesthetic sensitivity, taste, intelligence, aesthetic ability, meta-analysis

Despite the difficulty of finding objectivity in aesthetic judgments, empirical aesthetics, and psychology researchers have relentlessly attempted to identify the features of consensually preferred stimuli (e.g., Locher & Nodine, 1989; Marković & Gnozdenovi, 2001; Reber, 2002; Silvia & Barona, 2009; Vitz, 1966). While a lot of this research has been focused on aesthetic universals (e.g., Arnheim, 1954; Graves, 1951) and their flexibility (e.g., McManus, 2005), the *person* perspective (Jacobsen, 2006) is interested in individual differences in aesthetic judgments (e.g., Barron, 1953; Eysenck & Furnham, 1993; Furnham & Avison, 1997; Furnham & Bunyan, 1988; Rawlings, Twomey, Burns, & Morris, 1998). A special—and controversial (Eysenck, 1997; Gear, 1986)—approach to these individual differences consists in studying the extent to which some individuals may be *better* equipped to judge aesthetic stimuli.

What Is "T"?

Although there are different approaches to aesthetic ability—including aesthetic expertise (Chatterjee, Widick, Sternschein, Smith, & Bromberger, 2010; Plucker, Kaufman, Temple, & Qian, 2009; Silvia, 2007; Smith & Smith, 2006), sensitivity to complexity (Barron & Welsh, 1952), aesthetic chills (Silvia & Nusbaum, 2011), and exploration tendencies (Nodine, Locher, & Krupinski, 1993)—one that has recently regained interest (Chamorro-Premuzic & Furnham, 2004; Furnham & Chamorro-Premuzic, 2004; Myszkowski, Storme, & Zenasni, 2016; Myszkowski et al., 2014; Summerfeldt, Gilbert, & Reynolds, 2015) is the aesthetic sensitivity—or "T"—approach. This approach consists in measuring and studying the ability to identify the aesthetic value of stimuli by directly presenting stimuli that vary in aesthetic quality, such quality being either measured through consensual or expert agreement (Child, 1962; Eysenck, 1940), or manipulated through the "controlled alteration" (Meier, 1928, p. 188) of stimuli (that consists in altering the stimuli to make it of lesser aesthetic quality). Despite the psychometric challenges that such an approach raises, notably in terms of content validity (Myszkowski et al., 2014), researchers have recently found interest in "T"'s suggested relations with various dispositional variables, including personality traits (Chamorro-Premuzic & Furnham, 2004; Frois & Eysenck, 1995; Myszkowski et al., 2014), creativity (Myszkowski et al., 2014), artistic training (Frois & Eysenck, 1995), obsessive–compulsive disorder tendencies (Summerfeldt et al., 2015), and intelligence (Bezruczko & Frois, 2011; Myszkowski et al., 2014).

Although "T" previously existed (Meier, 1928), evidence supporting the relevance of the "T" factor approach was found in the early 1940s, when, using exploratory factor analyses of the judgments of a variety of aesthetic stimuli, a general factor was found to explain most of the variability in individual visual aesthetic preferences across wide arrays of stimuli (Eysenck, 1940, 1941a, 1941b, 1942, 1968). Originally named "T" for good Taste (Ey-

This article was published Online First December 12, 2016.

Nils Myszkowski, Department of Psychology, Pace University; Pinar Çelik and Martin Storme, Laboratoire Adaptations Travail-Individu, Université Paris Descartes - Sorbonne Paris Cité.

Correspondence concerning this article should be addressed to Nils Myszkowski, Department of Psychology, Pace University, 41 Park Row, New York, NY 10038. E-mail: nmyszkowski@pace.edu

senck, 1983), this "T" factor translates Eysenck's findings—and later Child's (1962)—that individuals differ in their general degree of agreement with consensual aesthetic judgments of stimuli of various nature. In other words, people tend to agree with one another on what they consider of superior aesthetic value, and the individuals who agree more with such consensual preferences tend to do so across a wide variety of types of stimuli. Thus, people differ *consistently* across most visual domains in their degree of agreement with consensual judgments: That degree of agreement is "T." Later, "T" started becoming referred to as "aesthetic sensitivity" (Child, 1962, 1964), and was defined by Child (1964, p. 49) as the "the extent to which a person gives evidence of responding to relevant stimuli in some consistent and appropriate relation to the external standard."

Since Meier's early development of the controlled alteration method (Meier, 1928), three main measures of aesthetic sensitivity have been developed with this method and studied: The Meier Art Tests (Meier, 1940, 1963), the Design Judgment Test (Graves, 1948, 1951), and the Visual Aesthetic Sensitivity Test (Götz, 1985; Götz, Borisy, Lynn, & Eysenck, 1979); all three are still used in current research (Chamorro-Premuzic & Furnham, 2004; Furnham & Chamorro-Premuzic, 2004; Myszkowski et al., 2014). Although alternate methods to measure "T" have been discussed, like Lifton's (1961) suggestion to ask participants to indicate whether two paintings were of the same artist, or like the PJB Test (Bamossy, Scamoon, & Johnston, 1983), and although some of the initial "T" measures did not involve stimulus manipulation (Child, 1962; Eysenck, 1940), the measure of "T" through standardized testing is typically achieved through the "controlled alteration" method (Meier, 1928). In the case of these tests, pairs or triads of similar visual stimuli are created, one stimulus remaining unaltered, while the other(s) are altered to present defaults—support for content validity may be found in the consensus among a community sample, and/or in the agreement of experts (Götz, 1985). Study participants have to indicate which of the two or three presented designs is of the best quality (Götz et al., 1979; Myszkowski et al., 2014) or most appealing (Child, 1962).

"T" and g

At the time that "T" was conceptualized, the dimensionality of intelligence was at the heart of the debate for individual differences psychology researchers (Burt, 1940; Eysenck, 1939; Spearman, 1939; Thurstone, 1938), some of them—particularly Eysenck—being involved in research on both the general factor of intelligence, g, and "T" (Myszkowski et al., 2016). Therefore, it is not surprising that, since then, not only g and "T" have been investigated using similar statistical methods—both g and "T" strongly rely on exploratory factor analyses in their original conceptualization (Eysenck, 1941a, p. 91; Spearman, 1939)—but also possible connections between g and "T" have been considered from the beginning: In the early characterization of "T," Eysenck described "T" as "merely the manifestation of "g" when aesthetic material is used" (Eysenck, 1940, p. 101), leading to early searches for correlations between g and "T."

Common Processes

Eysenck's description of good taste as a facet of intelligence is cavalier—"T" and g measures being after all very different in nature. Nevertheless, there are cognitive processes involved in psychometric intelligence that "T" tasks may tap into.

Attention shifting. A first example of such common processes could be the inhibition of nonrelevant features and the attention-shift to relevant ones. Indeed, when presented with an aesthetic stimulus during a "T" task, the capacity of individuals to shift attention to specific relevant features of the stimulus may be crucial to their performance. Typically, in the case of the "controlled alteration" method used for "T" tests (e.g., Götz, 1985), to better perform, individuals have to shift their attention, from the attributes that are equivalent in both stimuli to the features that were altered—the very locations of the altered stimulus' flaws. This tentative explanation of attention-shifting as a process common to g and "T" is consistent with the finding that art-trained viewers are better able to spot differences between an original stimulus and its altered version (Nodine et al., 1993), consistent with the identification of attention-shifting as a key element in the perceptual advantage of artists (Kozbelt & Seeley, 2007), and consistent with the finding that inhibiting tasks that require attention-shifting (e.g., Stroop tasks) are correlated with g measures (Friedman et al., 2006).

Reflective processing. Related to the idea that inhibition may be an important common process between g and "T," and in line with the idea that art expertise is characterized by the tendency to adopt an aesthetic stance that is "emotionally distanced" (Leder, Gerger, Brieber, & Schwarz, 2014, p. 1138) when judging aesthetic stimuli, the observed correlations between g and "T" could be partially due to the ability to inhibit a mode of aesthetic processing that is focused on direct emotional arousal and the protection of existing knowledge, to facilitate a different mode of processing, which, although partially relying on emotions, would be more analytic and focused on knowledge acquisition. In this view, Cupchik (1995) contrasts the *reactive model* of aesthetic processing—where pleasure and emotional arousal have the most importance—with the *reflective model*— where emotional responses are elements that help in the analysis and interpretation of stimuli. This distinction can be related to the distinction made by Armstrong and Detweiler-Bedell (2008) between pursuing *prevention goals* (seeking to prevent confusion and maintain existing knowledge) and *promotion goals* (seeking to expand knowledge) when judging aesthetic stimuli. However, it should not be excluded that emotional factors play a role in "T" measures, and it should not be considered that higher "T" scores are achieved through the total inhibition of emotion: Processing complex aesthetic stimuli is not simply a problem-solving task (Muth, Hesslinger, & Carbon, 2015), as it notably relies to a great extent on affective rewards associated to experiencing insights in judging stimuli. In this view, recent research relating personality traits to "T" measures (Myszkowski et al., 2014) notably indicated that openness to aesthetics, openness to fantasy, openness to feelings, and sensation seeking are related to higher "T" scores.

Goal management. Another potential process that could be commonly implied in g and "T" measures is goal management. Indeed, goal management—spawning subgoals from goals and tracking the success or not in pursuing each subgoal to achieve a higher goal—underlies success in g tests (Carpenter, Just, & Shell, 1990), and may also be implied in "T" type tasks. It can be argued that participants who are more able to spawn subgoals from a "T" test item (e.g., successively analyzing and evaluating symmetry, organizational balance, line breaks, contrast, etc.) are better able to deconstruct the very process of evaluating the aesthetic quality of a stimulus into separate steps, thus focusing more distinctly on the different aspects, and achieving better performance in identifying flaws at each step: This could consequently lead to a higher ability to achieve the higher goal of identifying aesthetic quality, leading to higher performance in "T" tests. Once again, this should be nuanced by the nonequivalence between an aesthetic judgment task and a problem-solving task (Muth et al., 2015): It is possible that, when facing the aesthetic stimuli used in "T" tests, the ability to adopt a *gestalt* approach, rather than to decompose a problem into subproblems, is more efficient—especially when the "T" task is presented with a time limit. Alternatively, it is also possible that higher performance in "T" tasks is achieved when the subgoals that are pursued are not purely the result of a cognitive "management" process, but when instead individuals follow insights about what subgoals need to be addressed primarily.

Abstraction. Finally, we could also propose abstraction—that consists of "the construction of representations that are only loosely tied to perceptual inputs and instead are more dependent on high-level interpretations of inputs that provide a generalization over space and time" (Carpenter et al., 1990, p. 428)—as a cognitive process that is both involved in psychometric intelligence tests and in "T." Indeed, as participants analyze stimuli, they may need to detach themselves from direct perceptual inputs to use abstract rules and use the *structural skeleton* (Arnheim, 1954) as a stronger basis for judgment. In this view, art experts are often found to be more able to focus on the elements of an aesthetic stimulus that are less directly available: They for example direct their attention more toward the relationships between objects than on the objects as individual elements (Nodine et al., 1993), and to focus on the order and dynamics of visual structure (Winston & Cupchik, 1992). In other words, individuals with higher "T" levels may use inputs of higher-order than individuals with lower levels, which would direct their attention on inputs that are more directly available in the stimulus.

Observed correlations

The question of a possible correlation between g and "T" has been investigated multiple times in the visual domain, on various samples (children and adults), and with various measures of both constructs (Bezruczko & Frois, 2011; Chamorro-Premuzic & Furnham, 2004; Child, 1962; Eysenck, 1940; Frois & Eysenck, 1995; Furnham & Chamorro-Premuzic, 2004; Götz, 1987; Götz et al., 1979; Myszkowski et al., 2014). However, it has not yet led to a clear answer: A range of negative or positive weak and nonsignificant correlations to strong positive correlations have been observed (Bezruczko & Frois, 2011; Child, 1962; Frois & Eysenck, 1995). These studies were conducted with relatively small sample sizes, with the largest sample size being 130 participants (Myszkowski et al., 2014). Consequently, the literature provides contradictory results about the relationship between g and "T" with approximately half of the studies yielding significant positive correlations and the other half yielding nonsignificant correlations.

Not only have the observed correlations and their significance varied, but the conclusions about the relationship between the two constructs have also varied, from discarding intelligence as an "irrelevant feature of the individual" (Eysenck, 1983, p. 230) for judgments made with "T" measures, to seeing in these correlations a sign that intelligence may facilitate aesthetic judgment through cognitive processes (Myszkowski et al., 2014). Facing mixed evidence for the direction and magnitude of the correlation between the two constructs, researchers have been—and are still—cautious in their conclusions, discussing psychometrical challenges (Bezruczko & Frois, 2011; Furnham & Chamorro-Premuzic, 2004; Iwawaki, Eysenck, & Götz, 1979; Myszkowski et al., 2014).

The Present Research

Our aim in this article was to provide, through a systematic literature review and meta-analysis, an estimate for the magnitude and direction of the correlation between the general factor of intelligence, g, and "T." As we explained earlier, there is a lot of diversity in the research on the topic: Studies have been conducted (a) with various measures of g, (b) various measures of "T," (c) on very different samples of children and adults from different countries, (d) have led to a wide range (from negative weak to positive strong) of correlations, which (e) has itself led to a wide array of conclusions.

As we have earlier explained that common cognitive processes may be involved in "T" and g measures, we hypothesized that "T" measures are positively correlated with intelligence measures. We believe that a demonstration through systematic review and meta-analysis can clear the confusion created by such heterogeneous research, provide guidelines for future research, as indicate that intelligence should not be discarded as playing a role in "T" measures.

Method

Literature Search

The PsycArticles, PsycInfo, and ERIC databases were searched independently by two researchers, using the following keywords, previously identified during a prereview of the literature: "(A)esthetic ability," "(A)esthetic judgment," "(A)esthetic sensitivity," "Art judgment," and "Art ability." Furthermore, three additional searches were performed, with in each, as a keyword, one of the three main tests of visual "T": the Visual Aesthetic Sensitivity Test (Götz, 1985), the Design Judgment Test (Graves, 1948), and the Meier Art Tests (Meier, 1940, 1963). In addition, all reference lists of the obtained articles were inspected for studies that could have been missing from the databases. The titles and abstracts were screened for the potential inclusion of a "T" and a g measure. The search was not date restricted, but the publication date of the studies that met the inclusion criteria ranged between 1940 and 2014. The search was not country, age, culture, or language specific.

Inclusion Criteria

All the studies that were included in the meta-analysis had to meet the following criteria: (a) The studies included a judgment measure—either an ad hoc measure (e.g., Eysenck, 1940) or a standardized test (Götz, 1985; Graves, 1948; Meier, 1940, 1963)—of "T" that *directly* operationalized Child's definition

(Child, 1964, p. 49), which consists, as we earlier explained, of asking individuals to judge or compare stimuli based on their beauty or aesthetic pleasantness, and then confronting their judgments with external standards; (b) the stimuli that were used to measure "T" had to be visual stimuli of artistic or decorative nature (e.g., the participants were not asked to rate the beauty of individuals); (c) the research included an intelligence test of any sort; (d) regarding the relation between "T" and g, enough information (typically the correlation coefficient and sample size) was reported to estimate effect size and its variance; and (e) the participants were not sampled from a clinical population. The total number of included studies was 23, and the pooled total sample size was 1,531. The summary of the different measures used to measure "T" and g in the included studies is reported in Table 1.

Data Input

We input the effect size and sample size in the R-packages '*metafor*' (Viechtbauer, 2010; Viechtbauer & Viechtbauer, 2015) and '*meta*' (Schwarzer, 2007; Schwarzer & Schwarzer, 2015). All but one of the included studies reported correlation coefficients and sample sizes, which were consequently directly input.

The study (Chamorro-Premuzic & Furnham, 2004) that did not report a g–"T" correlation coefficient investigated the relation between g and "T" through hierarchical regression (incrementally adding predictors to a regression model). In this study, the researchers entered intelligence as a predictor of "T" after having previously added art interests. They reported the incremental part of variance explained (R^2-Change) because of entering intelligence. We transformed the R^2-Change into a correlation coefficient metric by using its root square, before including it in the meta-analysis. Even though this study was not an outlier in the meta-analysis, it can be noted that the presence of art interests as predictors in the regression models of this study may have led to slightly misestimating the correlation between g and "T."

Whenever the studies investigated and presented g–"T" correlations as different results obtained on different samples (Child, 1962; Frois & Eysenck, 1995; Götz, 1987; Götz et al., 1979), we included the results on each sample directly, without pooling them (because the samples were independent). In same-sample situations, however, we averaged correlation coefficients: When, in the same sample, the correlation between g and "T" was reported when "T" was both measured using the agreement with experts and with the agreement with consensus (Child, 1962), or when, in the same sample, a correlation between "T" and two different intelligence tests was presented (Bezruczko & Frois, 2011), we input the average of the two correlations and sample size.

Statistical Analyses

The meta-analysis was performed using the '*metacor*' function of the R-package '*meta*' (Schwarzer, 2007; Schwarzer & Schwarzer, 2015) and with the '*rma.uni*' function of '*metafor*' (Viechtbauer, 2010; Viechtbauer & Viechtbauer, 2015). The reasons why two different packages were used were that they had different capabilities: Specifically, "*meta*" allowed more control over the plots presented in Figure 1 and Figure 2, but did not allow moderator analyses, which '*metafor*' allowed. Besides, using two packages allowed the results of our statistical procedure to be cross validated.

Model. The heterogeneity in the research methods used—that we earlier underlined—clearly called for choosing a priori the random-effects estimation method of the correlation between g and "T." For the random-effects model, we used the Hartung-Knapp-Sidik-Jonkman (Hartung, 1998; Hartung & Knapp, 2001a, 2001b; Sidik & Jonkman, 2005) estimation method, which has been showed to outperform the classic DerSimonian-Laird method (DerSimonian & Laird, 1986), especially in cases similar to ours where the number of studies is small ($k = 23$) and there is heterogeneity (IntHout, Ioannidis, & Borm, 2014; Sidik & Jonkman, 2005).

Heterogeneity. Heterogeneity was measured through Higgins and Thompson's I^2 (Higgins & Thompson, 2002; Higgins, Thompson, Deeks, & Altman, 2003), which has been shown to be a more accurate and straightforward estimate of heterogeneity that is because of variation between studies than Cochran's Q (Cochran, 1954).

Table 1
Summary of the Measures Used in the Included Studies

Publication	g measure	"T" measure
Eysenck (1940)	Verbal intelligence test (White, 1931)	Ad hoc measure
Child (1962)	"Scholastic Aptitude Test Verbal" (not referenced)	Ad hoc measure
Götz et al. (1979)	"IQ" (not further specified or referenced)	Visual Aesthetic Sensitivity Test (Götz et al., 1979)
Götz (1987)	Standard Progressive Matrices (Raven, 1941)	Visual Aesthetic Sensitivity Test (Götz, 1985)
Frois and Eysenck (1995)	Standard Progressive Matrices (Raven, 1941)	Visual Aesthetic Sensitivity Test (Götz, 1985)
Furnham and Chamorro-Premuzic (2004)	Wonderlic Personnel Test (Wonderlic, 1992)	Design Judgment Test (Graves, 1948)
Chamorro-Premuzic and Furnham (2004)	Wonderlic Personnel Test (Wonderlic, 1992) and Standard Progressive Matrices (Raven, 1941)	Meier Art Tests (Meier, 1942)
Bezruczko and Frois (2011)	Collective Scale of Intellectual Level (Benedetto & Clerc, 1969) and Standard Progressive Matrices (Raven, 1941)	Visual Aesthetic Sensitivity Test (Götz, 1985)
Myszkowski et al. (2014)	Standard Progressive Matrices (Raven, 1941)	Visual Aesthetic Sensitivity Test (Götz, 1985)

Study and sample (sorted by year of publication)	Sample size	r	95% C.I.	Weight (Random)
Eysenck (1940)	18	0.25	[-0.25; 0.64]	1.8%
Child (1962) - Stanford	22	0.35	[-0.08; 0.68]	2.2%
Child (1962) - Yale	22	0.12	[-0.32; 0.52]	2.2%
Götz et al. (1979) - Boys	204	0.07	[-0.07; 0.21]	7.5%
Götz et al. (1979) - Girls	165	0.20	[0.05; 0.34]	7.0%
Götz (1987) - Private Gymnasium	102	0.37	[0.19; 0.53]	5.9%
Götz (1987) - Commercial College	73	0.62	[0.45; 0.74]	5.1%
Frois & Eysenck (1995) - Boys (10)	61	0.33	[0.09; 0.54]	4.6%
Frois & Eysenck (1995) - Boys (11)	49	0.28	[0.00; 0.52]	4.0%
Frois & Eysenck (1995) - Boys (12)	40	0.21	[-0.11; 0.49]	3.5%
Frois & Eysenck (1995) - Boys (13)	39	0.34	[0.03; 0.59]	3.4%
Frois & Eysenck (1995) - Boys (14)	36	0.21	[-0.13; 0.50]	3.2%
Frois & Eysenck (1995) - Boys (15)	40	0.13	[-0.19; 0.42]	3.5%
Frois & Eysenck (1995) - Girls (10)	50	0.26	[-0.02; 0.50]	4.1%
Frois & Eysenck (1995) - Girls (11)	49	0.29	[0.01; 0.53]	4.0%
Frois & Eysenck (1995) - Girls (12)	55	0.45	[0.21; 0.64]	4.3%
Frois & Eysenck (1995) - Girls (13)	51	0.07	[-0.21; 0.34]	4.1%
Frois & Eysenck (1995) - Girls (14)	50	0.25	[-0.03; 0.49]	4.1%
Frois & Eysenck (1995) - Girls (15)	51	0.50	[0.26; 0.68]	4.1%
Furnham & Chamorro-Premuzic (2004)	74	0.28	[0.06; 0.48]	5.1%
Chamorro-Premuzic & Furnham (2004)	102	0.33	[0.14; 0.49]	5.9%
Bezruczko & Frois (2011)	48	0.50	[0.25; 0.68]	4.0%
Myszkowski et al. (2014)	130	0.27	[0.10; 0.42]	6.5%
Fixed effect model	**1531**	**0.28**	**[0.23; 0.33]**	–
Random effects model		**0.30**	**[0.23; 0.36]**	**100%**

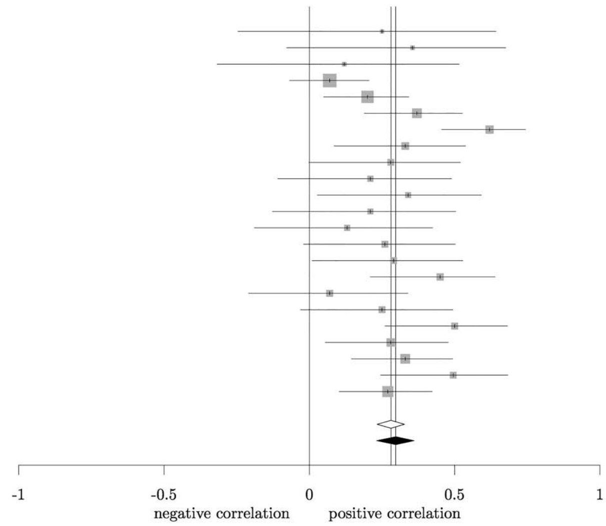

Figure 1. Forest plot with 95% confidence intervals and weights.

Publication bias. Because the g–"T" relation is typically presented as a "side" result in many studies (e.g., Bezruczko & Frois, 2011; Child, 1962; Eysenck, 1940), there was no reason to expect that there could be a substantial publication bias. However, we investigated publication bias empirically through both the visual inspection of the contour-enhanced funnel plot (Peters, Sutton, Jones, Abrams, & Rushton, 2008) and the Egger regression test for publication bias (Sterne & Egger, 2005).

Additional moderation analyses. Although the main focus of this meta-analysis was the correlation between g and "T," and although the reviewed literature did not indicate that the relation between g and "T" could be moderated by any factor, we decided to perform additional moderation analyses using '*metafor*,' with the same estimation methods explained earlier. The reason for exploring for such moderation effects was to gather information on whether or not the correlation between the two constructs was generalizable to a variety of samples and measures. In other words, we wanted here to explore if there were signs that the heterogeneity of the observed correlations could be because of different characteristics of the samples and different methods, or if it could be imputable to sampling error.

We tested the moderation effect of (a) the year of publication, (b) the gender of the participants—some articles (Frois & Eysenck, 1995; Götz et al., 1979) reported correlations separately for males and females—(c) the mean age of the participants—that was not always available (Child, 1962; Eysenck, 1940), although when only the range was reported (Chamorro-Premuzic & Furnham, 2004; Furnham & Chamorro-Premuzic, 2004), we used the mean between the maximum and minimum— (d) the type of "T" measure—an ad hoc measure (Child, 1962; Eysenck, 1940) or a standardized test (Götz, 1985; Götz et al., 1979; Graves, 1948; Meier, 1940, 1963)—and (e) the type of g measure—verbal, nonverbal, or composite.

Results

g–"T" Correlation

We found a significant positive weak to moderate correlation between g and "T" (ρ = .30, 95% confidence interval (CI) [.23, .36], z = 9.00, $p < .001$).

The forest plot presented in Figure 1 gives an overview of the results of the meta-analysis. It presents each study through its effect size (the central point), the 95% CI of its effect size (the horizontal segment), and its weight in the meta-analytic random effects model (the size of the gray squares surrounding the central point). The fixed and random effects estimates of the correlation and their confidence intervals (respectively, the white and the black diamonds) are also plotted.

The weights of the studies did not indicate that a few studies contributed substantially more than the others, which could have biased the effect size estimate. However, the forest plot shows heterogeneity in effect size and statistical significance across the studies, with a number of confidence intervals including 0. The percentage of heterogeneity because of variation between studies was moderate (I^2 = 42.5%, 95% CI [5.7%, 65%]), which was expected, considering not only the inconsistency in the observed correlations that led to this meta-analysis, but also all the differences between these studies, in terms of procedures and sample characteristics.

The contour-enhanced funnel plot (reported in Figure 2) plots the size of each study—as recommended (Sterne & Egger, 2001), *SE* was used as a measure of study size—as a function of its observed effect size. It is inspected for asymmetry, potentially indicating a relation between study size and effect size, as well as for the potential suppression of nonsignificant findings. Here, the plot showed no asymmetry, as confirmed by the nonsignificance of the Egger regression test ($t(21)$ = 1.20, p = .24), and did not suggest the suppression of nonsignificant findings. These results indicate no evidence of publication bias.

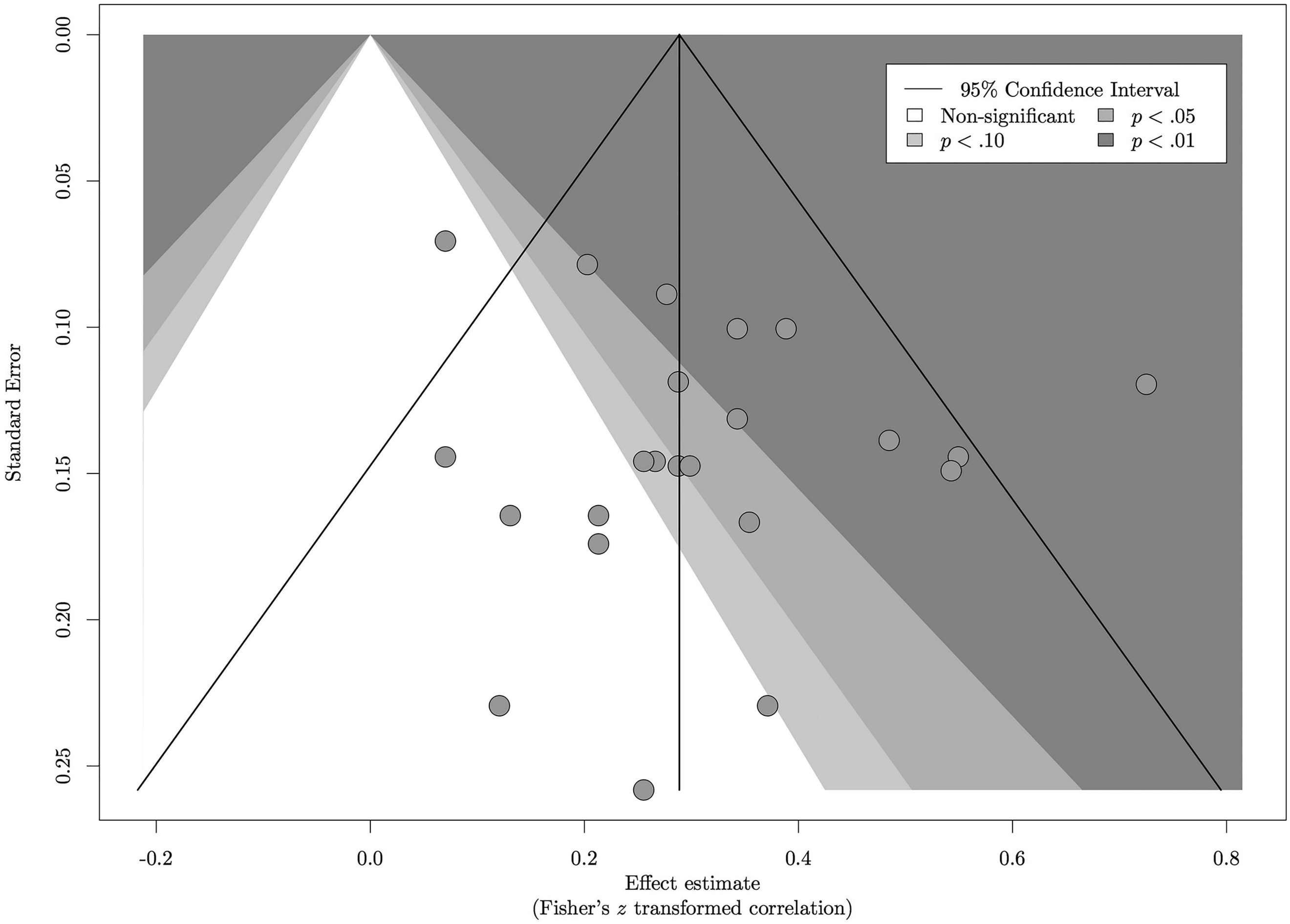

Figure 2. Contour-enhanced funnel plot.

Moderators

None of the tested moderators—the year of publication, $F(1, 21) = 0.98$, $p = .33$, the gender of the participants, $F(1, 12) = 1.27$, $p = .28$, the age of the participants, $F(1, 18) = 0.29$, $p = .60$, the type of "T" measure, $F(1, 21) = 0.67$, $p = .67$, the type of g measure, $F(2, 20) = 1.27$, $p = .52$—had a significant effect on the correlation between the two constructs. Although these tests do not confirm null relationships (only the inability to reject null relationships), the obtained results suggest that the correlation between g and "T" does not depend on any of these characteristics.

Discussion

For more than 70 years (Eysenck, 1940; Myszkowski et al., 2014), researchers have been interested in the correlation between intelligence and "T" scores in the visual domain. They have been largely confused by the variety—in terms of magnitude, statistical significance and even direction (e.g., Child, 1962)—of the correlations that have been observed between the two constructs.

We believe that the present results, obtained through a systematic review and meta-analysis, have cleared the confusion: Although about half of the studies found nonsignificant correlations, our effect size estimate allows to conclude that there is a positive correlation between g and "T" measures, and that it is not of negligible magnitude (Cohen, 1988). More specifically, the estimated confidence intervals indicate that the correlation is small to moderate. We can add that it appears from our publication bias analyses that this result is probably not because of publication bias. Furthermore, no significant part of the heterogeneity in the observed correlations was explained by either year of publication, gender, age, the type of "T" test, or the type of g test. This strengthens our point, as it suggests that the variability in the observed correlations is not due to the characteristics of the sample or the procedure: To the contrary, it suggests that such heterogeneity could be imputable to sampling error.

Implications

Although this is a meta-analytic investigation of correlational results, the most important implication of this study is that it suggests that common processes are implied in both psychometric intelligence and "T" measures. The present results do not indicate which processes these may be, but the fact that the type of g measure (verbal, nonverbal, or composite) had no significant moderation effect on the g–"T" correlation indicates that these processes are probably not specific to one single type of task.

Concerning these common g–"T" processes, when introducing the topic of this research, we advanced a (nonexhaustive) list of possibilities. Although it appears from the very instructions used (e.g., Götz, 1985) that "T" measures do not necessitate high abilities in maintaining and updating information in working memory—that are often described as the center of focus of psychometric intelligence tests (Belacchi, Carretti, & Cornoldi, 2010; Friedman et al., 2006)—psychometric intelligence tests actually involve other processes that "T" measures could tap into. More specifically, we proposed that "T" involves one or more of the following processes also involved in g (Carpenter et al., 1990; Friedman et al., 2006): Attention-shifting—high "T" individuals would be more able to shift their attention toward different elements, while low "T" individuals would remain more "stuck" on some elements that are not necessarily relevant—reflective processing—high "T" individuals would be able to adopt a mode of processing that aims at extending knowledge, while low "T" individuals would adopt a mode of processing that seeks direct emotional arousal—goal management—high "T" individuals would be more able to spawn subgoals from the task, being thus able to focus specifically on aesthetic features than low "T" individuals—and abstraction—high "T" individuals would be more focused on higher-order features, the *structural skeleton*, of the aesthetic stimuli, while low "T" individuals would focus more on the individual objects that are more directly available in the stimulus. Future research could investigate the individual cognitive processes involved in "T" tasks that could explain the correlation through our meta-analysis, using either a psychometrical approach—correlating individual differences in "T" with the performance at tasks that specifically measure the abilities associated with each cognitive process—or an experimental approach—using for example a dual-task paradigm, where individuals would have to perform two tasks simultaneously, one being a "T" type task, and the other task involving a specific cognitive process that would reduce the cognitive resources available for the "T" task.

We could add that, on one hand, recent research (Myszkowski et al., 2014) has indicated that, in addition to being correlated with g, "T" measures may be related to both openness and creativity, reinforcing the idea that artistic reception and creation use the same structural framework (Tinio, 2013), and that artists have higher perceptual skills (Kozbelt & Seeley, 2007; Kozbelt, Seidel, ElBassiouny, Mark, & Owen, 2010); on the other hand, creativity, openness, and intelligence have been found to be also correlated (Feist, 1998; Gignac, Stough, & Loukomitis, 2004; Kim, 2008; Plucker, Esping, Kaufman, & Avitia, 2015). The present evidence for a correlation between g and "T" therefore may be explained by potential presence of a common "intellect" or "achievement" factor behind altogether g, "T," creativity, and openness: We think that further investigation on how "T" relates to openness and creativity will enrich the understanding of how it relates to g.

Limitations

Our meta-analytic investigation has limitations. First, our investigation was only carried over studies that measured "T" in the visual domain with a specific procedure, which consists in comparing participants' aesthetic judgments with standard judgments provided by consensual agreement or experts. This restricts our conclusion to that type of measure in particular, which is certainly quite specific and cannot fully encompass aesthetic ability, expertise, or fluency. As was recently noted (Myszkowski & Zenasni, 2016), the current state of research on these different constructs calls for a more multifactorial approach of aesthetic capacity—that would include not only "T" tests, but also the depth of aesthetic exploration tendencies (Koide, Kubo, Nishida, Shibata, & Ikeda, 2015; Nodine et al., 1993), art knowledge (Silvia, 2007; Smith & Smith, 2006), sensitivity to complexity (Barron & Welsh, 1952), also named "K" (Eysenck, 1941b; Myszkowski et al., 2016), and the tendency to experience aesthetic chills (Silvia & Nusbaum, 2011)—that would probably altogether more accurately and comprehensively represent the ability to judge art, or "good taste," than comparisons of individual taste with common or expert taste. Another limitation is that "T" measures are known to face psychometric challenges (Liu, 1990), notably a lack of reliability (Bezruczko & Frois, 2011), which could have attenuated many of the correlations that were observed. Indeed, it may be that g and "T" are even more associated than we think, and future research should aim at improving "T" measures to gain a better accuracy in the estimation of the correlation between the two constructs.

Conclusion

The scatteredness of the literature on the relations between g and "T" called for the meta-analytic review provided here. Despite such variety, our review and meta-analysis have clearly indicated that there is a correlation that researchers should not overlook. Despite a very heterogeneous body of research, the main result is rather straightforward and leaves little doubt as to the presence of a correlation. However, we do not believe that researchers should consider the relation between g and "T" as completely established. To the contrary, the results of this meta-analysis raise new questions, and call for the further investigation of the cognitive processes involved in "T" measures.

References

Armstrong, T., & Detweiler-Bedell, B. (2008). Beauty as an emotion: The exhilarating prospect of mastering a challenging world. *Review of General Psychology, 12,* 305–329. http://dx.doi.org/10.1037/a0012558

Arnheim, R. (1954). *Art and visual perception: A psychology of the creative eye*. Oakland, CA: University of California Press.

Bamossy, G., Scamoon, D. L., & Johnston, M. (1983). A preliminary investigation of the reliability and validity of an aesthetic judgment test. *Advances in Consumer Research Association for Consumer Research, 10,* 685–690.

Barron, F. (1953). Complexity-simplicity as a personality dimension. *The Journal of Abnormal and Social Psychology, 48,* 163–172. http://dx.doi.org/10.1037/h0054907

Barron, F., & Welsh, G. S. (1952). Artistic perception as a possible factor in personality style: Its measurement by a figure preference test. *The Journal of Psychology, 33,* 199–203. http://dx.doi.org/10.1080/00223980.1952.9712830

Belacchi, C., Carretti, B., & Cornoldi, C. (2010). The role of working memory and updating in Coloured Raven Matrices performance in typically developing children. *European Journal of Cognitive Psychology, 22,* 1010–1020. http://dx.doi.org/10.1080/09541440903184617

Benedetto, P., & Clerc, P. (1969). Enquête nationale sur le niveau intellectuel des enfants d'age scolaire [National Survey of the intellectual levels of children of school age]. *Population, 24,* 1123. http://dx.doi.org/10.2307/1529737

Bezruczko, N., & Frois, J. P. (2011). Comparison of several artistic judgment aptitude dimensions between children in Chicago and Lisbon. *Visual Arts Research, 37,* 1–15. http://dx.doi.org/10.5406/visuartsrese.37.1.0001

Burt, C. (1940). *The factors of the mind: An introduction to factor-analysis in psychology*. London, England: University of London Press, Limited.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97,* 404–431. http://dx.doi.org/10.1037/0033-295X.97.3.404

Chamorro-Premuzic, T., & Furnham, A. (2004). Art judgment: A measure related to both personality and intelligence? *Imagination, Cognition and Personality, 24,* 3–24. http://dx.doi.org/10.2190/U4LW-TH9X-80M3-NJ54

Chatterjee, A., Widick, P., Sternschein, R., Smith, W. B., & Bromberger, B. (2010). The assessment of art attributes. *Empirical Studies of the Arts, 28,* 207–222. http://dx.doi.org/10.2190/EM.28.2.f

Child, I. L. (1962). Personal preferences as an expression of aesthetic sensitivity. *Journal of Personality, 30,* 496–512. http://dx.doi.org/10.1111/j.1467-6494.1962.tb02319.x

Child, I. L. (1964). Observations on the meaning of some measures of esthetic sensitivity. *The Journal of Psychology, 57,* 49–64. http://dx.doi.org/10.1080/00223980.1964.9916671

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics, 10,* 101–129. http://dx.doi.org/10.2307/3001666

Cohen, J. (1988). *Statistical power analysis for the behavioral* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cupchik, G. C. (1995). Emotion in aesthetics: Reactive and reflective models. *Poetics, 23,* 177–188. http://dx.doi.org/10.1016/0304-422X(94)00014-W

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7,* 177–188. http://dx.doi.org/10.1016/0197-2456(86)90046-2

Eysenck, H. J. (1939). Review of L. L. Thrustone's "Primary Mental Abilities." *The British Journal of Educational Psychology, 9,* 270–275. http://dx.doi.org/10.1111/j.2044-8279.1939.tb03214.x

Eysenck, H. J. (1940). The general factor in aesthetic judgements. *The British Journal of Psychology General Section, 31,* 94–102. http://dx.doi.org/10.1111/j.2044-8295.1940.tb00977.x

Eysenck, H. J. (1941a). The empirical determination of an aesthetic formula. *Psychological Review, 48,* 83–92. http://dx.doi.org/10.1037/h0062483

Eysenck, H. J. (1941b). "Type" factors in aesthetic judgements. *The British Journal of Psychology General Section, 31,* 262–270. http://dx.doi.org/10.1111/j.2044-8295.1941.tb00992.x

Eysenck, H. J. (1942). The experimental study of the "good gestalt": A new approach. *Psychological Review, 49,* 344–364. http://dx.doi.org/10.1037/h0057013

Eysenck, H. J. (1968). An experimental study of aesthetic preference for polygonal figures. *The Journal of General Psychology, 79,* 3–17. http://dx.doi.org/10.1080/00221309.1968.9710447

Eysenck, H. J. (1983). A new measure of "good taste" in visual art. *Leonardo, 16,* 229. http://dx.doi.org/10.2307/1574921

Eysenck, H. J. (1997). *Rebel with a cause: The autobiography of Hans Eysenck* (Revised Updated edition). New Brunswick, NJ: Transaction Publishers.

Eysenck, H. J., & Furnham, A. (1993). Personality and the Barron-Welsh art scale. *Perceptual and Motor Skills, 76,* 837–838. http://dx.doi.org/10.2466/pms.1993.76.3.837

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review, 2,* 290–309. http://dx.doi.org/10.1207/s15327957pspr0204_5

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17,* 172–179. http://dx.doi.org/10.1111/j.1467-9280.2006.01681.x

Frois, J. P., & Eysenck, H. J. (1995). The visual aesthetic sensitivity test applied to Portuguese children and fine arts students. *Creativity Research Journal, 8,* 277–284. http://dx.doi.org/10.1207/s15326934crj0803_6

Furnham, A., & Avison, M. (1997). Personality and preference for surreal paintings. *Personality and Individual Differences, 23,* 923–935. http://dx.doi.org/10.1016/S0191-8869(97)00131-1

Furnham, A., & Bunyan, M. (1988). Personality and art preferences. *European Journal of Personality, 2,* 67–74. http://dx.doi.org/10.1002/per.2410020106

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

Gear, J. (1986). Eysenck's Visual Aesthetic Sensitivity Test (VAST) as an example of the need for explicitness and awareness of context in empirical aesthetics. *Poetics, 15,* 555–564. http://dx.doi.org/10.1016/0304-422X(86)90011-2

Gignac, G. E., Stough, C., & Loukomitis, S. (2004). Openness, intelligence, and self-report intelligence. *Intelligence, 32,* 133–143. http://dx.doi.org/10.1016/j.intell.2003.10.005

Götz, K. O. (1985). *VAST: Visual Aesthetic Sensitivity Test* (4th ed.). Dusseldorf, Germany: Concept Verlag.

Götz, K. O. (1987). Visual aesthetic sensitivity and intelligence. *Perceptual and Motor Skills, 65,* 422. http://dx.doi.org/10.2466/pms.1987.65.2.422

Götz, K. O., Borisy, A. R., Lynn, R., & Eysenck, H. J. (1979). A new Visual Aesthetic Sensitivity Test: I. Construction and psychometric properties. *Perceptual and Motor Skills, 49,* 795–802. http://dx.doi.org/10.2466/pms.1979.49.3.795

Graves, M. E. (1948). *Design judgment test*. New York, NY: Psychological Corporation.

Graves, M. E. (1951). *The art of color and design* (1st ed.). New York, NY: McGraw-Hill Book Company, Inc.

Hartung, J. (1998). *An alternative method for meta-analysis*. Technical Report Number, SFB 475: Komplexitätsreduktion in Multivariaten Datenstrukturen, Universität Dortmund. Retrieved from http://www.econstor.eu/handle/10419/77295

Hartung, J., & Knapp, G. (2001a). A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Statistics in Medicine, 20,* 3875–3889. http://dx.doi.org/10.1002/sim.1009

Hartung, J., & Knapp, G. (2001b). On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Statistics in Medicine, 20,* 1771–1782. http://dx.doi.org/10.1002/sim.791

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539–1558. http://dx.doi.org/10.1002/sim.1186

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557–560. http://dx.doi.org/10.1136/bmj.327.7414.557

IntHout, J., Ioannidis, J. P., & Borm, G. F. (2014). The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology, 14,* 25. http://dx.doi.org/10.1186/1471-2288-14-25

Iwawaki, S., Eysenck, H. J., & Götz, K. O. (1979). A new Visual Aesthetic Sensitivity Test (VAST): II. cross-cultural comparison between England and Japan. *Perceptual and Motor Skills, 49,* 859–862. http://dx.doi.org/10.2466/pms.1979.49.3.859

Jacobsen, T. (2006). Bridging the arts and sciences: A framework for the psychology of aesthetics. *Leonardo, 39,* 155–162. http://dx.doi.org/10.1162/leon.2006.39.2.155

Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *The Journal of Creative Behavior, 42,* 106–130. http://dx.doi.org/10.1002/j.2162-6057.2008.tb01290.x

Koide, N., Kubo, T., Nishida, S., Shibata, T., & Ikeda, K. (2015). Art expertise reduces influence of visual salience on fixation in viewing abstract-paintings. *PLoS ONE, 10,* e0117696. http://dx.doi.org/10.1371/journal.pone.0117696

Kozbelt, A., & Seeley, W. P. (2007). Integrating art historical, psychological, and neuroscientific explanations of artists' advantages in drawing and perception. *Psychology of Aesthetics, Creativity, and the Arts, 1,* 80–90. http://dx.doi.org/10.1037/1931-3896.1.2.80

Kozbelt, A., Seidel, A., ElBassiouny, A., Mark, Y., & Owen, D. R. (2010). Visual selection contributes to artists' advantages in realistic drawing. *Psychology of Aesthetics, Creativity, and the Arts, 4,* 93–102. http://dx.doi.org/10.1037/a0017657

Leder, H., Gerger, G., Brieber, D., & Schwarz, N. (2014). What makes an art expert? Emotion and evaluation in art appreciation. *Cognition and Emotion, 28,* 1137–1147. http://dx.doi.org/10.1080/02699931.2013.870132

Lifton, W. M. (1961). The development of a music reaction test to measure affective and aesthetic sensitivity. *Journal of Research in Music Education, 9,* 157–166. http://dx.doi.org/10.2307/3344311

Liu, F.-J. (1990). Critique of three tests of aesthetic judgment: Maitland Graves Design Judgment Test; The Meier Art Tests: I, Art Judgment; and The Meier Art Tests: II, Aesthetic Perception. *Visual Arts Research, 16,* 90–99.

Locher, P. J., & Nodine, C. (1989). The perceptual value of symmetry. *Computers & Mathematics with Applications, 17,* 475–484. http://dx.doi.org/10.1016/0898-1221(89)90246-0

Marković, S., & Gnozdenovi, V. (2001). Symmetry, complexity and perceptual economy: Effects of minimum and maximum simplicity conditions. *Visual Cognition, 8,* 305–327. http://dx.doi.org/10.1080/13506280143000025

McManus, I. C. (2005). Symmetry and asymmetry in aesthetics and the arts. *European Review, 13,* 157–180. http://dx.doi.org/10.1017/S1062798705000736

Meier, N. C. (1928). A measure of art talent. *Psychological Monographs, 39,* 184–199. http://dx.doi.org/10.1037/h0093346

Meier, N. C. (1940). *The Meier Art Tests: I, Art Judgment*. Iowa City, IA: Bureau of Educational Research and Service, University of Iowa.

Meier, N. C. (1942). *The Meier Art Tests. I. Art judgment: Examiner's manual*. Iowa City, IA: Bureau of Educational Research and Service, State University of Iowa.

Meier, N. C. (1963). *The Meier Art Tests: II, Aesthetic Perception*. Iowa City, IA: Bureau of Educational Research and Service, University of Iowa.

Muth, C., Hesslinger, V. M., & Carbon, C.-C. (2015). The appeal of challenge in the perception of art: How ambiguity, solvability of ambiguity, and the opportunity for insight affect appreciation. *Psychology of Aesthetics, Creativity, and the Arts, 9,* 206–216. http://dx.doi.org/10.1037/a0038814

Myszkowski, N., Storme, M., & Zenasni, F. (2016). Order in complexity: How Hans Eysenck brought differential psychology and aesthetics together. *Personality and Individual Differences, 103,* 156–162. http://dx.doi.org/10.1016/j.paid.2016.04.034

Myszkowski, N., Storme, M., Zenasni, F., & Lubart, T. (2014). Is visual aesthetic sensitivity independent from intelligence, personality and creativity? *Personality and Individual Differences, 59,* 16–20. http://dx.doi.org/10.1016/j.paid.2013.10.021

Myszkowski, N., & Zenasni, F. (2016). Individual differences in aesthetic ability: The case for an Aesthetic Quotient. *Frontiers in Psychology, 7,* 750. http://dx.doi.org/10.3389/fpsyg.2016.00750

Nodine, C. F., Locher, P. J., & Krupinski, E. A. (1993). The role of formal art training on perception and aesthetic judgment of art compositions. *Leonardo, 26,* 219. http://dx.doi.org/10.2307/1575815

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61,* 991–996. http://dx.doi.org/10.1016/j.jclinepi.2007.11.010

Plucker, J. A., Esping, A., Kaufman, J. C., & Avitia, M. J. (2015). Creativity and intelligence. In S. Goldstein & J. Naglieri (Eds.), *Handbook of intelligence* (pp. 283–291). New York, NY: Springer. Retrieved from http://link.springer.com/chap./10.1007/978-1-4939-1562-0_19

Plucker, J. A., Kaufman, J. C., Temple, J. S., & Qian, M. (2009). Do experts and novices evaluate movies the same way? *Psychology & Marketing, 26,* 470–478. http://dx.doi.org/10.1002/mar.20283

Raven, J. C. (1941). Standardization of progressive matrices, 1938. *British Journal of Medical Psychology, 19,* 137–150. http://dx.doi.org/10.1111/j.2044-8341.1941.tb00316.x

Rawlings, D., Twomey, F., Burns, E., & Morris, S. (1998). Personality, creativity, and aesthetic preference: Comparing psychoticism, sensation seeking, schizotypy, and openness to experience. *Empirical Studies of the Arts, 16,* 153–178. http://dx.doi.org/10.2190/8GVJ-ERL3-HYUM-EH88

Reber, R. (2002). Reasons for the preference for symmetry. *Behavioral and Brain Sciences, 25,* 415–416. http://dx.doi.org/10.1017/S0140525X02350076

Schwarzer, G. (2007). Meta: An R package for meta-analysis. *R News, 7,* 40–5. Retrieve from https://cran.r-project.org/doc/Rnews/Rnews_2007-3.pdf

Schwarzer, G., & Schwarzer, M. G. (2015). *Package "meta."* New York, NY: Springer.

Sidik, K., & Jonkman, J. N. (2005). Simple heterogeneity variance estimation for meta-analysis. *Journal of the Royal Statistical Society, Series C: Applied Statistics, 54,* 367–384. http://dx.doi.org/10.1111/j.1467-9876.2005.00489.x

Silvia, P. J. (2007). Knowledge-based assessment of expertise in the arts: Exploring aesthetic fluency. *Psychology of Aesthetics, Creativity, and the Arts, 1,* 247–249. http://dx.doi.org/10.1037/1931-3896.1.4.247

Silvia, P. J., & Barona, C. M. (2009). Do people prefer curved objects? Angularity, expertise, and aesthetic preference. *Empirical Studies of the Arts, 27,* 25–42. http://dx.doi.org/10.2190/EM.27.1.b

Silvia, P. J., & Nusbaum, E. C. (2011). On personality and piloerection: Individual differences in aesthetic chills and other unusual aesthetic experiences. *Psychology of Aesthetics, Creativity, and the Arts, 5,* 208–214. http://dx.doi.org/10.1037/a0021914

Smith, L. F., & Smith, J. K. (2006). The nature and growth of aesthetic fluency. In P. Locher, C. Martindale, & L. Dorfman (Eds.), *New directions in aesthetics, creativity and the arts* (pp. 47–58). Amityville, NY: Baywood Publishing Co.

Spearman, C. (1939). Thurstone's work re-worked. *Journal of Educational Psychology, 30,* 1–16. http://dx.doi.org/10.1037/h0061267

Sterne, J. A. C., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology, 54,* 1046–1055. http://dx.doi.org/10.1016/S0895-4356(01)00377-8

Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 99–110). Hoboken, NJ: Wiley.

Summerfeldt, L. J., Gilbert, S. J., & Reynolds, M. (2015). Incompleteness, aesthetic sensitivity, and the obsessive-compulsive need for symmetry. *Journal of Behavior Therapy and Experimental Psychiatry, 49,* 141–149. http://doi.org/10.1016/j.jbtep.2015.03.006

Thurstone, L. L. (1938). *Primary mental abilities. Psychometric Monograph*. Chicago, IL: University of Chicago Press.

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. http://dx.doi.org/10.1037/a0030872

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36,* 1–48.

Viechtbauer, W., & Viechtbauer, M. W. (2015). *Package "metafor."* The comprehensive R archive network. Retrieved from http://cran.r-project.org/web/packages/metafor/metafor.pdf

Vitz, P. C. (1966). Preference for different amounts of visual complexity. *Behavioral Science, 11,* 105–114. http://dx.doi.org/10.1002/bs.3830110204

White, H. D. J. (1931). An application of mental tests to university students. *British Journal of Educational Psychology, 1,* 279–295. http://dx.doi.org/10.1111/j.2044-8279.1931.tb02673.x

Winston, A. S., & Cupchik, G. C. (1992). The evaluation of high art and popular art by naive and experienced viewers. *Visual Arts Research, 18,* 1–14.

Wonderlic, E. (1992). *Manual of the wonderlic personnel test and scholastic level exam II*. Libertyville, IL: Wonderlic Personnel Test.

Received April 4, 2016
Revision received August 24, 2016
Accepted October 4, 2016 ■