literature

MDPI

Article

Heinrich von Kleist's Extremely Complex Syntax: How Does It Affect Aesthetic Liking?

Winfried Menninghaus [1,*], Vanessa Kegel [2], Kirill Fayn [2] and Wolff Schlotz [3,†]

[1] Max Planck Institute for Empirical Aesthetics, 60322 Frankfurt, Germany
[2] Independent Researcher, Sydney 2006, Australia
[3] WMR Prenatal Adversity & Brain Development, University of the Southampton, Southampton SO17 1BJ, UK
* Correspondence: w.m@ae.mpg.de
† In memoriam: Wolff Schlotz, psychologist, methods expert, and one of the founding members of the Max Planck Institute for Empirical Aesthetics, made from the very beginning decisive contributions to the research design, the methodology and the data analysis of the present study. Sadly, he did not live to see it completed. We are very grateful for what he did to make this study fly.

Abstract

Ease of cognitive processing is an important predictor of aesthetic liking. However, many acclaimed artworks are fairly complex and require substantial cognitive effort. Are they aesthetically liked despite or because of this increased cognitive challenge? The present study pursued this question experimentally. The high syntactic complexity of Heinrich von Kleist's narratives provided the test case. According to literary scholars, this high syntactic complexity should support increased levels of how "suspenseful," "intense," "interesting," and evocative of a sense of "urgency" the texts are perceived, and it should thereby also support higher overall aesthetic liking. This expectation is in line with recent models in empirical aesthetics according to which higher ease of processing and higher cognitive challenge are not mutually exclusive, but can conjointly drive aesthetic liking to higher levels. The standard hypothesis of cognitive fluency instead predicts a disfluency-driven negative effect on aesthetic liking. We tested these two predictions in two studies by presenting excerpts from Kleist's narratives in their original vs. syntactically simplified versions to participants. Results differ substantially depending on how the target variables are statistically modeled. If ease of processing and cognitive challenge are modeled separately as predictors of the aesthetically evaluative ratings, higher ease of processing is a strong positive and higher cognitive challenge a largely negative predictor. However, when the two complementary cognitive variables are modeled conjointly, they are both positive predictors of the aesthetically evaluative ratings. Their predictive power differs, however, significantly. Only the positive effect of ease of processing is pervasive across all readers. That of cognitive challenge is substantially modified by individual differences. Specifically, it was observed for readers who (1) are of higher age, (2) like to read narratives in general, and (3) reported prior positive experiences with Kleist. Supporting the ecological validity of our findings, readers meeting these criteria are more likely than others to actually read Kleist outside the laboratory.

Keywords: syntactic complexity; cognitive challenge; processing fluency; suspense; intensity; urgency; liking; individual differences

check for updates

Academic Editor: Gillian Dooley

Received: 14 May 2025
Revised: 10 June 2025
Accepted: 17 June 2025
Published: 30 September 2025

Citation: Menninghaus, Winfried, Vanessa Kegel, Kirill Fayn, and Wolff Schlotz. 2025. Heinrich von Kleist's Extremely Complex Syntax: How Does It Affect Aesthetic Liking? *Literature* 5: 25. https://doi.org/10.3390/literature5040025

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The standard hypothesis of cognitive fluency/disfluency in aesthetic processing stipulates that fluency-enhancing features are aesthetically preferred over disfluent ones (Reber et al. 2004; Winkielman and Cacioppo 2001). There is substantial empirical evidence in support of this hypothesis (Anglada-Tort et al. 2019; Belke et al. 2010; Forster et al. 2013; Obermeier et al. 2016). At the same time, several models of aesthetic processing include provisions for the compatibility, at least within certain limits, of higher disfluency on one dimension of stimulus processing with higher fluency on other dimensions, resulting in a potentially favorable balance of the two antagonistic ingredients (Giora et al. 2004; Jacobs 2014; Graf and Landwehr 2015; Menninghaus et al. 2015; Miall and Kuiken 1994). For example, the implementation of rhyme and meter is often achieved at the expense of deviant word forms and word order that render the processing of word morphology and syntax more demanding (Wallot and Menninghaus 2018; Menninghaus et al. 2023). Still, the co-presence of higher disorder in poetic language drives aesthetically evaluative ratings, specifically for "beauty" and "melodiousness," to higher levels (Menninghaus et al. 2017). High ratings on scales for "succinctness" (Menninghaus et al. 2015), "humorousness" (Menninghaus et al. 2014), and "interestingness" (Fayn et al. 2015b; Knoop et al. 2016; Silvia 2005) have likewise been shown to be compatible with increased cognitive demand.

These findings underlie what we henceforth call the revised as opposed to the standard hypothesis of cognitive dis/fluency in aesthetic processing. The study reported here tested these hypotheses directly against one another. Six excerpts from Heinrich von Kleist's narratives (1777–1811) were the test case. Kleist ranks among the famous authors of German literature. His narratives and dramas—which until recently were nearly mandatory readings in German school curricula—tend to focus on highly violent events, including social and religious conflicts and uproar, extreme gender relations, duels, and large-scale natural disasters. Kleist's suicide at age 34, together with an agreed-upon shooting of his beloved, contributed to his fame as a rebellious (anti-)classic. Importantly for our study, Kleist was also decidedly extreme in his handling of German syntax. The stylistic features of interest were the following:

(1) The sentences feature complex cascades of embeddings (subordinate clauses), including many embeddings within embeddings, which in linguistics are technically called "center-embeddings."

(2) As a consequence of (1), Kleist's sentences also feature an excessive number of punctuation marks, and

(3) their absolute length is very high.

All these stylistic features (for more details and references, see the Method section of Study 1) tend to render syntactic processing cognitively more demanding. Crucially for our study, these features can easily be statistically quantified, and they can also be readily manipulated to varying degrees, while keeping both the words used and the total word count identical. This enabled us to create less syntactically complex and potentially also less cognitively demanding versions of the texts and to contrast these with the original versions. Furthermore, we developed three statistical measures of the syntactic complexity under consideration. This enabled us to treat the text-specific levels of syntactic complexity as a continuous predictor of the subjective ratings of the different text versions.

2. Hypotheses

To begin with, we expected that our participants would on average rate the original text versions as more complex, more cognitively demanding, and deviating more from ordinary language use than the modified versions, and consequently would rate them as higher in overall cognitive challenge (=Hypothesis 1).

Higher cognitive challenge is likely to reduce ease of processing. Accordingly, we expected that Kleist's original versions of the texts should be less easy to process than our self-produced experimental versions of the texts (=Hypothesis 2).

Regarding the expected effects of Kleist's extraordinary syntax on aesthetic valuation, we tested seven hypotheses that were adopted from Kleist scholarship or developed by the first author of this study. These hypotheses invariably stipulate that Kleist's special style, for all its syntactic complexity, should boost rather than reduce the aesthetic enjoyment of his texts. In contrast, the standard hypothesis of cognitive fluency stipulates that the high syntactic complexity of Kleist's original texts should have a negative effect on aesthetic evaluations and overall liking.

As the syntactic trajectories of Kleist's sentences go through extended periods of unresolved tensions, they are likely, by virtue of their linguistic tension-resolution structure, to enhance subjective impressions of narrative suspense (=Hypothesis 3; for general accounts of tension-resolution trajectories in language—which all rely on our capacities for predictive coding—see Frazier and Fodor 1978; Pickering and van Gompel 2006; Traxler et al. 2018 for the musical analogue, see Huron 2006; Meyer 1961; for stylistic analyses in literary scholarship on Kleist that support these assumptions, see Fries 1904; Sembdner 1994; Stenzel 1966).

Kleist is, moreover, an author with a pronounced focus on states of emergency, be they natural catastrophes such as an earthquake, war, social upheaval, violent political and religious conflicts, or extreme individual predicaments such as coping with the consequences of having been raped while in a state of unconsciousness (the case of the *Marquise of O*...). States of emergency call for a high sense of alertness and an urgency in survival-enhancing responses, and the intensity of experiencing such exceptional situations is likely to be very high. In Kleist's verbal execution of his narrative plots, a style-driven increase of feelings of urgency and intensity (cf. Carrière 1984; Gumbrecht and Knüpling 2004) may add to the plot-driven sense of urgency and intensity. Accordingly, we predicted that Kleist's original narratives would also be rated higher in subjectively perceived urgency (=Hypothesis 4) and intensity (=Hypothesis 5) than the modified variants. As a result of this resonance of content and style, we expected that Kleist's original sentences would likewise be perceived as higher in representational vividness (German *Lebendigkeit*; for the meaning of this concept in poetics and aesthetics, see Menninghaus 2009; Belfi et al. 2018) compared to their syntactically modified variants (=Hypothesis 6).

Classical poetics distinguishes two antithetical types of sentence rhythm: smooth and austere (Dionysius of Halicarnassus 1985, p. 169; see also von Hellingrath 1911). The first derives its aesthetic power from a pleasant ease of fluent processing. The second—which in extreme cases pushes sentences almost to the point of becoming failed sentences—thrives on the rewards of resistance overcome; it is believed to constitute a rhythm of its own kind, a more austere one, but still a pronounced rhythm. In our understanding, Kleist's original sentences can be assigned to the category of austere rhythm, as readers are likely to integrate the consistent detours from a smooth sentence flow into their expectations (predictions) and may end up discovering distinct rhythms of their own in Kleist's tense syntax. Expecting that our experimental simplification of Kleist's sentences would destroy this particular type of rhythm without necessarily implementing a distinctly rewarding smooth rhythm, we suspected that the original sentences, for all their complexity, should still be rated higher in perceived rhythm than the modified versions (=Hypothesis 7).

We also tested the assumption that Kleist's particular prose style might intensify well-established general dimensions of readerly involvement in narratives (Cohen 2001; Lee 2004; Zillmann 1996), specifically, attentional focus, feelings of being in another world (immersion, transportation), and empathy with the protagonists (=Hypothesis 8).

Moreover, we took recourse to a modern category of aesthetic appeal that became topical only in Kleist's time, namely, the "interesting" (Schlegel 2001; for recent research, see Fayn et al. 2019; Silvia 2010). We tested the assumption that Kleist's original narratives might be perceived as more interesting than their stylistically less ambitious variants (=Hypothesis 9).

Finally, if Hypotheses 3 through 9—or at least some of these hypotheses—hold true for the respective individual dimensions of aesthetic reward, then overall self-reported liking should also be higher for Kleist's original sentences compared to our self-produced modified versions (=Hypothesis 10).

Because Hypotheses 3 through 9 all bear on dimensions of aesthetic evaluation, we henceforth refer to these in a summarizing fashion as aesthetic evaluations. While we did expect positive confirmations of Hypotheses 1 and 2, Hypotheses 3 through 9 were all tested in an exploratory fashion.

Notably, there is already ample evidence that increased cognitive disfluency may contribute to a more profound semantic understanding of messages (see the review of the pertinent research in Alter 2013) and may hence be associated with positive effects on processing and memory. In this line of research, however, inherent aesthetic processing rewards are not the key issue. Against this background, the present study extends the existing research on cognitive benefits of cognitive disfluency in language processing to the potential enhancing effects of cognitive disfluency on overall aesthetic liking as well as on specific dimensions of aesthetic evaluation.

3. Study 1

3.1. Method

3.1.1. Participants

In total, 201 participants completed this study. One participant indicated that he was not a German native speaker; his responses were excluded from the analyses. The final sample thus comprised N = 200 participants (75% female, age: M = 37.24 years, SD = 18.08, range: 18–78 years). Among these, one participant did not respond to the age and gender questions, and another participant did not respond to the question regarding reading frequency.

Participants were recruited via announcements under the heading "Do You Like Reading?—The Art of Story Telling" via the research participant database of the Max Planck Institute for Empirical Aesthetics as well as via flyers and posters distributed at art centers and bookstores. The inclusion criteria were a minimum age of 18 years and German as a native language (176 were German monolingual, and 24 bilingual with German and another language). Participants were predominantly highly educated (92% had earned the general entrance qualification for studying at a university). On average, the participants reported that they read novels and/or short narratives several times per month (M = 3.37, SD = 1.75, response scale: 0 = *never*, 1 = *once per year*, 2 = *several times per year*, 3 = *several times per month*, 4 = *once per week*, 5 = *several times per week*, 6 = (*almost) daily*) and that they liked reading novels and/or short narratives (M = 2.05, SD = 1.37, response scale: −3 = *dislike very much*, −2 = *dislike*, −1 = *rather dislike*, 0 = *neutral*, +1 = *rather like*, +2 = *like*, +3 = *like very much*).

3.1.2. Stimuli

The selection of excerpts from Kleist's narratives was driven by three criteria. First, the excerpts should present sequences of events that can be read as self-contained episodes. Second, they should, for the purposes of our experimental study, not exceed an average total length of 700 words. Third, both the original and the modified versions should vary

with respect to quantitative descriptors of the stylistic features under scrutiny, i.e., sentence length, syntactic complexity, and prosodic fluency. Guided by these criteria, we presented all participants of Study 1 with excerpts from Kleist's narratives *Die heilige Cäcilie oder die Gewalt der Musik* (*St. Cecilia or the Power of Music*; von Kleist 1997), *Das Erdbeben in Chili* (*The Earthquake in Chile*; von Kleist 1993), *Michael Kohlhaas* (Michael Kohlhaas; von Kleist 1990), *Die Marquise von O. . .* (*The Marquise of O. . .*; von Kleist 1989), and *Der Zweikampf* (*The Duel*; two excerpts; von Kleist 1994),which were adopted from the critical edition of Kleist's works. The average text length was 680 words (ranging from 620–773 words). In three cases, the excerpts were taken from the very beginning of the narratives; the other three excerpts were preceded by a short summary of what had happened prior to the beginning of the excerpt. The selection of the texts was not guided by preferences for some of Kleist's texts over others, but by the need to identify text passages that we expected to be well-readable as self-contained episodes.

In both the original and the modified versions, early 19th-century orthography as found in the critical edition of the narratives was adjusted to modern orthography. In a few cases, antiquated vocabulary was replaced in order to avoid reading difficulties that result exclusively from linguistic forms that have become historically obsolete.

Regarding the modified versions, great care was taken to leave not only the content, but also the wording, as identical as possible (for details, see the section "Text Modification").

Textual Target Features and the Quantitative Measures That Capture Them. We studied, by means of experimental modifications, the cognitive and aesthetic effects of the following stylistic characteristics of Kleist's sentences:

(1) The single most distinctive stylistic feature is the *very high number of subordinate clauses* (cf. Bischoff 1899), or *embeddings*, including many infinitive clauses and participle clauses, per main clause (cf. Admoni 1987, p. 84). Importantly, for all their syntactic complexity, there is nothing syntactically ambiguous, let alone incorrect, about Kleist's extremely complex sentences (cf. Admoni 1987, p. 83). Specifically, Kleist made frequent, and quite extraordinary, use of subordinate clauses that are *center-embedded* and/or nested in other subordinate clauses. Center-embedded clauses—and even more so center-embedded clauses nested in other center-embedded clauses—are known to render syntactic parsing particularly demanding (for linguistic research on the processing of such clauses see Bach et al. 1986; Chomsky 1965; Gibson 1998; Karlsson 2007; Larkin and Burns 1977; Makuuchi et al. 2009; Yngve 1960). We quantified this stylistic feature by computing for each text the average number of embeddings per main clause. These embeddings comprise all subordinate clauses of a sentence, regardless of whether they were center-embedded, initial-embedded, or final-embedded.

(2) In German writing, many embeddings are mandatorily marked by commas. Multiple syntactic embeddings on different levels therefore tend to result in a high number of commas. In the example given in Figure 1, immediately after the initial word (in this case, the name of the protagonist), the sentence is sent on a lengthy detour through cascades of subordinate clauses before the noun phrase is completed by the verb phrase ("Kohlhaas [. . .] asked"). The *mandatory commas* following one upon another in this example at very short intervals are held to prime a reduced sentence flow (cf. Fries 1904). Kleist's *optional commas* further increase this tendency by strengthening otherwise minor or purely facultative prosodic pauses, most notably between noun or verb phrases, on the one hand, and adjectival or adverbial phrases, on the other (cf. Klausnitzer 2012; Nebrig and Spoerhase 2012; Sembdner 1994; Stenzel 1966). To be sure, Kleist's optional punctuation marks make the syntactic roles of the sentence constituents they delimit more explicit; in this capacity they are helpful for correctly decoding the grammatical dependencies in very long sentences. Still, the unusually high number of punctuation marks requires the processing

of additional items of non-phonetic visual input. In this capacity, they may well have a negative effect on the subjectively perceived fluency of processing.

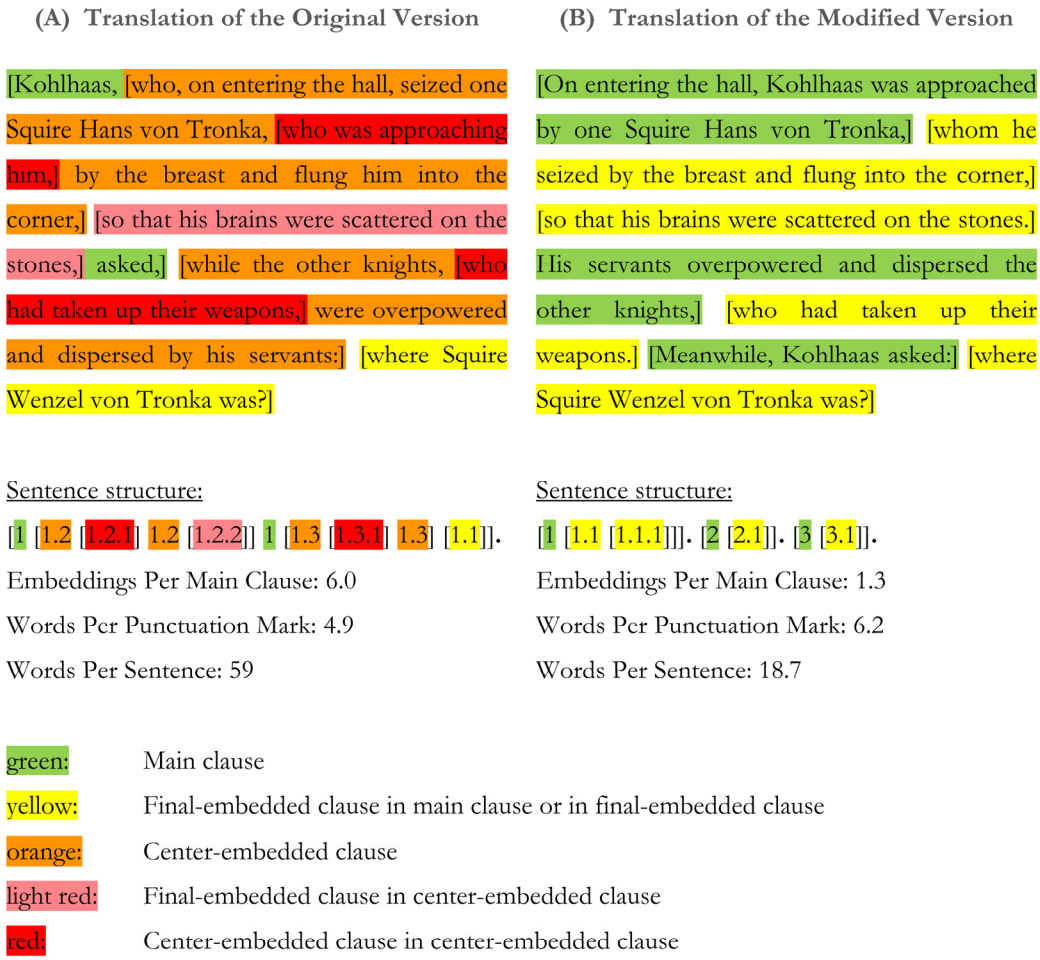

Figure 1. An exemplary sentence from Kleist's novella *Michael Kohlhaas* in its original and modified versions. Note. In the original version (**A**), the noun phrase ("Kohlhaas") and the verb phrase ("asked") of the main clause (=1), as well as the actual question asked ("where Squire Wenzel von Tronka was?"), are set apart by multiple center and final embeddings. In the modified version (**B**), the propositional information has been split up into three main clauses that are complemented only by final-embedded clauses.

We operationalized this stylistic feature by computing the number of words per punctuation mark, without distinguishing between mandatory and optional punctuation marks.

A typical Kleistian sentence is *higher* than usual in the number of embeddings per main clause and—largely for this very reason—*lower* in the average number of words between punctuation marks compared to stylistically less deviant sentences.

(3) Kleist's original sentences are also very long as measured by the number of words per sentence. The unit of the sentence—which comprises both the main clauses and all their embeddings—was defined as extending between the periods or question/exclamation marks preceding and concluding it, regardless of how many syntactic phrases this unit includes. (Text-initial sentences obviously have no preceding period mark). The stylistic feature of high sentence length is largely a function of the massive insertions of embeddings in both main and subordinate clauses. It is therefore likely to be less distinctive of Kleist's style than is the extraordinarily high number of embeddings per main clause.

Text Modifications. In our text modifications, we took great care to leave all words of the original texts (apart from conjunctions and other function words) and the overall seman-

tic meaning unchanged. It turned out that Kleist's extraordinary sentences can very easily be "normalized" by reorganizing the word order alone, i.e., without replacing, omitting, or adding words. The three stylistic target features were modified in the following ways:

(1) The extremely high number of embeddings per main clause—and specifically of center embeddings—was experimentally reduced by transforming several of the subordinate clauses into main clauses while largely avoiding center embeddings.

(2) The changes that reduced the high number of embeddings per main clause simultaneously increased the very low number of words per punctuation mark. Removing most of the instances of optional punctuation further increased the number of words per punctuation mark.

(3) The transformation of subordinate clauses into main clauses as described under (1) also led to strongly reduced absolute sentence lengths (as measured by words per sentence).

The changes in words per punctuation mark and words per sentence are largely an invariable consequence of experimentally modifying the embeddings per main clause. We therefore expected that the most distinctive stylistic feature of Kleist's sentences—i.e., the extremely frequent embeddings per main clause—should also be the most powerful predictor of the effects of Kleist's complex syntax on cognitive processing and aesthetic evaluation.

Figure 1 presents an example of how we modified the stylistic target features of the original texts. It is based on an excerpt from the narrative Michael Kohlhaas (von Kleist 1990). The translation of the original German version is based on John Oxenford's translation (von Kleist 2019); it is mainly used for illustrative purposes (cf. Appendix A Figure A1 for this example in its original German language version as used in our studies). The translation of the modified version is ours.

Comparison of the Text Versions Based on our Objective Measures of Textual Complexity. Our three text measures show substantial variation not only between the original and the modified text versions, but also between the original texts. Specifically, two of these (*The Earthquake in Chile* and *The Marquise of O*. . .) show far less extreme values for both embeddings per main clause and words per sentence compared to the other excerpts.

3.1.3. Design

We implemented a 2 × 6 mixed factorial design with *Version* (original or modified version—between-person) and *Text* (six excerpts from five narratives—within-person) fully crossed. To control for potential order effects, we used a Latin square design. The Latin square was based on five texts instead of six, as two of the excerpts that came from the same narrative (*Der Zweikampf*) were not presented individually but always in a fixed order. This pair was assigned the position of the fifth text in the Latin square-based text sequence.

Participants were randomly assigned to two conditions, with 100 participants reading the original texts and 100 reading the modified texts. We initially decided on the number of participants based on a power analysis. The analysis suggested a total sample size of $N = 120$ to detect a small- to medium-sized effect ($f = 0.15$, which was determined based on a pilot study) of a between-subjects factor using a repeated measure analysis of variance comprising six measurements, with input parameters $\alpha = 0.05$ and power $1 - \beta = 0.80$. We decided to increase the sample size to 200 to achieve a higher power for detecting an effect of this size and to reduce a potential sampling bias within each individual text sequence.

Each of the 10 Latin-square balanced sequences was read by 10 participants per condition. However, due to the erroneous assignment of one participant and the retroactive exclusion of a non-native speaker we ended up with a slight imbalance for one sequence ($n = 11$ vs. $n = 9$ for another sequence) in the modified text condition. There were no

significant differences between the two groups in terms of age, gender, levels of education, or the degree to which they reported liking reading novels (all $p > 0.05$).

3.1.4. Procedure

The experimental sessions were conducted in groups of up to eight participants and took about one hour. Each participant was seated at an individually comfortable distance in front of a computer. The experiment was programmed in Presentation (Neurobehavioral Systems Inc., Albany, CA, USA).

Before the start of the experiment, participants read an information sheet about the general background and procedure of this study, gave their written informed consent, and filled out a short demographic questionnaire on paper. They were then randomly assigned to one of the two experimental conditions (original or modified versions of the texts). At the beginning of the experiment, participants completed a practice trial in order to familiarize themselves with the 27 questions to be answered after each text. Subsequently, the presentation of the first of the six texts started. Each text was presented in its entirety on one screen in a book page layout, and the presentation was finished by each participant individually clicking on an on-screen button. Participants were not informed of the author of the texts.

After reading each text, participants were asked whether they had read the narrative before and/or had an idea who the author was. Next, participants completed a computerized questionnaire that presented 27 items regarding the cognitive, affective, and aesthetic aspects of their individual perception of the text. This procedure was repeated for each of the six texts. At the end of the experimental sessions, participants answered a questionnaire regarding their reading behavior and preferences. Finally, participants were thanked and received the compensation for their participation.

3.1.5. Measures

Ratings. The participants' perception and evaluation of the presented narratives were assessed using nine measures, with a total of 27 items (see Appendix A Table A1 for the wording of all items). All items were presented in randomized order with a response scale ranging from 0 (*not at all*) to 6 (*extremely*). Cognitive challenge was assessed using four items, whereas liking, suspense, urgency, intensity, vividness, rhythm, and readerly involvement were assessed using three items each. Interest and ease of processing were assessed using one item each. McDonald's omega (ω) was used as a metric of reliability using multilevel confirmatory factor analysis (Geldhof et al. 2014). All multi-item scales had good to excellent reliability at both the within- and between-subject levels: cognitive challenge ($\omega_{within} = 0.73$, $\omega_{between} = 0.83$), liking ($\omega_{within} = 0.92$, $\omega_{between} = 0.96$), suspense ($\omega_{within} = 0.90$, $\omega_{between} = 0.96$), urgency ($\omega_{within} = 0.77$, $\omega_{between} = 0.96$), intensity ($\omega_{within} = 0.88$, $\omega_{between} = 0.92$), vividness ($\omega_{within} = 0.75$, $\omega_{between} = 0.88$), rhythm ($\omega_{within} = 0.75$, $\omega_{between} = 0.91$), and readerly involvement ($\omega_{within} = 0.84$, $\omega_{between} = 0.87$).

Text familiarity. We assessed whether a participant was familiar with a text or knew its author by asking "Did you know this text prior to reading it in the present study?" (*yes*/*no*) and "Do you have an idea of who the author of this text might be?" (*yes*/*no*), followed by a free text field in which the name of the presumed author could be entered. This input allowed to check for the correctness of author recognition.

Reading behavior and motivation. After reading all six texts and giving ratings for the predefined items, participants filled in a questionnaire regarding their reading behavior, specifically, frequency and liking of reading novels/short narratives, poems, dramas, textbooks, and journalistic texts. Finally, they answered some items regarding their motivations for reading, which were not analyzed for the current study.

3.1.6. Statistical Analysis

We tested our hypotheses using mixed-effects regression models with crossed random effects for participants and texts and tested for the main effects of text modification (factor Version) using Wald-type tests. We tested 10 separate models using mean scores of items for the factors cognitive challenge, ease of processing, liking, interest, suspense, urgency, intensity, vividness, rhythm, and readerly involvement as outcome variables. We tested text order effects by adding the text position and its interaction with the factor Version to the model. As we used a Latin square design for balancing text position effects, any order effects can be attributed to repeated reading independent of specific text sequences.

We performed sensitivity analyses to test whether the results were dependent on potentially confounding variables. First, we included gender and frequency of reading novels/short narratives. Second, we analyzed reading times and checked whether the results were affected by excluding 10 outlying observations, defined as a participant's reading time for a specific text deviating more than 3 standard deviations from the average reading time across all texts and participants. As these sensitivity analyses did not change the findings, we do not report their results here.

In addition, we performed two exploratory analyses. First, we tested Text by Version interactions to investigate whether any differences between original and modified versions were text-specific. Second, we explored the effects of the quantitative text characteristics identified above. The mixed-effects regression model included crossed random factors for participants and texts, along with separate within-version-centered predictors of the three indicators embeddings per main clause, words per punctuation mark, and words per sentence for the original and modified text versions. The centered predictors were constructed by standardizing the variables across all six original and all six modified text versions. The parameter associated with a standardized predictor variable reflects the average effects of text-specific deviations from the version mean on the reader.

3.2. Results

3.2.1. Reading Times, Text Familiarity, and Author Recognition

Altogether, we collected 1200 text readings (200 participants, 6 texts each). On average, it took participants 3 min 24 s to read one text (M = 204.50 s, SD = 73.10 s, Mdn = 201 s, IQR = 80 s). Reading times for the original texts were not significantly different from those for the modified versions (Mdiff = 6.78 s, SE = 8.60, $p = 0.43$). At the same time, they show a trend towards shorter reading times for texts at a later presentation position (coefficient = −6.77, SE = 0.72, $p < 0.001$).

The participants indicated that they had not known most of the texts before they read them in this study (N = 1023, 85%). At 15% for the original versions and 14% for the modified versions, self-indicated familiarity with the texts was not significantly different between the versions ($\chi 2 = 0.33$, $p = 0.60$).

Correct author attributions for the presented texts were likewise in the range of 15% and not significantly affected by our experimental modification (original version = 16%, modified version = 14%, $\chi 2 = 1.99$, $p = 0.37$).

3.2.2. Mean Version-Dependent Differences in Ratings Across All Texts

As predicted in Hypothesis 1, the original excerpts were on average rated higher than their modified variants for the factor cognitive challenge, b = 0.23, SE = 0.10, $p = 0.02$. However, considering that the differences between text versions as quantified by our three statistical measures embeddings per main clause, words per punctuation mark, and words per sentence are fairly large, the mean increases in self-reported cognitive challenge are quite small (0.23 points on a scale of 0–6), and the effect on self-reported ease of processing

was not significant, b = 0.03, SE = 0.11, $p = 0.80$. Moreover, the manipulation did not affect overall liking, b = 0.09, SE = 0.15, $p = 0.53$, nor did it affect the ratings for suspense, b = 0.07, SE = 0.11, $p = 0.54$, urgency, b = 0.10, SE = 0.11, $p = 0.37$, intensity, b = 0.06, SE = 0.12, $p = 0.59$, vividness, b = 0.16, SE = 0.98, $p = 0.10$, rhythm, b = 0.04, SE = 0.11, $p = 0.73$, readerly involvement, b = 0.09, SE = 0.13, $p = 0.49$, or interest, b = 0.09, SE = 0.15, $p = 0.53$.

Considering that the higher self-reported cognitive challenge had no adverse effect on ease of processing and the aesthetic evaluations, the results do not support the standard hypothesis of cognitive fluency. At the same time, compared to the modified versions, Kleist's syntactically complex original texts received on average no significantly higher ratings on the scales for the aesthetic evaluation factors. Thus, the results are inconclusive regarding the two hypotheses. They provide neither full support for the standard hypothesis nor for the revised hypothesis of cognitive fluency.

3.2.3. Version-Dependent Differences in Ratings for the Individual Texts

In the next step, we separately analyzed the results for all individual texts. In the following Figure 2, positive values indicate higher scores for the original compared to the modified versions.

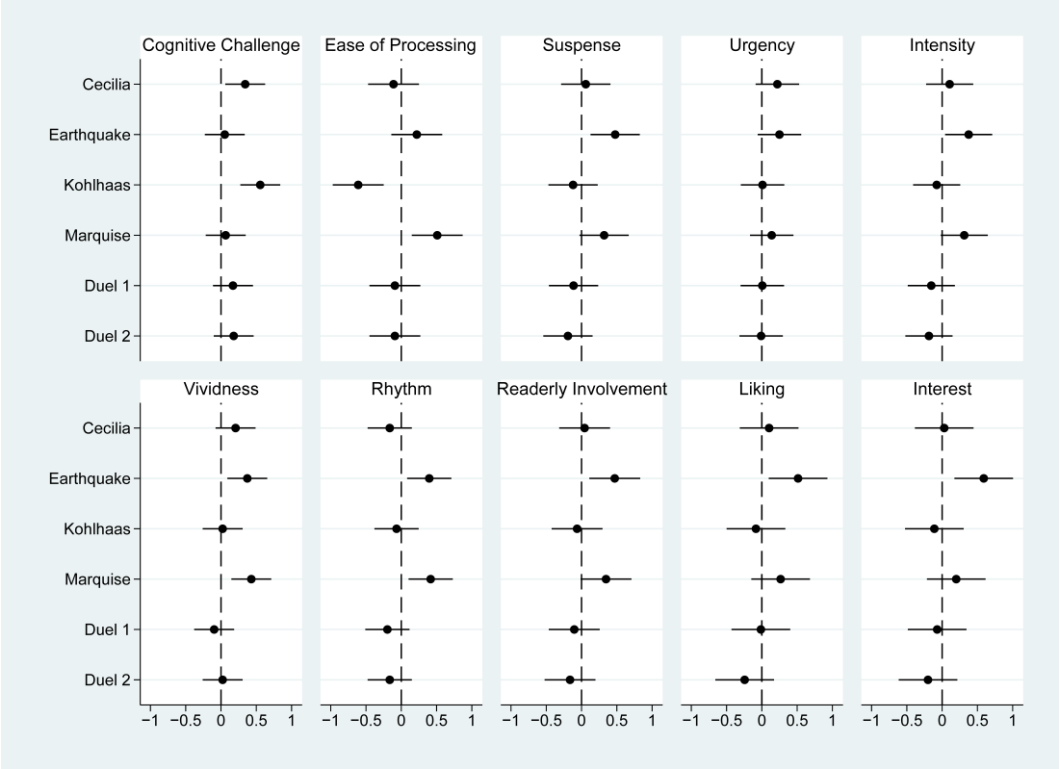

Figure 2. Mixed-model-derived text-specific differences between the modified and original versions of the texts. Note: Whiskers indicate 95% confidence intervals.

The results show that the slightly higher mean ratings for cognitive challenge for the original compared to the modified versions in the overall analysis are largely driven by the significant version effects obtained for the excerpt from *Michael Kohlhaas* and, to a lesser degree, that from *Cecilia.* Notably, even regarding the excerpt from *Michael Kohlhaas*, higher ratings for cognitive challenge and lower ratings for ease of processing were not associated with significant differences in the aesthetic evaluations.

For the four other excerpts from Kleist's narratives, the levels of cognitive challenge reported by our study participants did not significantly differ between the versions. Regarding ease of processing, a significant adverse effect of the original version was obtained

only for the excerpt from *Michael Kohlhaas*. The original excerpt from *The Marquise of O...* was actually rated as being higher in ease of processing than the less complex modified version, and for the other four texts the results show no version differences.

Moreover, two of the original texts (*The Earthquake in Chile* and *The Marquise of O...*) were rated more positively on several of the aesthetic evaluation scales as compared to their modified counterparts. Specifically, the original version of *The Earthquake in Chile* was rated significantly higher for liking, interest, suspense, intensity, vividness, rhythm, and readerly involvement, and the original version of *The Marquise of O...* was rated significantly higher for vividness and readerly involvement, with no adverse effects for the other aesthetic ratings.

Notably, *The Marquise of O...* is the least syntactically complex original text as indicated by our statistical measures of syntactic complexity (see Table 1), and *The Earthquake in Chile* is second-lowest on these measures. Thus, it was only for those texts in which the objectively measured syntactic complexity exceeded moderate levels that the Kleistian syntax turned from an aesthetic virtue into a liability and was dispreferred. This negative turnaround is in line with theoretical assumptions as well as empirical findings showing that increases in complexity tend to support higher aesthetic enjoyment only up to certain threshold levels (Berlyne 1971, pp. 219–20; Güçlütürk et al. 2016).

Table 1. Quantitative text characteristics by text version: original vs. modified.

	Words per Sentence			**Words per Punctuation Mark**			**Embeddings per Main Clause**		
	O	M	Δ	O	M	Δ	O	M	Δ
Text									
Cecilia	59.5	26.5	33.0	4.5	8.0	−3.5	2.9	1.2	1.7
Earthquake	36.1	20.6	15.5	5.3	7.6	−2.3	1.3	0.7	0.6
Kohlhaas	69.7	17.5	52.2	3.9	7.0	−3.1	2.6	0.8	1.8
Marquise	35.6	27.2	8.4	4.2	7.2	−3.0	1.8	0.8	1.0
Duel 1	77.3	25.3	52.0	5.3	10.2	−4.9	2.2	0.7	1.5
Duel 2	56.4	27.4	29.0	4.9	10.2	−5.3	2.3	0.7	1.6
Summary statistics									
Mean	55.8	24.1	31.7	4.7	8.4	−3.7	2.2	0.8	1.4
Range	41.7	9.9	43.8	1.4	3.2	3.0	1.6	0.5	1.2

Note. O = original text version; M = modified text version; Δ = the difference O–M; Cecilia = *St. Cecilia or the Power of Music*; Earthquake = *The Earthquake in Chile*; Kohlhaas = *Michael Kohlhaas*; Marquise = *The Marquise of O...*; Duel = *The Duel* (two excerpts).

The fact that we observed positive effects of Kleist's original syntax only for two of the six original texts is certainly disappointing, considering that Kleist's complex syntax is a distinctive trait of his prose style. At the same time, the results are also far from providing strong and conclusive support for the standard hypothesis of "ease of processing"-driven aesthetic liking.

3.2.4. Objective Measures of Syntactic Complexity as Predictors of the Subjective Ratings

The mixed-model-derived estimates for the effects of our measures of syntactic complexity on the ratings for cognitive challenge, ease of processing, and the aesthetic evaluations are reported in Figure 3. Given the substantial overlap between the three measures of syntactic complexity and thus probable multicollinearity, separate models were estimated for each of the three measures. Because of the between-participants design and the great differences in syntactic complexity between the two conditions, the influence of the three measures of syntactic complexity on the aesthetic evaluations was modeled separately for the original and the modified text versions.

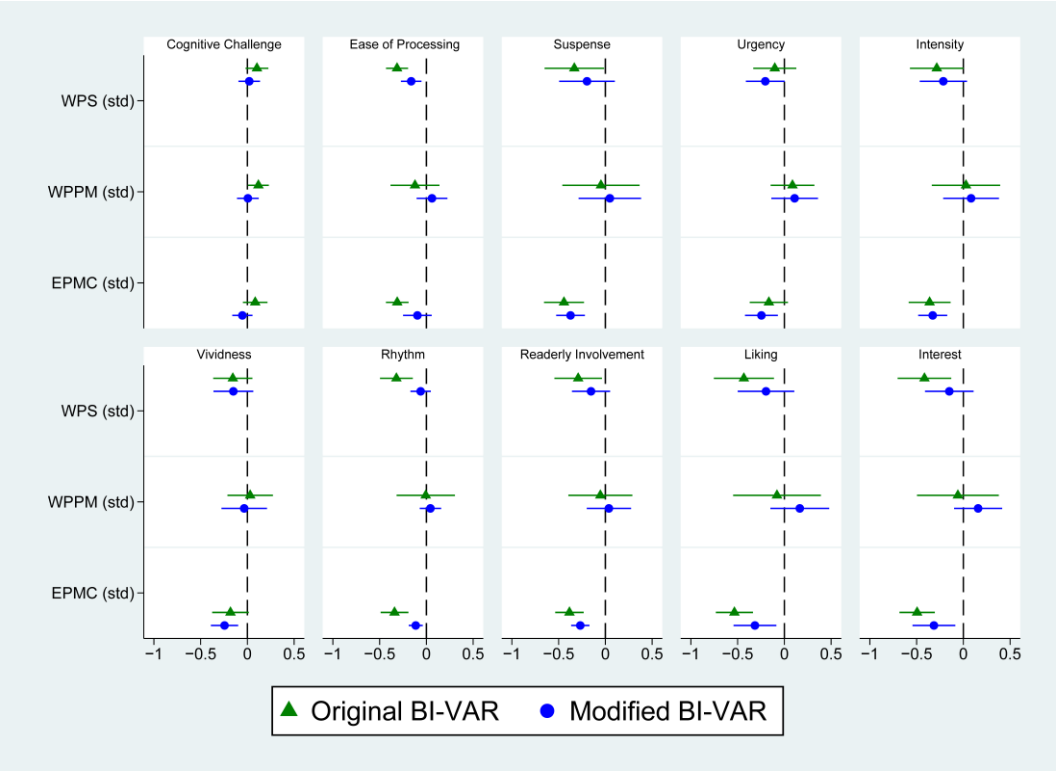

Figure 3. Mixed–model-derived estimates for the effects of within-condition standardized text measures on the mean subjective ratings for all texts. Note: WPS = words per sentence, WPPM = words per punctuation mark, EPMC = embeddings per main clause. Whiskers indicate 95% confidence intervals.

For the original text versions, higher values for embeddings per main clause and words per sentence predicted lower levels of perceived ease of processing, suspense, rhythm, readerly involvement, liking, and interest. Further, higher values for embeddings per main clause were a significant predictor of lower intensity and vividness. Higher values for words per punctuation mark predicted higher ratings for cognitive challenge.

For the modified text versions, higher values for embeddings per main clause were a significant predictor of lower suspense, urgency, intensity, vividness, rhythm, readerly involvement, liking and interest, but were not related to cognitive challenge or ease of processing. Higher values for words per sentence were associated with lower ease of processing. Differences in words per punctuation mark did not predict differences on any of the scales.

In general, the effects of our quantitative measures of syntactic complexity were stronger for the original text versions. Still, higher values for embeddings per main clause were consistently related to lower aesthetically evaluative ratings across both text versions.

Taken together, these findings suggest that when text differences are considered at the within-person level and hence separately for the two text versions, higher values for embeddings per main clause and, to a smaller extent, also higher values for Words per Sentence, are associated with less positive aesthetic evaluations for both the original and the modified text versions. Unlike the results of the between-participants comparisons, these results clearly support the standard hypothesis of cognitive fluency.

3.2.5. Relations Between Ease of Processing, Cognitive Challenge, and Aesthetic Evaluations

As reported in the Introduction, there is not only in art scholarship, but also in psychological aesthetics a long-standing tradition that considers cognitively challenging

engagement in artworks to be conducive to more profound and, in the end, also more rewarding experiences compared to mere processing ease. Moreover, experts and individuals higher in openness to experience have been shown to respond more positively to cognitive challenges (Fayn et al. 2017, 2019; Silvia et al. 2009).

In light of these assumptions and findings, we tested—in analogy to a study by Forster et al. (2013)—whether cognitive challenge and ease of processing might, in fact, both be positive predictors of aesthetic evaluations. To this end, we analyzed cognitive challenge and ease of processing both individually and conjointly as predictors of the aesthetic evaluations collected in our study.[1] Figure 4 visualizes the results of these three analyses. Green circles indicate the predictive power of cognitive challenge for the aesthetically evaluative ratings, blue circles the predictive power of ease of processing and red triangles the joint predictive power of both cognitive challenge and ease of processing.

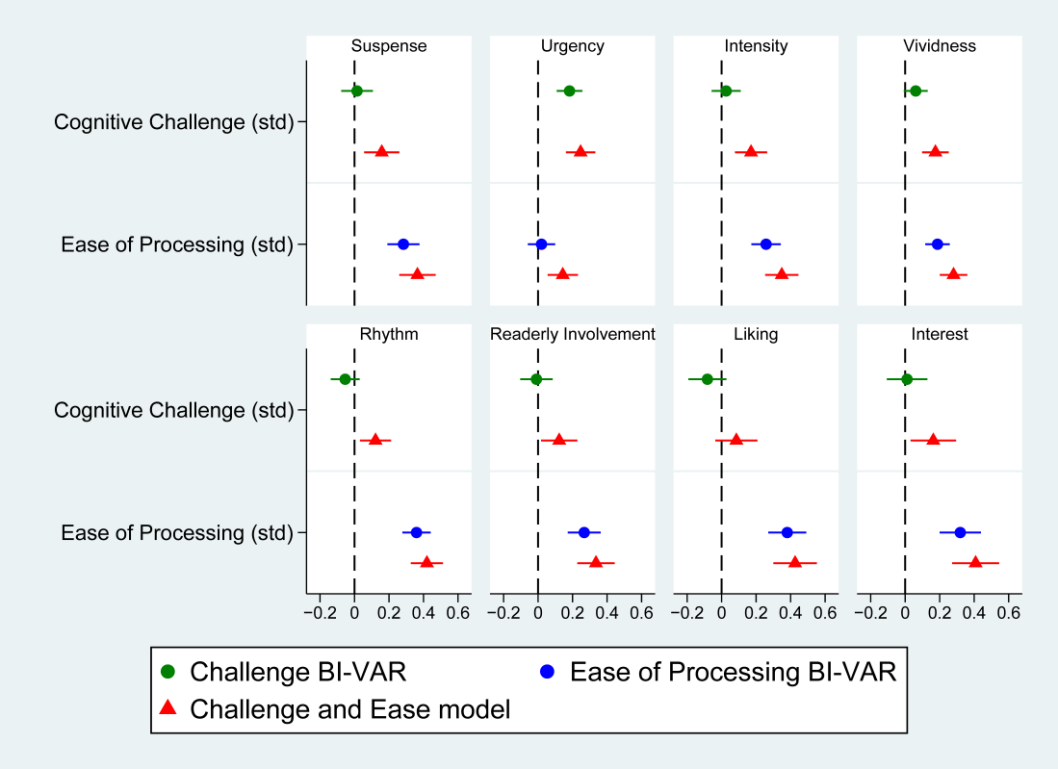

Figure 4. Mixed-model-derived coefficients for the prediction of the aesthetic evaluations by within-person standardized ratings of cognitive challenge and ease of processing for the original versions of the text in Study 1. Note: Whiskers indicate 95% confidence intervals.

When modeled independently, higher *cognitive challenge* ratings were positively associated with greater perceived *urgency* (supporting Hypotheses 4), negatively predicted *rhythm* (rejecting Hypothesis 7), and were uninformative regarding the other aesthetically evaluative ratings. In contrast, higher ratings for ease of processing predicted higher ratings on all aesthetic evaluations except for *rhythm*. However, when *cognitive challenge* and *ease of processing* were modeled conjointly as predictors of the aesthetic evaluations, results changed substantially. In this analysis, the two seemingly opposite cognitive variables are positive predictors of all aesthetic evaluations, and their joint positive effects are stronger than those of ease of processing only.

Findings of this type—where the inclusion of a third variable changes the magnitude and/or direction of another variable—are known as suppression effects (Paulhus et al. 2004). Such effects can result when two measures have partly opposite relations to a dependent variable. Regarding the present case, we propose that the measure of cognitive challenge

captures both a lack of ease of processing—which is negatively related to the aesthetic evaluations—and a judgment of complexity that is in line with the revised cognitive fluency hypothesis, i.e., a type of cognitive challenge that is appreciated by the reader, at least within certain limits. Therefore, when the variance associated with ease of processing is partialed out, the ratings for cognitive challenge are likely to be dissociated from a mere lack of understanding, and the remaining variance largely amounts to a judgment of complexity that allows for a positive evaluation—in line with both the theorizing of literary scholars and the revised hypothesis of cognitive fluency.

The hypotheses that we largely adopted from Kleist scholarship all stipulate a positive nexus between syntactic complexity (and, by implication, increased cognitive challenge) and ratings for intensity, readerly involvement, liking, etc. Critically, however, these hypotheses typically do not consider the potential importance of ease of processing. As a consequence, they are not informative regarding the critical question of how ease of processing and syntactic complexity-driven cognitive challenge might join forces rather than being mutually exclusive.

Moreover, none of the psychological models of aesthetic processing that stipulate a potentially aesthetically favorable balance of ingredients of enhanced and reduced cognitive fluency (Giora et al. 2004; Jacobs 2014; Graf and Landwehr 2015; Menninghaus et al. 2015; Miall and Kuiken 1994) have specifically proposed that a joint modeling of ratings for ease of processing and cognitive challenge might provide a way of circumventing the potential suppression effect identified above.

As a result, our interpretation of the data needs to be treated with caution. For this reason, we took recourse to data from an earlier study that involved similar variables and hence allowed for retesting the suppression effect observed in our Kleist-data. We identified a study by Fayn et al. (2015a) as a parallel case. In this study, measures of pleasure and beauty were predicted by subjective ratings of the comprehensibility and complexity of artworks. When modeled separately, complexity was a nonsignificant predictor of pleasure ($b = -0.03$, $SE = 0.03$, $p = 0.27$) and a negative predictor of beauty ratings ($b = -0.06$, $SE = 0.03$, $p = 0.03$). However, as we modeled—for these earlier data—complexity conjointly with comprehensibility, complexity turned out to be a significant positive predictor of both pleasure ($b = 0.12$, $SE = 0.03$, $p < 0.001$) and beauty ($b = 0.08$, $SE = 0.03$, $p = 0.003$), in full analogy to the results of the present study. This replication of our analysis based on a previously published dataset supports the findings we made based on the joint modeling of ease of processing and cognitive challenge as predictors of the aesthetic evaluations.

3.3. Summary and Discussion of Study 1

In a between-person analysis, the text manipulation led, as predicted, to higher mean ratings for the factor cognitive challenge for the original texts compared to the modified versions. However, the increase was only slight, and it had on average no significant effect on both self-reported ease of processing and the aesthetic evaluations. As a result, these data support neither the standard nor the revised hypothesis regarding the nexus between cognitive challenge, ease of processing, and the aesthetic evaluation factors.

In contrast to the between-person models, within-person analyses suggest that the cognitively less demanding text versions were aesthetically preferred in both the original and modified conditions, thus providing support for the standard hypothesis of cognitive fluency. This result is congruent with previous findings that within-person effects of fluency are stronger than between-person effects (Forster et al. 2015).

At the same time, a closer inspection of the data on the level of the individual texts reveals that for the two original texts which are only moderately higher in objectively

measured syntactic complexity compared to their respective modified versions (*Marquise of O. . .*, *The Earthquake in Chile*) and, moreover, lower in syntactic complexity compared to the four other original texts, the ratings show no relevant change in cognitive challenge and even higher values for ease of processing (in the case of *The Marquise of O. . .*), along with higher ratings for the aesthetic evaluations compared to the modified versions. These results provide some support for the revised hypothesis of cognitive dis/fluency. They are in line with the assumption that increases in complexity can support higher enjoyment up to certain threshold levels and hence that the relation between the two is nonlinear.

Thus, all in all, the results are inconclusive regarding the two hypotheses under scrutiny, although by no means contradictory.

4. Study 2

In order to further investigate how differences in the experimental design and levels of analyses would affect the results obtained for the competing hypotheses tested in our study, we performed a second experiment in which all readers read both the original and the modified text versions and rated them on the same rating items that were used in Study 1.

Beyond the possible role of the experimental design, we tested two further possibilities that could explain the relatively small effect sizes observed in Study 1.

(1) Guided by previous findings that show a stronger role of subjective fluency ratings compared to objective stimulus measures in predicting liking ratings (Forster et al. 2013), we tested the possibility that the lack of consistent evidence for either hypothesis in Study 1 might be due to a similar result pattern in our case. We did so by directly comparing the predictive validity of the subjective ratings of ease of processing and cognitive challenge with the predictive power of embeddings per main clause, i.e., the objective measure of syntactic complexity that was the strongest predictor of the aesthetically evaluative ratings in Study 1.

(2) Preferences for cognitively demanding kinds of artworks and other stimuli have been shown to vary between individuals. Specifically, context-based expertise and the personality trait of openness to experience relate to greater positive sensitivity to subjectively perceived complexity and processing ease (Fayn et al. 2015a, 2015b, 2017; Silvia 2013; Silvia and Berg 2011; Silvia et al. 2009). Therefore, we tested whether relations between the ratings for the aesthetic evaluation factors and the subjective reports for ease of processing and cognitive challenge varied as a function of (a) a general preference for reading novels and shorter narratives—which has been shown to be related to higher openness to experience (Medford and McGeown 2012), (b) prior positive reading experiences specifically with Kleist, and (c) age.

4.1. Method

4.1.1. Participants

Participants were recruited using the same channels and announcements as Study 1. A total of 128 participants completed this study (65% female, age: M = 38.45 years, SD = 18.24, range: 18–85 years). The inclusion criteria were a minimum age of 18 years and having German as a native language (114 were German monolingual, and 14 bilingual with German and another language).

Participants were again predominantly highly educated (92% had the general qualification for university entrance). On average, they reported reading novels and/or short narratives several times per month (M = 2.87, SD = 1.70, response scale: 0 = *never*, 1 = *once per year*, 2 = *several times per year*, 3 = *several times per month*, 4 = *once per week*, 5 = *several times per week*, 6 = (*almost) daily*), and liking reading novels and/or short narratives (M = 1.76,

SD = 1.40, response scale: −3 = *dislike very much*, −2 = *dislike*, −1 = *rather dislike*, 0 = *neutral*, +1 = *rather like*, +2 = *like*, +3 = *like very much*).

4.1.2. Stimuli

All participants were presented with both the original and the modified versions of the excerpts from Kleist's narratives *Caecilia*, *The Earthquake in Chile*, *Michael Kohlhaas*, and *The Marquise of O*... that were used in Study 1, resulting in a total of eight text readings per participant. In order to keep the study length manageable, Study 2 resorted to only four of the six texts that were presented in Study 1. Specifically, we selected two texts for which our quantitative measures showed only a modest version difference and two texts for which the quantitative version difference was high. In Study 1, these two groups of stimuli were not balanced, as four of the six narratives showed a high version difference. Moreover, of the four text excerpts with high version differences that were presented in Study 1, we dropped the two excerpts that were the only ones to be extracted from the same narrative, i.e., *The Duel* 1 and *The Duel* 2.

For all details regarding the four original narrative excerpts and their modified variants, see the pertinent sections of Study 1.

4.1.3. Design

Rather than the 2 × 6 between-participants factorial design of Study 1, we implemented a 2 × 4 within-participants factorial design, again with the factors *Version* (original or modified version) and *Text* (four excerpts from four narratives) fully crossed and with a Latin square design controlling for potential effects of presentation order. Each row of the 4 × 4 Latin square was realized 32 times. As the original texts might be rated differently after reading the modified text versions than before reading these (and the other way round), we controlled for potential order effects of the factor *Version* by crossing the Latin square conditions with 16 combinations of the two versions of the four texts. All text versions were rated on the same scales as in Study 1. For all further details, see Study 1, Section "Method."

In all other regards, the procedural details of Study 2 were identical to those of Study 1.

4.1.4. Statistical Analysis

We tested all of our hypotheses using mixed-effects regression models with crossed random effects for participants and texts, testing for main effects of text modification (factor Version). As we used a Latin square design for balancing text order effects, any order effects can be attributed to repeated reading independent of specific text sequences.

We used sensitivity analyses to test whether the results were dependent on potentially confounding variables. First, we included gender. Second, we analyzed reading times and checked whether the results were affected by excluding 10 outlying observations with reading times that deviated more than 3 *SD*s from the average reading time across all texts and participants. As neither sensitivity analysis changed the findings, we do not report their results here.

To replicate the results from Study 1, we again compared all ratings for the original and modified text versions. Subsequently, we analyzed the continuous effect of the experimental text manipulation by modeling the relationship between all aesthetically evaluative ratings and embeddings per main clause, i.e., the objective measure of syntactic complexity that best predicted differences in the ratings in Study 1. Next, we tested how self-reported cognitive challenge and ease of processing were on average related to the self-reported aesthetic evaluations.

4.2. Results

4.2.1. Reading Times, Text Familiarity, and Author Recognition

We obtained data for a total of 1023 text readings (128 participants each read 8 texts; one participant did not complete one of the readings). On average, it took participants 3 min 22 s to read one text (M = 202.39 s, SD = 73.39 s). As in Study 1, reading times for the original texts were not significantly different from those for the modified texts (M_{diff} = 3.85 s, SE = 3.60, p = 0.29). Analogously, linguistic studies on the processing of optional punctuation marks in English texts showed no differences in overall reading times for sentences with and without such marks (cf. Stine-Morrow et al. 2010; Angele et al. 2024).[2]

We observed a general trend towards shorter reading times for texts presented later in the trajectory of the experiment (coefficient = −6.00, SE = 0.76, $p < 0.001$). These results are very similar to those obtained for the same tests in Study 1.

With 22% familiarity for the original versions and 20% for the modified versions, our participants' self-indicated familiarity with the presented texts was not significantly different for the two versions ($\chi 2 = 0.89$, $p = 0.35$). Correct author attributions regarding the texts amounted to 14%; they were likewise not significantly affected by our experimental modification (original = 14%, modified = 14%, $\chi 2 = 0.17$, $p = 0.92$).

4.2.2. Version-Dependent Differences Between the Ratings for the Four Texts

In accord with Hypothesis 1, the original versions were again rated higher on the factor cognitive challenge, $b = 0.56$, $SE = 0.06$, $p < 0.001$, and lower for ease of processing, $b = -0.69$, $SE = 0.07$, $p < 0.001$. The effect size (Westfall et al. 2014) of this result is small to medium ($d = 0.28$), compared to the very small effect size observed in Study 1 ($d = 0.09$).

Contrary to the results of Study 1, and supporting the standard hypothesis of cognitive fluency, the modified texts were on average rated higher on the scales for overall liking, $b = -0.28$, $SE = 0.07$, $p < 0.001$, suspense, $b = -0.16$, $SE = 0.06$, $p = 0.007$, vividness, $b = -0.15$, $SE = 0.05$, $p = 0.003$, rhythm, $b = -0.50$, $SE = 0.06$, $p < 0.001$, readerly involvement, $b = -0.17$, $SE = 0.06$, $p = 0.002$, and interest, $b = -0.18$, $SE = 0.07$, $p = 0.01$, but not for urgency, $b = 0.06$, $SE = 0.05$, $p = 0.29$, and intensity, $b = -0.11$, $SE = 0.06$, $p = 0.08$.

In Study 1, version-dependent differences in cognitive challenge were observed for only two of the texts (*Cecilia, Michael Kohlhaas*), and differences in ease of processing were either not significant or pointed in opposite directions. In contrast, the within-person design of Study 2 yielded results that are much more consistently in accord with the standard hypothesis of cognitive fluency: Higher ratings for ease of processing predicted higher aesthetic evaluations, and higher ratings for cognitive challenge predicted lower aesthetic evaluations.

Differences between the versions regarding the aesthetic evaluations were less consistent. For the syntactically most challenging text (*Michael Kohlhaas*), the original version was rated lower on all aesthetic evaluation factors except urgency and interest. However, version differences for the other texts were not substantial, with the exception of significantly lower ratings for rhythm for the original versions of three of the four texts.

An additional analysis using that measure of syntactic complexity which was the strongest predictor of rating differences between versions in Study 1 (i.e., embeddings per main clause) yielded results that are very similar to those reported in Study 1 (for details, see Appendix A Figure A2). Again, more embeddings per main clause predicted higher cognitive challenge and lower ease of processing, along with lower suspense, intensity, vividness, rhythm, readerly involvement, liking, and interest. Thus, for all ratings except urgency, the standard hypothesis of cognitive fluency was supported in Study 2.

4.2.3. Ease of Processing, Cognitive Challenge, and Aesthetic Evaluations

As for the data of Study 1, we also performed for those of Study 2 two analyses regarding the effects of ease of processing and cognitive challenge on the aesthetic evaluations. In the first analysis, ease of processing and cognitive challenge were modeled as separate and in the second as joint predictors of the aesthetic evaluations. The results are shown in Figure 5.

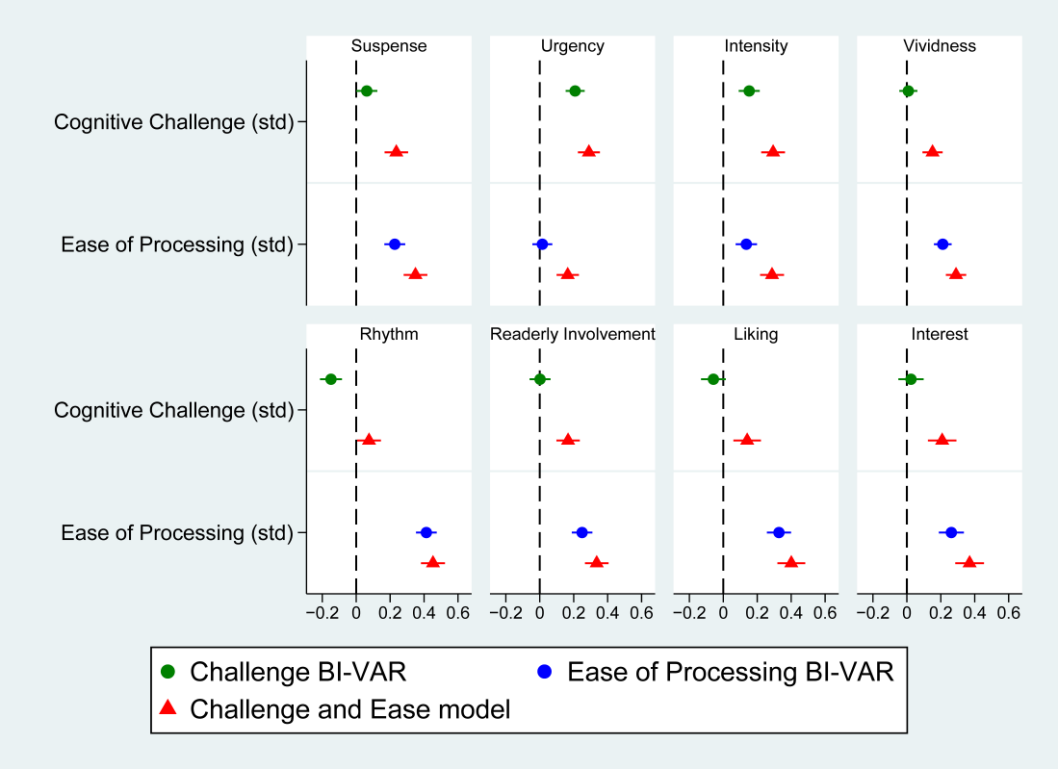

Figure 5. Mixed-model-derived coefficients for the prediction of aesthetic evaluation ratings by within-person standardized ratings for cognitive challenge and ease of processing. Note: Whiskers indicate 95% confidence intervals.

When modeled independently, higher cognitive challenge ratings were associated with greater urgency and intensity (supporting Hypotheses 4 and 5) and lower rhythm (further supporting the rejection of Hypothesis 7), whereas higher ease of processing was associated with higher ratings on all aesthetic evaluations except for rhythm. Regarding the other aesthetic evaluations, the ratings for cognitive challenge had no predictive power. These separate analyses hence suggest that higher ease of processing fairly consistently supports higher positive aesthetic evaluations. At the same time, there is also evidence in support of the revised hypothesis of cognitive dis/fluency insofar as higher perceived cognitive challenge does not consistently compromise aesthetic liking and even supports higher ratings on two of the relevant scales for aesthetic evaluations, i.e., urgency and intensity.

When cognitive challenge and ease of processing were modeled conjointly as predictors of the aesthetic evaluations, results are again substantially different. In this analysis, cognitive challenge, too, is a positive predictor of all aesthetic evaluations except for rhythm.

This replication of the pertinent findings of Study 1 in a within-participants comparison of the original and modified text versions further supports our hypothesis of a "suppression effect" in the separate modeling of ease of processing and cognitive challenge.

4.2.4. Objective Textual Complexity Versus Subjectively Perceived Fluency

We next tested a model that also includes the strongest objective predictor of aesthetic evaluations, namely, the number of embeddings per main clause. The results of this analysis are shown in Figure 6. Embeddings per main clause were again a significant negative predictor of liking, readerly involvement, rhythm, vividness, intensity, and suspense. However, these effects are much smaller than the positive effects observed for the two subjectively measured variables, i.e., ease of processing and cognitive challenge. These findings replicate previous work showing that subjective fluency of processing is a stronger predictor of subjective liking than are objective measures of fluency (Forster et al. 2013).

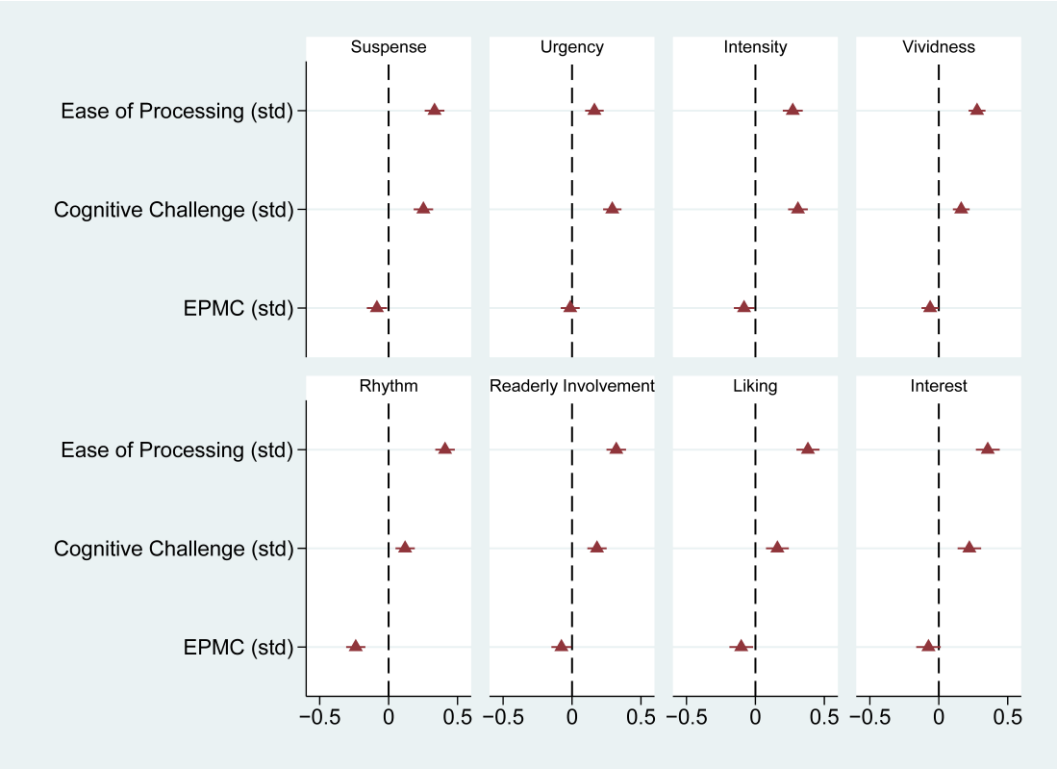

Figure 6. Mixed-model-derived coefficients for predicting aesthetic evaluation scales by within-person standardized ratings of cognitive challenge and ease of processing and the measure of embeddings per main clause. Note: EPMC = embeddings per main clause. Whiskers indicate 95% confidence intervals.

Given that our objective text measures exclusively reflect dimensions of syntactic complexity, it is not surprising that subjective ratings of ease of processing and cognitive challenge capture a larger variance of the overall readerly impressions when reading Kleist's texts.

It is all the more noteworthy that our purely formalist text measure, embeddings per main clause, distinguished two text groups in Study 1: We did find positive effects of our stylistic target features on aesthetic evaluations, but only for original versions with intermediate scores for syntactic complexity. Original versions that are high to very high in levels of syntactic complexity had negative effects on the aesthetic evaluations. Thus, in line with the"inverted u-shape" effect (Berlyne 1971, pp. 219–20), we found a positive effect of Kleist's syntax only up to a certain threshold level.

4.2.5. Individual Differences in Relations Between Cognitive Challenge and Aesthetic Evaluations

In line with previous findings showing that expertise in a specific domain and greater openness to experience are associated with greater positive reactions to cognitive challenge, we tested in a next step whether the relations of the aesthetic evaluations with cognitive challenge and ease of processing differed as a function of specific reader characteristics. Specifically, we tested:

(1) whether the quality of previous experiences with Kleist was related to more positive responses to greater cognitive challenge;

(2) whether the frequency and liking of reading novels predicted more positive experiences in responses to greater cognitive challenge; and

(3) whether older people—who should have more reading experience as well as more refined vocabularies and more extended general knowledge (Baltes et al. 1999; Birney and Sternberg 2011; McArdle et al. 2002)—responded more positively to greater cognitive challenge.

In line with the last prediction, age was negatively correlated with average self-reported cognitive challenge: Older persons tended to find the texts less challenging overall. We therefore controlled for average cognitive challenge within the age models. No predictions were made for the moderation of the relations between the aesthetic evaluations and ease of processing.

Previous Experiences with Kleist. Mixed-model-derived interaction results are presented in Table 2. Previous experiences with Kleist significantly moderated the relations between cognitive challenge and suspense, vividness, rhythm, liking, and interest, suggesting that those participants who had more positive previous experiences with Kleist responded more positively to higher cognitive challenge in these aesthetic evaluations. At the same time, none of the relations between ease of processing and the aesthetic evaluations were moderated by previous experiences with Kleist.

Table 2. Individual differences in moderation of the influence of cognitive challenge and ease of processing on aesthetic evaluations.

Variable	Kleist Experience		General Enjoyment of Reading Narratives		Age	
	Ease Slope	Challenge Slope	Ease Slope	Challenge Slope	Ease Slope	Challenge Slope
Suspense	−0.004	0.11 **	0.04	0.10 **	−0.06	0.07
Urgency	0.02	0.05	−0.01	0.04	0.03	0.04
Intensity	−0.08	0.06	−0.02	0.05	−0.09 **	0.03
Vividness	−0.02	0.09 *	−0.004	0.05	−0.03	0.07 *
Rhythm	−0.06	0.13 **	0.03	0.11 **	−0.04	0.08 *
Readerly Involvement	−0.05	0.08	0.03	0.08 *	−0.11 **	0.08 *
Liking	−0.06	0.15 **	0.01	0.18 ***	−0.06	0.15 ***
Interest	−0.06	0.11 *	0.01	0.16 ***	−0.08	0.09 *

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (all two-tailed).

Follow-up simple slope analyses were performed to further understand the significant interactions. The slopes were compared for participants one standard deviation above and one standard deviation below the mean on previous Kleist liking (high and low likers). For the factors suspense and vividness, the slopes were positive for high and low likers. For rhythm, liking, and interest, high likers had positive relations to cognitive challenge, whereas nonsignificant relations were obtained for low likers.

Enjoyers and Frequent Readers of Novels and Short Stories. Participants' ratings of overall enjoyment and frequency of engagement with novels and short stories (i.e.,

irrespective of any special focus on their potential prior readings of Kleist's narratives) show an effect pattern that is similar to the effects of previous experiences with Kleist. These participants show more positive relations between cognitive challenge and suspense, rhythm, readerly involvement, liking, and interest (Table 2). At the same time, similar to prior experiences with Kleist, none of the relations between ease of processing and the aesthetic evaluations were moderated by prior reading engagement with novels and short stories in general. These relations thus appear to hold independently of individual differences, whereas the effects of cognitive challenge on the aesthetic evaluations are strongly dependent on individual differences.

Follow-up simple slope analyses were performed to further understand the significant interactions. Slopes were compared for those one standard deviation above and one standard deviation below the mean on reading tendencies. For suspense, regardless of reading tendencies, the slopes for cognitive challenge were positive. For rhythm, readerly involvement, liking, and interest, more avid readers had positive slopes for cognitive challenge, whereas less avid readers' slopes were nonsignificant.

Age. Higher age predicted more positive relations between cognitive challenge and vividness, rhythm, readerly involvement, liking, and interest. Higher age also had a positive effect on the relations between ease of processing, on the one hand, and intensity and readerly involvement, on the other. Follow-up simple slope analyses were performed to further understand the significant interactions. Slopes were compared for participants one standard deviation above and one standard deviation below the mean for age. For interest, the cognitive challenge slopes were positive for both older and younger participants. For vividness, rhythm, readerly involvement, and liking, older participants had positive relations with cognitive challenge, whereas the slopes for younger participants were nonsignificant. For readerly involvement and intensity, the ease of processing slopes were positive for both younger and older participants.

4.3. Summary and Discussion of Study 2

Compared to the results of Study 1, the direct comparison of the two text versions within participants resulted in a pronounced contrast effect: The original versions were rated higher for cognitive challenge, lower for ease of processing, and also lower for some of the aesthetic evaluations. This result pattern supports the standard hypothesis of cognitive fluency-driven aesthetic liking. Notably, however, despite the substantial differences between the text versions regarding their objectively measured syntactic complexity, significant adverse effects on the aesthetic evaluations were very limited in terms of size, and some important aesthetic evaluations were not negatively affected at all by higher syntactic complexity.

Moreover, when modeling ease of processing and cognitive challenge conjointly as predictors of the aesthetic evaluations, the apparently contradictory variables actually joined forces in positively predicting the aesthetic evaluations. This result replicates the findings of Study 1. The fact that negative effects of higher cognitive challenge on the aesthetic evaluations were only obtained in the within-person comparison of two versions of the same text also supports concerns that these effects may primarily be an artifact of the experimental method. After all, in natural reading situations, readers barely ever read two syntactically different versions of content-wise convergent texts one after the other.

Furthermore, our analysis of individual differences also yielded substantial support for the revised hypothesis of cognitive fluency. Individuals who reported earlier enjoyable encounters with Kleist's texts and/or a general tendency to frequently read and enjoy novels as well as those who were older showed a clear tendency to report the rating pattern predicted by the revised hypothesis of cognitive dis/fluency—that is, avid readers were

positively sensitive to cognitive challenge, whereas the other readers were not. These results support the notion that higher cognitive challenge can well be an asset and enticement for enjoyment as long as it is compatible with rather than disruptive of the rewards of ease of processing. The limits of this interaction of ease of processing and cognitive challenge appear to be largely determined by individual differences.

A recent study on the effects of poetic diction in responses to poems, proverbs, and humoristic couplets (Menninghaus et al. 2023) focused in a similar fashion on two opposite feature groups: features of parallelistic diction (rhyme, meter, alliteration, etc.) that enhance predictability and ease of (cognitive) processing and features of deviation (such as incomplete syntax and word forms, inversion of the canonical word order, etc.) that reduce predictability and ease of cognitive processing. Positive effects of parallelism were observed across all readers, whereas positive effects of deviations were moderated by individual differences: only participants that scored comparatively high in poetry preference and rhythmic ability showed positive effects of increased cognitive challenge.

5. General Discussion

This study expands the understanding of the relations between cognitive challenge, ease of processing, and the aesthetic reward dimensions in responses to artworks, specifically, to high-art literature. From the perspective of both literary authors and scholars of literature, pronounced detours from standard language uses should enhance rather than reduce aesthetic appreciation. After all, such features not only contribute to markedly individual author styles, but are in all likelihood aimed at supporting specific positive aesthetic effects. At the same time, the standard hypothesis of cognitive fluency-driven aesthetic liking typically predicts that higher cognitive challenge is likely to reduce ease of processing and thereby aesthetic liking.

The results of our two studies support both hypotheses. More importantly, they show that the two apparently competing hypotheses are actually not mutually exclusive, but rather complement each other in a meaningful way. When more challenging and de-complexified versions of the same texts are compared for different levels of cognitive challenge within persons, the chances are high that robust adverse effects of the more demanding/disfluent text versions on aesthetic evaluations will be obtained. However, in a between-person design and hence in the absence of a direct contrast—which is an experimental artefact not found in natural reading situations—, such adverse effects do not necessarily materialize, even when the levels of syntactic complexity and cognitive challenge are as high as they are for readers of Kleist's narratives. Further highlighting the compatibility of our results with earlier findings, the subjective reports of fluency or ease of processing were overall stronger predictors of the aesthetic evaluations than were the objectively measured differences in text-based syntactic complexity.

When ease of processing and cognitive challenge were modeled conjointly rather than separately as predictors of the aesthetic evaluations, both variables related positively to the aesthetic evaluations. In this analysis, the two apparently competing hypotheses cease being mutually exclusive; rather, both hold at the same time. This result lends further support to an earlier study (Fayn et al. 2015a) that reported a suppression effect for the positive role of cognitive challenge only in analyses that modeled ease of processing and cognitive challenge separately as predictors of aesthetic liking, but not for the joint modeling of these two predictors. Moreover, our result is in accord with the theorizing of literary scholars. At the same time, it does by no means reject the standard hypothesis of cognitive fluency, as higher ease of processing supported higher aesthetic evaluations in this analysis as well.

The positive effects of cognitive challenge observed in this study were particularly driven by three person variables: previous enjoyable encounters with Kleist's narratives, a general tendency to enjoy reading narratives, and older age. These results are in line with well-replicated findings according to which expertise and the personality trait of openness to experience moderate people's experiences of cognitive challenge and ease of processing in response to complex artworks and films (e.g., Fayn et al. 2017; Silvia 2005, 2007, 2013; Silvia and Berg 2011). Specifically, experts and those with a higher openness to experience tend to be more positively affected by complexity and less negatively affected by reduced ease of processing. While we did not measure expertise and openness in the study reported here, enjoyment, general frequency of reading, and previous positive experiences with reading Kleist's narratives specifically are in line with greater expertise and higher openness.

Individuals build up their tastes and preferences over time, and authors do not cater equally to the tastes of all individuals. Therefore, data obtained from readers who have previously and repeatedly engaged in Kleist readings on their own are particularly valuable data if the goal is to understand "real" Kleist readings outside of highly artificial laboratory contexts. Importantly, as most literarily ambitious authors cannot expect more than a fairly select constituency of readers, this selective support for our Hypotheses 3 through 10 is strong enough to safeguard our findings against outright pessimism regarding the fate of cognitively demanding literature.

To be sure, it is also interesting to learn how individuals who read less, let alone repeatedly read Kleist on their own, respond to his texts. However, the results of our study strongly suggest that mean ratings for ambitious literary texts obtained in a laboratory context from individuals irrespective of their earlier experiences with and preferences for literature are very likely not to reflect self-motivated reading processes in the wild. Therefore, analyzing relevant differences on the part of the participating individuals is strongly called for. Only if this rule is observed will data collected in the lab or via the Internet have the potential to reveal profiles of "real" reading experiences. Novice-expert studies could reveal more about this dimension of individual differences in responses to literary artworks of high linguistic complexity. Moreover, in order to arrive at a deeper understanding of individual differences in responses to cognitively challenging artworks, qualitative studies are called for.

Finally, without recourse to humanist scholarship on Kleist, we could not have developed the hypotheses we tested. Humanist scholars typically write about artworks they like or even admire. This positivity bias is anything but a mere limitation; it is also an enabling mechanism for appreciating the inherent virtues of artworks. Our analyses of individual differences suggest that those readers who actually read and often reread, in a self-motivated fashion, the narratives under consideration in the present study actually have aesthetic experiences that support the assumptions advocated in humanist scholarship on Kleist. Other empirical studies on actual art reception may therefore be well advised not only to pursue the more general hypotheses of psychological aesthetics, but also to consult humanist scholarship on the arts, along with considering the importance of individual differences.

Author Contributions: Conceptualization, W.M., V.K. and W.S.; methodology, W.S. and K.F.; formal analysis, W.S. and K.F.; investigation, V.K. and K.F.; resources, Max Planck Institute of Empirical Aesthetics; data curation, V.K. and K.F.; writing—original draft preparation, W.M.; writing—review and editing, W.M. and K.F.; visualization, K.F.; supervision, W.M.; project administration, W.S. and K.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Max Planck Institute of Empirical Aesthetics.

Institutional Review Board Statement: All experimental procedures were approved by the Ethics Council of the Max Planck Institute for Empirical Aesthetics (approval 12/2017).

Informed Consent Statement: All experimental procedures were undertaken with the written informed consent of each participant. Participants in Study 1 received € 10 for their participation, whereas participants in Study 2 received €20, as we expected that Study 2 would take longer than Study 1 since each participant was assigned eight instead of six text readings.

Data Availability Statement: All original data are stored on the server of the Max Planck Institute for Empirical Aesthetics and can be made accessible upon request.

Conflicts of Interest: There are no conflicts of interest to declare.

Appendix A

1 Cognitive Challenge, Ease of Processing, and Aesthetic Evaluation Items

Table A1. German items and English translations.

#	Variable	German Wording	English Translation
10	Cognitive Challenge	Wie komplex wirkt der Text auf Sie?	How complex does the text seem to you?
11		Wie anspruchsvoll fanden Sie die Lektüre des Textes?	How challenging did you find it to read of the text?
19		Wie markant ist der Stil der Erzählung?	How distinctive is the style of the narrative?
21		Wie sehr weicht der Stil des Textes von normalem Sprachgebrauch ab?	How much does the style of the text deviate from normal usage?
12	Ease of Processing	Wie leicht ist der Text zu verstehen?	How easy is it to understand the text?
01	Suspense	Wie spannend finden Sie die Geschichte?	How exciting do you find the story?
02		Wie dramatisch finden Sie die Art des Erzählens?	How dramatic do you find the style of narration?
03		Wie packend finden Sie die Erzählung?	How gripping do you find the narrative?
04	Urgency	Wie temporeich finden Sie die Abfolge der Ereignisse?	How fast-paced do you find the sequence of events?
05		Wie sehr vermittelt Ihnen das Fortschreiten der Erzählung einen Eindruck von Dringlichkeit?	How much does the progression of the narrative give you a sense of urgency?
06		Wie gedrängt ist die Schilderung des Geschehens?	How compressed is the description of the events?
22	Intensity	Wie intensiv war das Leseerlebnis für Sie?	How intense was the reading experience for you?
25		Als wie kraftvoll empfinden Sie die Sprache der Erzählung?	How powerful do you find the language of the narrative?
26		Wie stark ist der Eindruck, den der Text bei Ihnen hinterlassen hat?	How strong an impression did the text leave on you?
07	Vividness	Wie lebendig treten Ihnen beim Lesen die Figuren bzw. die geschilderten Ereignisse vor Augen?	How vividly do the characters or the events described appear to your mind while reading?
08		Wie anschaulich ist Ihr Eindruck von den berichteten Ereignissen?	How vivid is your impression of the reported events?
09		Wie detailreich finden Sie die Geschichte geschrieben?	How rich in detail do you find the story?
20	Rhythm	Wie elegant finden Sie die Erzählung geschrieben?	How elegantly written do you find the story?
23		Wie rhythmisch finden Sie die Sprache des Textes?	How rhythmic do you find the language of the text?
27		Wie fließend finden Sie die Sprache des Textes?	How fluent do you find the language of the text?
13	Readerly Involvement	Wie sehr hat die Erzählung Ihre Aufmerksamkeit gefesselt?	How much did the narrative hold your attention?
14		Wie sehr hatten Sie das Gefühl, beim Lesen in einer anderen Welt gewesen zu sein?	How much did you feel like you were in another world while reading?
15		Wie sehr haben Sie sich beim Lesen in die Gefühle und Gedanken der (Haupt-) Figuren hineinversetzt?	How much did you empathize with the feelings and thoughts of the (main) characters while reading?
16	Liking	Wie gut gefällt Ihnen der Text?	How much do you like the text?
17		Wie gern würden Sie die gesamte Erzählung lesen?	How much would you like to read the entire narrative?
18		Wie sehr würden Sie den Text einem Freund/einer Freundin empfehlen?	How much would you recommend the text to a friend?
24	Interest	Wie interessant finden Sie den Text?	How interesting do you find the text?

2 Figures

(A) Original	(B) Modified
Kohlhaas, der, beim Eintritt in den Saal, einen Junker Hans von Tronka, der ihm entgegen kam, bei der Brust faßte, und in den Winkel des Saals schleuderte, daß er sein Hirn an den Steinen verspritzte, fragte, während die Knechte die anderen Ritter, die zu den Waffen gegriffen hatten, überwältigten, und zerstreuten: wo der Junker Wenzel von Tronka sei?	Beim Eintritt in den Saal kam Kohlhaas ein Junker Hans von Tronka entgegen, den er bei der Brust faßte und in den Winkel des Saals schleuderte, daß er sein Hirn an den Steinen verspritzte. Die Knechte überwältigten und zerstreuten die anderen Ritter, die zu den Waffen gegriffen hatten. Währenddessen fragte Kohlhaas: wo der Junker Wenzel von Tronka sei?
Sentence structure:	Sentence structure:
[1 [1.2 [1.2.1] 1.2 [1.2.2]] 1 [1.3 [1.3.1] 1.3] [1.1]].	[1 [1.1 [1.1.1]]]. [2 [2.1]]. [3 [3.1]].
Sentences: 1; Words: 58; Commas: 12	Sentences: 3; Words: 58; Commas: 3
Embeddings per main clause: 6.0	Embeddings per main clause: 1.3
Words per punctuation mark: 4.1	Words per punctuation mark: 8.3
Words per sentence: 58	Words per sentence: 19.3

green:	Main clause
yellow:	Final-embedded clause in main clause or in final-embedded clause
orange:	Center-embedded clause
light red:	Final-embedded clause in center-embedded clause
red:	Multiple center embedding (center-embedded clause in center-embedded clause)

Figure A1. German version of an exemplary sentence structure from Kleist's novella Michael Kohlhaas in the original and modified versions. Note: (**A**) In the original version, the processing of the main clause 1 is interrupted by a combination of multiple center- and final-embeddings. The object complement 1.1 of the verb "fragte" ("asked") is postponed by another instance of multiple center-embeddings. (**B**) In the modified version, the propositional information has been split up into three main clauses with final-embedded clauses. Unlike the case of central-embedding, the processing of the main clauses runs without interruption.

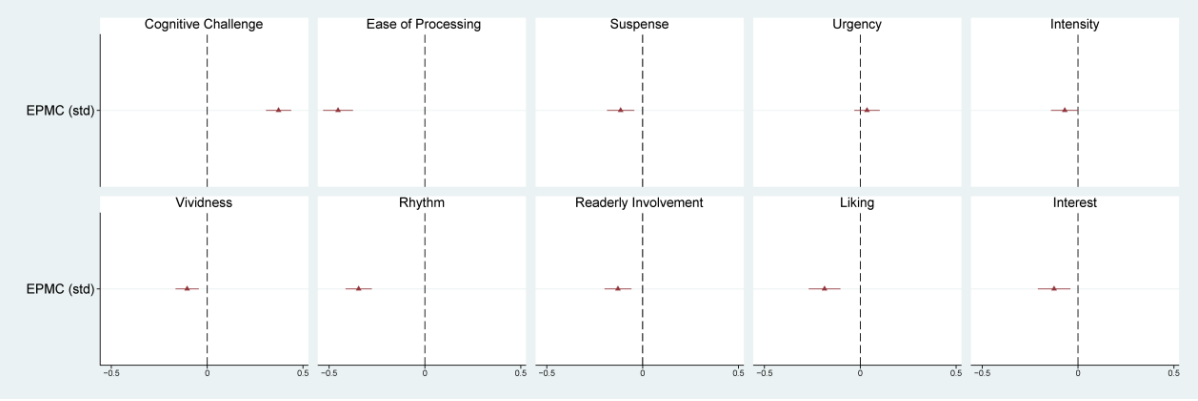

Figure A2. Study 2's mixed-model-derived estimates for the effects of standardized EPMC on the mean subjective ratings for all texts. Note: EPMC = embeddings per main clause. Whiskers indicate confidence intervals.

Notes

[1] The idea of modeling ease of processing and cognitive challenge conjointly as predictors of the aesthetic evaluations was not part of the a priori-design of the two studies reported in this article. It occurred to us only in retrospect. Importantly, the temporal genealogy of these additional analyses of our data does not affect their validity. We therefore report—for the purposes of better readability—all pertinent analyses as parts of one study.

[2] Still, local reading times measured for the immediate context of the optional punctuations did show slightly increased fixation times for the presence vs. absence of optional punctuations in these two studies.

References

Admoni, Wladimir G. 1987. *Die Entwicklung des Satzbaus der deutschen Literatursprache im 19. und 20. Jahrhundert*. Berlin: Akademie. ISBN 9783050001555.

Alter, Adam L. 2013. The Benefits of Cognitive Disfluency. *Current Directions in Psychological Science* 22: 437–42. [CrossRef]

Angele, Bernhard, Ismael Gutiérrez-Cordero, Manuel Perea, and Ana Marcet. 2024. Reading (,) With and Without Commas. *Quarterly Journal of Experimental Psychology* 77: 1190–200. [CrossRef]

Anglada-Tort, Manuel, Jochen Steffens, and Daniel Müllensiefen. 2019. Names and Titles Matter: The Impact of Linguistic Fluency and the Affect Heuristic on Aesthetic and Value Judgements of Music. *Psychology of Aesthetics, Creativity, and the Arts* 13: 277. [CrossRef]

Bach, Emmon, Colin Brown, and William Marslen-Wilson. 1986. Crossed and Nested Dependencies in German and Dutch: A Psycholinguistic Study. *Language and Cognitive Processes* 1: 249–62. [CrossRef]

Baltes, Paul B., Ursula M. Staudinger, and Ulman Lindenberger. 1999. Lifespan Psychology: Theory and Application to Intellectual Functioning. *Annual Review of Psychology* 50: 471–507. [CrossRef]

Belfi, Amy M., Edward A. Vessel, and G. Gabrielle Starr. 2018. Individual Ratings of Vividness Predict Aesthetic Appeal in Poetry. *Psychology of Aesthetics, Creativity, and the Arts* 12: 341–50. [CrossRef]

Belke, Benno, Helmut Leder, Tilo Strobach, and Claus Christian Carbon. 2010. Cognitive Fluency: High-level Processing Dynamics in Art Appreciation. *Psychology of Aesthetics, Creativity, and the Arts* 4: 214–22. [CrossRef]

Berlyne, Daniel E. 1971. *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts. ISBN 9780390086709.

Birney, Damian P., and Robert J. Sternberg. 2011. The Development of Cognitive Abilities. In *Developmental Science: An Advanced Textbook*. Edited by Marc H. Bornstein and Michael E. Lamb. London: Psychology Press, pp. 353–88. ISBN 978-1848726116.

Bischoff, Heinrich. 1899. Der Satzbau bei Heinrich von Kleist. *Zeitschrift für den deutschen Unterricht* 13: 713–20.

Carrière, Mathieu. 1984. *Für eine Literatur des Krieges*. Frankfurt/Basel: Stroemfeld. ISBN 9783878771517.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press. ISBN 9780262530071.

Cohen, Jonathan. 2001. Defining Identification: A Theoretical Look at the Identification of Audiences With Media Characters. *Mass Communication and Society* 4: 245–64. [CrossRef]

Dionysius of Halicarnassus. 1985. On Literary Composition. In *Dionysius of Halicarnassus: Critical Essays*. Translated by Stephen Usher. Cambridge, MA: Harvard University Press, vol. 2, ISBN 9780674995130.

Fayn, Kirill, Carolyn MacCann, Niko Tiliopoulos, and Paul J. Silvia. 2015a. Aesthetic Emotions and Aesthetic People: Openness Predicts Sensitivity to Novelty in the Experiences of Interest and Pleasure. *Frontiers in Psychology* 6: 1977. [CrossRef]

Fayn, Kirill, Niko Tiliopoulos, and Carolyn MacCann. 2015b. Interest in Truth Versus Beauty: Intellect and Openness Reflect Different Pathways Towards Interest. *Personality and Individual Differences* 81: 47–52. [CrossRef]

Fayn, Kirill, Paul J. Silvia, Carolyn MacCann, and Niko Tiliopoulos. 2017. Interested in Different Things or in Different Ways? Exploring the Engagement Distinction Between Openness and Intellect. *Journal of Individual Differences* 38: 265–73. [CrossRef]

Fayn, Kirill, Paul J. Silvia, Egon Dejonckheere, Stijn Verdonck, and Peter Kuppens. 2019. Confused or Curious? Openness/intellect Predicts More Positive Interest-Confusion Relations. *Journal of Personality and Social Psychology* 117: 1016–33. [CrossRef]

Forster, Michael, Gernot Gerger, and Helmut Leder. 2015. Everything's Relative? Relative Differences in Processing Fluency and the Effects on Liking. *PLoS ONE* 10: e0135944. [CrossRef]

Forster, Michael, Helmut Leder, and Ulrich Ansorge. 2013. It Felt Fluent, and I Liked It: Subjective Feeling of Fluency Rather Than Objective Fluency Determines Liking. *Emotion* 13: 280–89. [CrossRef]

Frazier, Lyn, and Janet Dean Fodor. 1978. The Sausage Machine: A New Two-Stage Parsing Model. *Cognition* 6: 291–325. [CrossRef]

Fries, Albert. 1904. Zu Heinrich von Kleists Stil. *Studien zur vergleichenden Literaturgeschichte* 4: 440–65.

Geldhof, G. John, Kristopher J. Preacher, and Michael J. Zyphur. 2014. Reliability Estimation in a Multilevel Confirmatory Factor Analysis Framework. *Psychological Methods* 19: 72–91. [CrossRef] [PubMed]

Gibson, Edward. 1998. Linguistic Complexity: Locality of Syntactic Dependencies. *Cognition* 68: 1–76. [CrossRef]

Giora, Rachel, Ofer Fein, Ann Kronrod, Idit Elnatan, Noa Shuval, and Adi Zur. 2004. Weapons of Mass Distraction: Optimal Innovation and Pleasure Ratings. *Metaphor and Symbol* 19: 115–41. [CrossRef]

Graf, Laura K. M., and Jan R. Landwehr. 2015. A Dual-Process Perspective on Fluency-Based Aesthetics: The Pleasure-Interest Model of Aesthetic Liking. *Personality and Social Psychology Review* 19: 395–410. [CrossRef]

Gumbrecht, Hans-Ulrich, and Friederike Knüpling, eds. 2004. *Kleist Revisited*. Paderborn: Fink. ISBN 9783846754689.

Güçlütürk, Yağmur, Richard H. A. H. Jacobs, and Rob van Lier. 2016. Liking Versus Complexity: Decomposing the Inverted U-Curve. *Frontiers in Human Neuroscience* 10: 112. [CrossRef]

Huron, David Brian. 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press. ISBN 9780262582780.

Jacobs, Arthur M. 2014. Towards a Neurocognitive Poetics Model of Literary Reading. In *Towards a Cognitive Neuroscience of Natural Language Use*. Edited by Roel M. Willems. Cambridge, MA: Cambridge University Press, pp. 135–95. ISBN 9781107323667.

Karlsson, Fred. 2007. Constraints on Multiple Center-Embedding of Clauses. *Journal of Linguistics* 43: 365–92. [CrossRef]

Klausnitzer, Ralf. 2012. "(die Schriftgelehrten mögen ihn erklären)": Zum Kommagebrauch des Heinrich von Kleist. In *Die Poesie der Zeichensetzung: Studien zur Stilistik der Interpunktion*. Edited by Alexander Nebrig and Carlos Spoerhase. Bern: Peter Lang, pp. 203–38. ISBN 9783034310000.

Knoop, Christine A., Valentin Wagner, Thomas Jacobsen, and Winfried Menninghaus. 2016. Mapping the Aesthetic Space of Literature From "Below". *Poetics* 56: 35–49. [CrossRef]

Larkin, Willard, and David Burns. 1977. Sentence Comprehension and Memory for Embedded Structure. *Memory & Cognition* 5: 17–22. [CrossRef]

Lee, Kwan Min. 2004. Presence, Explicated. *Communication Theory* 14: 27–50. [CrossRef]

Makuuchi, Michiru, Jörg Bahlmann, Alfred Anwander, and Angela D. Friederici. 2009. Segregating the Core Computational Faculty of Human Language From Working Memory. *Proceedings of the National Academy of Sciences* 106: 8362–67. [CrossRef]

McArdle, John J., Emilio Ferrer-Caja, Fumiaki Hamagami, and Richard W. Woodcock. 2002. Comparative Longitudinal Structural Analyses of the Growth and Decline of Multiple Intellectual Abilities Over the Life Span. *Developmental Psychology* 38: 115–42. [CrossRef]

Medford, Emma, and Sarah P. McGeown. 2012. The Influence of Personality Characteristics on Children's Intrinsic Reading Motivation. *Learning and Individual Differences* 22: 786–91. [CrossRef]

Menninghaus, Winfried. 2009. "Ein Gefühl der Beförderung des Lebens": Kants Reformulierung des Topos lebhafter Vorstellung. In *Vita aesthetica: Szenarien ästhetischer Lebendigkeit*. Edited by Armen Avanessian, Winfried Menninghaus and Jan Völker. Zürich/Berlin: Diaphanes, pp. 77–94. ISBN 9783037340752.

Menninghaus, Winfried, Isabel C. Bohrn, Christine A. Knoop, Sonja A. Kotz, Wolff. Schlotz, and Arthur M. Jacobs. 2015. Rhetorical Features Facilitate Prosodic Processing While Handicapping Ease of Semantic Comprehension. *Cognition* 143: 48–60. [CrossRef]

Menninghaus, Winfried, Isabel. C. Bohrn, Ulman Altmann, Oliver Lubrich, and Arthur M. Jacobs. 2014. Sounds Funny? Humor Effects of Phonological and Prosodic Figures of Speech. *Psychology of Aesthetics, Creativity, and the Arts* 8: 71–76. [CrossRef]

Menninghaus, Winfried, Valentin Wagner, Eugen Wassiliwizky, Thomas Jacobsen, and Christine A. Knoop. 2017. The Emotional and Aesthetic Powers of Parallelistic Diction. *Poetics* 63: 47–59. [CrossRef]

Menninghaus, Winfried, Valentin Wagner, Ines Schindler, Christine A. Knoop, Stefan Blohm, Klaus Frieler, and Mathias Scharinger. 2023. Parallelisms and Deviations: Two Fundamentals of an Aesthetics of Poetic Diction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 379: 20220424. [CrossRef] [PubMed]

Meyer, Leonard B. 1961. *Emotion and Meaning in Music*. Chicago: University of Chicago Press. ISBN 9780226521398.

Miall, David S., and Don Kuiken. 1994. Foregrounding, Defamiliarization, and Affect: Response to Literary Stories. *Poetics* 22: 389–407. [CrossRef]

Nebrig, Alexander, and Carlos Spoerhase. 2012. Für eine Stilistik der Interpunktion. In *Die Poesie der Zeichensetzung: Studien zur Stilistik der Interpunktion*. Edited by Alexander Nebrig and Carlos Spoerhase. Bern: Peter Lang, pp. 11–31. ISBN 9783034310000.

Obermeier, Christian, Sonja A. Kotz, Sarah Jessen, Tim Raettig, Martin von Koppenfels, and Winfried Menninghaus. 2016. Aesthetic Appreciation of Poetry Correlates With Ease of Processing in Event-Related Potentials. *Cognitive, Affective, & Behavioral Neuroscience* 16: 362–73. [CrossRef]

Paulhus, Delroy L., Richard W. Robins, Kali H. Trzesniewski, and Jessica L. Tracy. 2004. Two Replicable Suppressor Situations in Personality Research. *Multivariate Behavioral Research* 39: 303–28. [CrossRef] [PubMed]

Pickering, Martin J., and Roger P. G. van Gompel. 2006. Syntactic Parsing. In *Handbook of Psycholinguistics*, 3rd ed. Edited by Matthew J. Traxler and Morton A. Gernsbacher. Cambridge, MA: Academic Press, pp. 455–503. ISBN 9780123693747.

Reber, Rolf, Norbert Schwarz, and Piotr Winkielman. 2004. Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience? *Personality and Social Psychology Bulletin* 8: 364–82. [CrossRef]

Schlegel, Friedrich. 2001. *On the Study of Greek Poetry*. Translated by Stuart Barnett. Albany: SUNY Press. ISBN 9780791448298.

Sembdner, Helmut. 1994. Kleists Interpunktion. In *In Sachen Kleist, Beiträge zur Forschung*, 3rd expanded ed. Munich: Carl Hanser, pp. 149–75. ISBN 9783446176980.

Silvia, Paul J. 2005. What is Interesting? Exploring the Appraisal Structure of Interest. *Emotion* 5: 89–102. [CrossRef]

Silvia, Paul J. 2007. Knowledge-based Assessment of Expertise in the Arts: Exploring Aesthetic Fluency. *Psychology of Aesthetics, Creativity, and the Arts* 1: 247–49. [CrossRef]

Silvia, Paul J. 2010. Confusion and Interest: The Role of Knowledge Emotions in Aesthetic Experience. *Psychology of Aesthetics, Creativity, and the Arts* 4: 75–80. [CrossRef]

Silvia, Paul J. 2013. Interested Experts, Confused Novices: Art Expertise and the Knowledge Emotions. *Empirical Studies of the Arts* 31: 107–15. [CrossRef]

Silvia, Paul J., and Christopher Berg. 2011. Finding Movies Interesting: How Appraisals and Expertise Influence the Aesthetic Experience of Film. *Empirical Studies of the Arts* 29: 73–88. [CrossRef]

Silvia, Paul J., Robert A. Henson, and Jonathan L. Templin. 2009. Are the Sources of Interest the Same for Everyone? Using Multilevel Mixture Models to Explore Individual Differences in Appraisal Structures. *Cognition & Emotion* 23: 1389–406. [CrossRef]

Stenzel, Jürgen. 1966. *Zeichensetzung: Stiluntersuchungen an deutscher Prosadichtung*. Palaestra 241. Göttingen: Vandenhoeck & Ruprecht.

Stine-Morrow, Elizabeth A. L., Matthew C. Shake, Joseph R. Miles, Kenton Lee, Xuefei Gao, and George McConkie. 2010. Pay Now or Pay Later: Aging and the Role of Boundary Salience in Self-Regulation of Conceptual Integration in Sentence Processing. *Psychology and Aging* 25: 168–76. [CrossRef]

Traxler, Matthew J., Liv J. Hoversten, and Trevor A. Brothers. 2018. Sentence Processing and Interpretation in Monolinguals and Bilinguals: Classical and Contemporary Approaches. In *The Handbook of Psycholinguistics*. Edited by Eva M. Fernández and Helen Smith Cairns. Hoboken: Wiley Blackwell, pp. 320–44. [CrossRef]

von Hellingrath, N. 1911. *Pindar-Übertragungen von Hölderlin: Prolegomena zu einer Erstausgabe*. Munich: Diederichs.

von Kleist, Heinrich. 1989. *Sämtliche Werke. Berliner Ausgabe. Die Marquise von O…*. Edited by Roland Reuß and Peter Staengle. Frankfurt/Basel: Stroemfeld/Roter Stern, vol. II/2, ISBN 9783878773481.

von Kleist, Heinrich. 1990. *Sämtliche Werke. Berliner Ausgabe. Michael Kohlhass (1808)*. Edited by Roland Reuß and Peter Staengle. Frankfurt/Basel: Stroemfeld/Roter Stern, vol. II/1, ISBN 9783878773481.

von Kleist, Heinrich. 1993. *Sämtliche Werke. Brandenburger Ausgabe. Das Erdbeben in Chili*. Edited by Roland Reuß and Peter Staengle. Frankfurt/Basel: Stroemfeld/Roter Stern, vol. II/3, ISBN 9783878773504.

von Kleist, Heinrich. 1994. *Sämtliche Werke. Brandenburger Ausgabe. Der Zweikampf*. Edited by Roland Reuß and Peter Staengle. Frankfurt/Basel: Stroemfeld/Roter Stern, vol. II/6, ISBN 9783878773566.

von Kleist, Heinrich. 1997. *Sämtliche Werke. Brandenburger Ausgabe. Das Bettelweib von Locarno; Der Findling; Die heilige Cäcilie, oder, Die Gewalt der Musik*. Edited by Roland Reuß and Peter Staengle. Frankfurt/Basel: Stroemfeld/Roter Stern, vol. II/5, ISBN 9783878773542.

von Kleist, Heinrich. 2019. Michael Kohlhaas. In *Tales From the German, Comprising Specimens From the Most Celebrated Authors*. Edited by John Oxenford and C. A. Feiling. Translated by John Oxenford. Sydney: Wentworth Press, pp. 165–230. ISBN 978-1103095469.

Wallot, Sebastian, and Winfried Menninghaus. 2018. Ambiguity Effects of Rhyme and Meter. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 44: 1947–54. [CrossRef]

Westfall, Jacob, David A. Kenny, and Charles M. Judd. 2014. Statistical Power and Optimal Design in Experiments in Which Samples of Participants Respond to Samples of Stimuli. *Journal of Experimental Psychology: General* 143: 2020–45. [CrossRef]

Winkielman, Piotr, and John T. Cacioppo. 2001. Mind at Ease Puts a Smile on the Face: Psychophysiological Evidence That Processing Facilitation Leads to Positive Affect. *Journal of Personality and Social Psychology* 81: 989–1000. [CrossRef] [PubMed]

Yngve, Victor H. 1960. A Model and an Hypothesis for Language Structure. *Proceedings of the American Philosophical Society* 104: 444–66. Available online: https://www.jstor.org/stable/985230 (accessed on 9 July 2025).

Zillmann, Dolf. 1996. The Psychology of Suspense in Dramatic Exposition. In *Suspense: Conceptualizations, Theoretical Analyses, and Empirical Explorations*. Edited by Peter Vorderer, Hans Jürgen Wulff and Mike Friedrichsen. Hillsdale: Erlbaum, pp. 199–231. ISBN 9780805819663.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.