Psychology of entertainment
van Driel, K.

Citation for published version (APA):
In this chapter, the hypothesis is studied that esthetic judgments of art perceivers is a cognitive process, using the Cognitive Processing of Cultural Products model. In this model it is assumed that esthetic judgments are dependent on the strength of cognitive processes in relevant domains of perceiver’s consciousness. Four cognitive processes are identified as most important in esthetic judgments: encoding, comparison, modification and performance. These processes will be executed in three domains of consciousness: the rational, affective and imaginary. The 3x4 facet model is tested in the area of esthetic judgments of literary texts, i.e., books. The facets of the model are measured by using a structured questionnaire. Stronger cognitive processes during reading will be correlated with higher esthetic judgments of a literary text. In the first study, the questionnaire was completed by 215 psychology freshmen. A confirmatory factor analysis showed that the model could indeed be fitted (gfi = 0.98; agfi = 0.92). In the second study two groups (A and B) filled in the questionnaire with different instructions. Group A (N=49) was instructed to focus on a ‘good’ recently read book and group B (N=37) was instructed to focus on a ‘bad’ book. Perceived quality of the book was correlated with more intense cognitive activities of the reader’s consciousness.
INTRODUCTION

Why do I like this book? Why do I like this play?’ In the perception and experiencing of cultural products, these questions on esthetics are quintessential. Esthetic judgments by, for instance, art critics or literary scientists are usually stated in terms of ‘good’ or ‘bad’, ‘beautiful’ or ‘ugly’, of ‘Apollo’ versus ‘Dionysos’ (Nietzsche, 1871/1994, 20, 35, 99). However, a simple scientific demand, stated by Van het Reve (1979, 135), is that the description of a thing only describes that one thing and no other. In literary studies, he continues, this prerequisite is not so obvious: How to describe a good book - without using the word ‘good’ - such that the description does not relate to a worthless book and how to describe a bad book - not using the word ‘bad’ or similar words - such that good books are not described by it? Taking up Van het Reve’s challenge, this paper envisions esthetics not in terms of good or bad, but in terms of different levels of cognitive activity during processing.

A widespread view in both psychology and the arts is that readers and spectators look for consistency, whereas the artist strives to deregulate that process. At an optimum of resolved frustration, the art observer would find most satisfaction. Early examples of this assumption are found in Sklovskij (1965, 12), Striedter (1989, 23-24) and Mukarovsky (1964, 17-30). In line with Gestalt principles, Iser (1978) argued that readers fill in the blanks (‘Leerstellen’) if the good continuation of a story is deliberately hampered by artistic technique. Based on Sklovskij, Berlyne (1974) introduced a model, in which hedonic value and esthetic appreciation is predicted by the evoked arousal of artistic novelty. Following Berlyne (1974), Cupchik (1995; 1996, 189-197) argues that suspense, novelty, inconsistency and uncertainty urge readers and spectators to ‘deeper processing,’ thereby enhancing esthetic pleasure.

“The close linkage between stimulus features and response elements is very behavioral and is therefore subject to mechanisms such as stimulus generalization and satiation which govern such processes” (Cupchik & Leonard, 1997, 88-89).

Yet, these authors are mainly concerned with the input and output of a black
box, which leaves many questions unanswered. Foremost, what is inside the black box? And, in particular, which processes determine the output, evoked by surprise, suspense or novelty? Which of these 'deeper processes' and domains of human thinking are involved in judgments such as 'beautiful piece' and 'bad production'? How can these processes and domains be identified and measured?

The cognitive tradition in modern psychology may provide us with some answers to these questions. The Swiss Gestalt-psychologist Meili (1981) represents cognitive processes using four factors: analyzing information (Plastizität), restructuring information (Komplexität), synthesizing information (Globalisation), and performing (Flussigkeit). Following a similar scheme, Sternberg's (1982; 1985) componential structure of intelligence distinguishes processes such as meta-, performance-, and knowledge acquisition-components. In addition, in the social cultural context, Sternberg distinguishes three main processes: selective coding, selective combining and selective comparing. Implementing these theoretical descriptions to the hypothetical cognitive activities during the formation of esthetic judgments of art, the most striking cognitive processes are: encoding incoming information, retrieving relevant information from memory, modification of information and performance to acquire deep level analyses of the object. Esthetic judgments of art are comprised of conscious cognitive activity.

Cognitive psychologists tend to demarcate at least two domains of consciousness, especially in relation to (social) intelligence: a cognitive and an affective domain (Bermond, 1998). In the study of evaluation of objects, three domains are distinguished: cognitive, affective and behavioral (Kerlinger, 1984). In case of esthetic judgments a virtual behavioral domain, like fantasy or the imaginary, is more prominent than a real behavioral domain. So, we propose that in the case of esthetic judgments three domains (cognitive, affective and imaginary) play a part in the processing of cultural products and, consequently, the development of esthetic judgments.

The present study tries to specify these deeper processes from the black box, with the Cognitive Processing of Cultural Products model (or: the CPCP) to be presented here. This cognitive approach may fit in with some established assumptions from empirical esthetics. The CPCP assumes that observers reconstitute...
the work of art (cf. Boselie, 1979), which is necessary to fill in the blanks (cf. Iser, 1978). Since the amount of cognitive activation in perceiving art is likely to be positively related to esthetic pleasure, the measure of activation may function as a predictor of the esthetic experience of the art perceiver and, consequently, the quality of the work of art.

The remainder of the introduction will address the CPCP-model, followed by a short overview of questionnaire construction using the facet method. Two empirical studies were performed to test the model. The first study concerns the construction of a questionnaire for cognitive processing during reading and a large-scale validation study, to establish its reliability and validity. The second study focuses on testing specific hypotheses about the effect of the CPCP on the esthetic appreciation of books. Two specific hypotheses are derived from the hypothesis that perceived quality of a book will increase with more intense cognitive activity of the reader’s consciousness. In addition, three specific hypotheses are derived from the presumption that the distinction, made between the different domains of human thinking, may be relevant to the genre of the literary work.

**The Cognitive Processing Of Cultural Products Model**

Starting off with the computer-brain metaphor (Dennett, 1991), memory-structures (ROM), cognitive processes (CPU) and domains (RAM) may be distinguished. Therefore, the blanks that should be filled in by art observers may lie in the rational, affective and imaginary domains of human thinking. The cognitive processes involved may be encoding, retrieval, comparison and modification. Successful execution of these processes probably is dependent on the observer’s performance level, which in itself may be considered a process.

The rational domain involves logical inferencing, abstract thinking, literal and ‘scientific’ description, and mathematics. The affective domain refers to emotional states and affairs, intuition and morals, while the imaginary domain contains figurative, metaphorical aspects (e.g. Arnheim, 1971; Hoorn, 1997-a, 67-74; Hoorn, 1997-b, 20-23) and more notably fantasy and other not empirically evident appearances.

During encoding, an element in a work of art is consciously perceived. Such
an art-element may be a fictional character, a musical theme, a metaphor, a brush stroke, a happy ending, as well as a more complex idea or theory. A fictional character, for instance, has different features, which are sampled while reading or watching. It has a nose, hair, and a hunchback (rational domain), is nauseating (affective domain) and rides a unicorn (imaginary domain). The features in the combined domains form the feature set of the art-elements.

Upon encoding of the art-element, the retrieval process searches by association for memory-concepts of previously stored information associated with the art-element, including more general knowledge, ideas, theories, thoughts and personal memories. Its feature set includes the rational, affective and imaginary concepts that are ‘usually right’ or ‘right up till now’ about a character or happy ending.

Comparison of the two feature sets, i.e., the encoded art-elements and the retrieved memory-concepts, results in the intersection between both sets and two sets of distinctive features, the unique art-elements and the aberrant memory-concepts. The ratio of the intersection and the distinctive sets determines the coherence between the two.

If enough shared features are found, there is no need to modify earlier assumptions on, for instance, the story line. If too many distinctive features are found, disturbance urges modification of the feature sets of the art-elements and/or the memory-concept. Modification, therefore, reduces disturbance and increases coherence. Both qualities are important for the appreciation of the work, and depend on the individual’s own level of tolerance. A book that offers nothing new evokes too little disturbance: No puzzle to be solved (cf. Sternberg, 1988). A book that merely frustrates, shows too little coherence, will be condemned to the shelf of utter chaos (cf. Konijn & Hoorn, 1998).

Performance is an overall process, which relates to the successfulness of the recomposition of the work of art. Festinger’s (1957) cognitive dissonance theory claims that effort induces liking. Yet, there is an optimum to this relationship. When a book is too easy or too difficult, cognitive activation may be inhibited: in the first case because of boredom, in the second case because of incomprehensibility. Of course, individual abilities may affect the level of performance; thus, what presents a
Psychology of Entertainment

difficulty to one person may well be easy for another. For instance, listeners to a modern piece of music may not be able to sufficiently modify their ideas about music to appreciate the ‘foreign’ atonalities of the piece. Nevertheless, the paramount assumption of the CPCP model is that using more processes in more domains positively affects appreciation.

By measuring the activity of these processes, a no-value judgement of the work of art that evoked them, may be obtained.

**Questionnaire Construction**

**The Facet Method**

The ‘Questionnaire Cognitive Processing during Reading’, or the QCPR was constructed with the facet method (Canter, 1985; Borg, 1979). Oosterveld (1996) provides a comprehensive description of the facet method. Its aim is to optimize content validity with a systematic and, ideally, exhaustive specification of the concept. Its focus is the concept analysis (Oosterveld & Vorst, 1996), which describes the concept domain. Concept analysis consists of the following four steps. First, an inventory is made of behavioral aspects and underlying processes of the concept. Second, the essential aspects of the concept are defined as facets. Third, the elements of the facets are determined in such a way that they describe mutually exclusive categories within the facet, while the facet elements should be independent between the facets. Fourth, the final structure of the facet design is determined by a representation of the facets, their elements and the relations between the facets in a so-called mapping sentence. The mapping sentence is the verbal expression of the facet design and produces the different specific descriptions of the concept. Every single specific description (structuple) results from the combination of one facet element with one of the elements of each of the other. According to the specified mapping sentence, a number of items is written per structuple.

To design the QCPR, a concept analysis (Sloman, 1978) was performed, inventoring aspects related to beauty, cognition, perception and emotion. Three facets were defined: cognitive processes, domains of human thinking and memory
structures. Thereupon, encoding, comparison, modification and performance were determined as the elements of the process facet, whereas three domains were established, the rational, the affective and the imaginary. These first two facets form the CPCP model described in the previous paragraph and the (main) facet design (Figure 3.1).

**Figure 3.1  Facet design for cognitive processing**

<table>
<thead>
<tr>
<th>DOMAIN FACET</th>
<th>PROCESS FACET</th>
<th>Encoding</th>
<th>Comparison</th>
<th>Modification</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rational</td>
<td>RE</td>
<td>RC</td>
<td>RM</td>
<td>RP</td>
<td>scale 1</td>
</tr>
<tr>
<td>Affective</td>
<td>AE</td>
<td>AC</td>
<td>AM</td>
<td>AP</td>
<td>scale 2</td>
</tr>
<tr>
<td>Imaginary</td>
<td>IE</td>
<td>IC</td>
<td>IM</td>
<td>IP</td>
<td>scale 3</td>
</tr>
</tbody>
</table>

Note: Three DOMAINS are crossed with four processes. For the twelve resulting structuples (cells), items should broach the respective mental state.

Contrary to the full facet design, used for item production, the main facet design only describes the scales universal to the arts; on this level statistical analyses will be performed. On the other hand, the structures facet, being specific to each of the arts, covers literature in the present case. Consequently, this facet contains structural and stylistic principles that in Western society are considered elementary for the analysis of a novel. They were: Plot, Characters, Surroundings/Place, Events/Facts, Dialog, Theme, Time, Style and Miscellaneous (among others: Humor, Originality, Norms, Symbolism). Finally, the (full) facet design is presented in Figure 3.2, by the representation of the facets, their elements, and the relationship between the facets in the mapping sentence. Since items are directly derived from this mapping sentence, it also contains the structures facet.

Figure 3.1 shows the facet design for the QCPR. The four elements of the process facet are crossed with the three elements of the domain facet, resulting in 12 structuples. Each structuple reflects a specific mental state, in which the novel reader may be busy encoding affective features (AE) or modifying imaginary features (IM). By calculating the sum of rows and columns, seven scales for cognitive processing
can be created.

Note, that the scale for the memory retrieval process has not been included in the facet design. The retrieval of one's own personal memory or idea has no direct relation to the work of art like the other processes, although it certainly is evoked by the esthetic work and may also be enjoyed. Nevertheless, the retrieval process was implicitly measured in comparison and modification. Finally, it is important to note, that retrieval is a process element that is not to be confused with the memory structures, which form their own facet.

**Figure 3.2** *Mapping sentence or schematic model for a Questionnaire on the Cognitive Processing during Reading (QCPR)*

<table>
<thead>
<tr>
<th>Plot</th>
<th>Characters</th>
<th>activates (n)</th>
<th>Encoding</th>
<th>Comparision</th>
<th>Rational process in the Affective Imaginary domain,</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events/Facts</td>
<td>Dialog</td>
<td>Theme</td>
<td>Surroundings</td>
<td>Time</td>
<td>Style</td>
</tr>
</tbody>
</table>

which leads the reader to respond:

- "I disagree entirely" 1
- "I disagree" 2
- "I agree nor disagree" 3
- "I agree" 4
- "I agree entirely" 5

**Item Development**

For each of the twelve structuples, eight items for self-evaluation were written. Half the items were indicative (‘The tone of the dialogues made it easier to empathize with the characters’) and half contra-indicative (‘I noticed that the emotions of the main character hardly affected me’). These were followed by a 5-point scale (Figure 3.1). Items were written according to the mapping sentence (Figure 3.2). The resulting 96 items were tested in a pilot study among 51 A-level college students (age 16-18). The students completed the questionnaire and were instructed to criticize the items on word choice and content. Items were maintained that (1) contributed optimally to the highest possible Cronbach’s α for their structuple and the two related scales, (2) had a high correlation with their own scales.
and a substantially lower correlation with all other scales, (3) made an optimal contribution to the structural content in line with the facet design and (4) were not criticized on comprehensibility. Items that failed on one of these criteria were adjusted (20 items) or discarded and replaced (40 items).

As mentioned before, the remainder of this study consists of two parts. Part 1 is devoted to the construction and validation of the QCPR, whereas Part 2 focuses on what the QCPR was made for: measuring changing levels of cognitive activity and establishing the relation with the quality of the book.

**PART 1: CONSTRUCTION OF THE QCPR**

**Method**

The validation of the QCPR aimed for three goals: Test shortening, establishing scale reliability, and establishing construct validity of scales. The QCPR was administered to a large sample of students, supplemented by cognitive and personality measures for validation purposes.

**Sample**

The questionnaire was administered to 318 psychology freshmen. Subjects were randomly assigned to one of three conditions. Subjects in the first condition ($N=49$) filled in the questionnaire, while thinking of a good book they had read recently. The second group ($N=37$) did the same, while thinking of a bad book they had read recently. In the third condition ($N=215$) subjects answered the questionnaire while thinking of 'a literary book', irrespective of whether they considered it to be good or bad. Seventeen subjects were not considered in the analysis, because they reported they never had read a literary book. The validation study presented next is based on the data of the third condition ($N=215$), while the study on esthetic judgments (Part 2) utilizes the data of the first two conditions.

**Additional measures**

Four tests of intellectual abilities were applied, representing different
intellectual skills from Guilford's structure of intellect model (Elshout, 1976; Carroll, 1993). In addition, three personality inventories were administered, 1) the Dutch Personality Questionnaire (Luteijn, Starren & Van Dijk, 1975), 2) the Amsterdam Biographical Questionnaire (De Wilde, 1963/1970), and 3) the Bermond-Vorst Alexithymia Questionnaire (Bermond, Vorst & Oosterveld, 1992). Combined, the three personality inventories consisted of 16 scales, covering a wide variety of personality traits.

**Statistical analysis**

To shorten the test, two items were removed from each structuple according to the concept of content saturation (Jackson, 1971), i.e. items should show a high correlation with the intended scale and a low or at least a lower correlation with all other scales. Of each structuple, one item was removed from among the indicative items and one from among the contra-indicative items.

To establish the reliability of the scales, that is, the extent to which scale scores are subject to error, Cronbach's $\alpha$ was used. Scales with a low reliability do not measure a construct with enough precision. Nunnally (1967) considers a value of 0.70 as the lower bound for sufficient reliability.

To establish the construct validity of the scales, hypotheses of convergent and discriminant validities were tested. Two approaches were taken, a factor-analytical approach, which involves the convergent, and discriminant properties of the scores within the instrument and a correlational approach that focuses on the relations with external measures.

Concerning the first approach, Mellenbergh, Kelderman, Stijlen and Zondag (1979) showed that in the case of a facet-designed questionnaire, convergent and discriminant validity can be established by fitting a (confirmatory) factor model. The hypothesis underpinning such a model is that the responses can be described by a limited number of underlying factors. In a facet design, the number of factors equals the number of facet elements: three domain factors and four process factors. In the model, the structuples load on their constituent facet-element factors. In the present case, therefore, each structuple has two factor loadings: one on its process factor and one on its domain factor. Regarding the relations among factors, it is assumed that
the factors do not correlate between the facets, while they are allowed to correlate within the facets.

The acceptability of the model was tested with a $\chi^2$ test. This test should result in a non-significant $p$-value, since significance indicates misfit. In large samples, however, the $\chi^2$ test may have too much power and models with trivial misfit may be rejected falsely. Therefore, fit measures are available that are less affected by sample size, such as the goodness of fit index (GFI) and the adjusted goodness of fit index (AGFI). Both these measures should exceed 0.90 (Marsh, Balla, & McDonald, 1988). On a more detailed level, the factor loadings provide information on the strength of the relation between the structuples and the underlying factors. A low factor loading ($< 0.30$) indicates that the structuple does not sufficiently capture the supposed concept, giving detailed information on where to improve the questionnaire if necessary.

Regarding the second approach, it could be argued that two main sources of variance (the books and the readers) are present in the QCPR scores. This would implicate that the QCPR may be measuring reader characteristics as well. Since reader characteristics must be viewed as a confounding influence, the QCPR scores must be independent of the measures of these characteristics, i.e., personality and cognitive tests, in order to claim construct validity. More specifically, the QCPR scores must show discriminant validity with respect to these external measures. The influence of reader characteristics was examined by means of correlations between the scales of the QCPR and the cognitive and personality tests. Correlations smaller than 0.30 were considered adequate for discriminant validity.

Results

Scale reliabilities were calculated to remove two items from every structuple. Scale reliabilities of the reduced scales are presented in Table 3.1. The values of reliability all are well above the 0.70 boundary, and except for the 0.74 of rational, all values are even above 0.80.

Since subjects may have interpreted encoding ('While reading I saw...') as some sort of performance, the CPCP was fitted with encoding and performance combined into
Table 3.1  Reliabilities (Cronbach’s \( \alpha \)), means and standard deviations of the final scales of the QCPR

<table>
<thead>
<tr>
<th></th>
<th>Cronbach’s ( \alpha )</th>
<th>Mean</th>
<th>SD</th>
<th># Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rational</td>
<td>0.74</td>
<td>84.34</td>
<td>11.11</td>
<td>24</td>
</tr>
<tr>
<td>Affective</td>
<td>0.90</td>
<td>81.36</td>
<td>18.12</td>
<td>24</td>
</tr>
<tr>
<td>Imaginary</td>
<td>0.90</td>
<td>82.81</td>
<td>16.07</td>
<td>24</td>
</tr>
<tr>
<td>Encoding</td>
<td>0.86</td>
<td>65.81</td>
<td>12.47</td>
<td>18</td>
</tr>
<tr>
<td>Comparison</td>
<td>0.86</td>
<td>53.87</td>
<td>12.53</td>
<td>18</td>
</tr>
<tr>
<td>Modification</td>
<td>0.83</td>
<td>57.20</td>
<td>11.64</td>
<td>18</td>
</tr>
<tr>
<td>Performance</td>
<td>0.85</td>
<td>71.62</td>
<td>11.47</td>
<td>18</td>
</tr>
</tbody>
</table>

Note  \( N=213 \).

one variable. Fit measures run as follows: \( \chi^2 (df=36)= 70.38; \quad p= 0.00; GFI= 0.95; AGFI= 0.89 \). In addition, a model was fitted skipping encoding altogether \( \chi^2 (df=13)=23.64; \quad p= 0.035; GFI= 0.98; AGFI= 0.92 \), and fixing the unique variance

Table 3.2  Model fit and parameter estimates of the facet design minus encoding

| Factor Loadings and Unique Variances |
|-------------------|---|---|---|---|---|---|---|
| Rat. | Aff. | Ima. | Com. | Mod. | Per. | unique var |
| RC   | 0.51 | --  | --  | 0.57 | --  | --  | 0.41 |
| RM   | 0.82 | --  | --  | --   | 0.27 | --  | 0.26 |
| RP   | 0.08 | --  | --  | --   | --   | 0.64 | 0.59 |
| AC   | --   | 0.68| --  | 0.70 | --   | --  | 0.06 |
| AM   | --   | 0.54| --  | --   | 0.84 | --  | --   |
| AP   | --   | 0.46| --  | --   | --   | 0.75 | 0.23 |
| IC   | --   | --  | 0.70| --   | 0.42 | --  | 0.34 |
| IM   | --   | --  | 0.76| --   | --   | 0.22 | --   |
| IP   | --   | --  | 0.65| --   | --   | --   | 0.31 |

| Factor Correlations |
|---------------------|---|---|---|---|
| Rational          | 1.00          |     |     |     |
| Affective         | 0.66          | 1.00|     |     |
| Imaginary         | 0.75          | 0.88| 1.00|     |
| Comparison        | --            | --  | 1.00|     |
| Modification      | --            | --  | 0.05| 1.00|
| Performance       | --            | --  | 0.16| 0.49| 1.00|

Note: \( N=215 \), low factor loadings are in bold face, - - denotes a parameter restricted to 0
Chapter 3: Cognition of Literature

of the AM structure to zero Factor loadings, factor correlations and unique variances for this analysis are shown in Table 3.2. Among the three low factor loadings (bold), the 0.08 for RP is the lowest.

The correlations between the QCPR scales on the one hand and the 4 cognitive and 16 personality scales on the other hand were acceptable from the point of view of discriminant validity (< 0.30).

**Conclusion**

In general, the results show favorable psychometric features for the QCPR. The reliability estimates are high and most aspects of convergent and discriminant validity are supported. Nevertheless, one less favorable aspect of the QCPR was uncovered by confirmatory factor analysis. The association between encoding and performance was so strong that these could not be viewed as separate constructs. However, eliminating encoding from this analysis resulted in a model with adequate fit. It warrants further investigation, whether the intertwining of encoding and performance is an aspect of the QCPR, e.g., due to the wording of the items, or that the two constructs are in fact theoretically similar.

Of 140 correlations (7 QCPR scales * 20 cognitive and personality scales), none attained a value higher than 0.30, which indicates that cognitive abilities and personality traits do not substantially affect the results. These correlations provide support for the construct validity of the QCPR, as being an instrument that measures the effect of books on readers, rather than reader characteristics. It is noteworthy, that the QCPR and the Bermond-Vorst Alexithymia Questionnaire used similar items for affective and imaginary, the latter not specifically for literature. Yet, correlations between these scales are low, supporting the construct validity of the QCPR.

**Part 2: Good vs. Bad**

Whereas Part 1 examined the reliability and validity of the QCPR, Part 2 addresses what the QCPR was made for: differentiating good from bad books by measuring changing levels of cognitive activity. We have three hypotheses.

Hypothesis 1: If quality differs per book, the scores on QCPR scales should
Psychology of Entertainment

be higher with increasing quality of the book.

Hypothesis 2: Appreciation of a book is associated with more intense processing by the reader.

Hypothesis 3: Stronger expression of particular genres will occur in different domains of human thinking. More specifically: this hypothesis is specified for three genres: detective, psychological and science fiction novels will activate stronger expression in the rational, affective and imaginary domains, respectively.

Furthermore, the effects of gender and age on the QCPR scores were investigated.

Method

Sample

In order to investigate the first hypothesis, the aforementioned sample of the two smaller subject groups was utilized, i.e., the ‘good’ book condition (N=49) and the ‘bad’ book condition (N=37). For examining the other hypotheses as well as the exploratory analyses, the ‘no instruction’ sample (N=215) was used.

Measures

Data on various additional variables were collected. The respondent’s appreciation of the literary work was assessed with an appreciation scale and a school grade rating of the book. The appreciation scale (Table 3.1) consisted of eight judgmental items, referring to different quality aspects (e.g., ‘I thought the characters were strikingly well described’), whereas the school grade was obtained by asking the readers to rate the quality of the book on a scale from 1 (low quality) to 10 (high). Finally, genre of the recalled book and gender and age of the respondents were assessed.

Statistical analysis

Scale means between the good and the bad condition were compared. To establish an overall significance level, first a multivariate analysis of variance
(MANOVA) was performed, with group membership as the independent variable and the scales of the QCPR as the dependent variables. The multivariate test includes differences between conditions on all the scale-means, as well as differences between all possible combinations of scales (interaction effects). Next, univariate analyses of variance (ANOVA) will be used to assess single scale-mean differences between conditions. On both the multivariate and the univariate tests significance indicates group differences. Moreover, effect sizes (ES) were calculated, expressing scale mean differences in terms of their standard deviation. Cohen (1977) considers ES = 0.20 a small, ES = 0.50 an average, and ES = 0.80 a large effect.

To test the second hypothesis, the relations between the appreciation scale and school grade on the one hand, and the scales of the QCPR on the other hand, were correlated. A correlation above 0.30 was considered sufficient support for the hypothesis.

To investigate the third hypothesis, overall differences of genres on the scales of the QCPR, a MANOVA was performed, with genre as independent variable and the scales of the QCPR as dependent variables. This was followed by univariate tests per scale and, in case of significance, post hoc analyses with Tukey’s HSD procedure to identify the genres causing the effect.

Finally, MANOVA was used for tests of both gender and age effects on the scales of the QCPR. The overall tests were followed by univariate tests.

**Results**

Table 3.3 presents the results pertaining to hypothesis 1. The effects of good vs. bad are significant (Hotelling’s $T_{6,79}=1.71$, $p=0.00$), both for the overall multivariate test as well as for each single scale. Moreover, the effect sizes between ‘good’ and ‘bad’ conditions were very high, ranging from 1.25 to 2.30.

Table 3.4 presents correlations between QCPR scales and appreciation score and school grade rating for the no-instruction condition only. Only the correlation between comparison and appreciation does not exceed the 0.30 mark.

The multivariate test with the scales of the QCPR depending on genre as independent variable (Hotelling’s $T_{54,980}=0.31; p=0.59$), showed no significant differences between genres for the specific scales.
Table 3.3  Multivariate test, means, standard deviations, univariate signif. and univariate effect sizes of the conditions ‘good’ vs. ‘bad’

<table>
<thead>
<tr>
<th>Scales</th>
<th>Bad mean</th>
<th>Bad SD</th>
<th>Good mean</th>
<th>Good SD</th>
<th>p</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rational</td>
<td>71.54</td>
<td>7.90</td>
<td>86.74</td>
<td>8.79</td>
<td>0.00</td>
<td>1.82</td>
</tr>
<tr>
<td>Affective</td>
<td>55.73</td>
<td>12.38</td>
<td>87.00</td>
<td>16.13</td>
<td>0.00</td>
<td>2.19</td>
</tr>
<tr>
<td>Imaginary</td>
<td>63.81</td>
<td>15.04</td>
<td>87.59</td>
<td>11.21</td>
<td>0.00</td>
<td>1.81</td>
</tr>
<tr>
<td>Encoding</td>
<td>48.38</td>
<td>9.86</td>
<td>69.16</td>
<td>9.14</td>
<td>0.00</td>
<td>2.19</td>
</tr>
<tr>
<td>Comparison</td>
<td>43.43</td>
<td>9.99</td>
<td>55.96</td>
<td>10.11</td>
<td>0.00</td>
<td>1.25</td>
</tr>
<tr>
<td>Modification</td>
<td>44.78</td>
<td>9.18</td>
<td>61.71</td>
<td>10.70</td>
<td>0.00</td>
<td>1.70</td>
</tr>
<tr>
<td>Performance</td>
<td>54.49</td>
<td>9.51</td>
<td>74.49</td>
<td>7.91</td>
<td>0.00</td>
<td>2.30</td>
</tr>
</tbody>
</table>

| N               | 37       | 49      |

No significant overall effects on the QCPR scales were found for age (Hotelling’s $T_{90,1112} = 0.46 ; p = 0.64$) and gender (Hotelling’s $T_{6,201} = 0.06 ; p = 0.05$). Univariate tests for gender per scale suggest, however, stronger effects for women than men for affective ($p = 0.03 ; ES = 0.33$) and performance ($p = 0.01 ; ES = 0.40$).

Table 3.4  Correlations between QCPR-scales and appreciation (App.) or school grade (Sg)

<table>
<thead>
<tr>
<th>PROCESS</th>
<th>App.</th>
<th>Sg</th>
<th>DOMAIN</th>
<th>App.</th>
<th>Sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encoding</td>
<td>0.67</td>
<td>0.59</td>
<td>Rational</td>
<td>0.46</td>
<td>0.38</td>
</tr>
<tr>
<td>Comparison</td>
<td><strong>0.26</strong></td>
<td>0.34</td>
<td>Affective</td>
<td>0.63</td>
<td>0.60</td>
</tr>
<tr>
<td>Modification</td>
<td>0.49</td>
<td>0.44</td>
<td>Imaginary</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>Performance</td>
<td>0.61</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:  $N=207$, low correlations are in bold face

Conclusion

In this second part of the study specific hypotheses pertaining to the evaluative power of the QCPR were tested. The first two hypotheses were supported, implicating that the QCPR is able to discriminate between ‘good’ and ‘bad’ books and that the intensity of the cognitive processes active while reading are related to the appreciation of the literary work.

There is no evidence for specific score profiles for genre. Concerning this result, it should be noted that the distribution of books over the ten reported genres
was quite unbalanced. Only for the psychological novel \((N= 56)\) and the modern novel \((N= 44)\), the effect could be estimated with sufficient accuracy. Thus, whether the processing of literary texts is genre independent remains open for empirical evaluation. Such an investigation should guarantee sufficient sample size for each of the genres.

Regarding gender differences, not only do women report to be more affected by literature, but also, they seem to be the better performing readers. Age, does not significantly affect the results.

**DISCUSSION**

Regarding the CPCP structure of the QCPR, two versions of the model were fitted. One combined encoding and performance into one variable, the other discarded encoding altogether from the model. High measures of fit underscore that the CPCP can largely be recovered from the QCPR responses. However, encoding and performance remain insufficiently mutually exclusive. It could be that encoding an art-element (i.e. books) is highly correlated with an optimal performance, or that the QCPR or the subjects failed to make the distinction. Thorough semantic analysis of the scales encoding and performance is needed to improve their accuracy in the future.

Nevertheless, it may be objected that encoding and performance are closely related due to semantic overlap in item wording rather than triggering similar processes. A counterargument could be that the elements of the structures facet, i.e., structural and stylistic principles, are equally distributed across the respective scales. If subjects only judge word meanings, this structures facet would become a large source of variance in the QCPR scales, and the processes and domains facets would have a weaker presence in the responses.

Another counterargument is that the QCPR and the BVAQ for alexithymia used similar items for affective and imaginary. Yet, these scales correlate so low, that the hypothesis that the high correlation between encoding and performance is due to semantic overlap, is unlikely. In addition, the arguments against a halo-interpretation set out further below are also applicable here.

Two hypotheses were derived from the CPCP. Hypothesis 1 states that a book
with higher quality should increase the amount of cognitive processing, while hypothesis 2 claims that appreciation increases with the intensity of the cognitive processes. Strong evidence was found for the first hypothesis, that readers who recalled ‘good’ books show significantly higher scores on the cognitive-processing scales than those who read ‘bad’ books. For the second hypothesis, sufficiently high correlations were found between the appreciation scale and the school grade for quality on the one hand and the QCPR scales on the other hand. However, the comparison of an art-element (i.e. in books) with its memory concept does not seem to affect the esthetic quality of the work.

An interpretation of these results in terms of a halo-effect would imply that on the whole, readers judge ‘good’ books higher on the QCPR scales than ‘bad’ books. Yet, such a halo-interpretation would not account for differences in correlations, which are found between the QCPR scales and appreciation and school grade, nor differences in effect sizes; the difference between lowest and highest effect size is more than 1.00. Nevertheless, more support for the CPCP may be found if subjects are randomly assigned to books in an experiment.

How does the CPCP model relate to the field of empirical esthetics? Assumptions of theorists such as Sklovskij (1917/1965, 12), Striedter (1989, 23-24), Mukarovsky (1964, 17-30) and Iser (1978), that modifying concepts enhances reader appreciation, are given concrete form in the CPCP and are operationalized by means of the QCPR.

Likewise, Berlyne’s (1974) idea, that disturbance (e.g., by artistic novelty) urges readers to modify earlier concepts which may increase hedonic value, is also represented in the CPCP. However, the model clearly shows, that this is not due to the mere output of a black box (undifferentiated arousal). There is something in there and it is called cognitive processing.

There is also support for notions from cognitive psychology (cf. Kerlinger, 1984). Different domains are involved in human reasoning. In understanding the story line or a controversial character, the rational domain is called upon. The affective domain is activated where emotion is involved, which is in line with Bermondd (1998) and with Konijn (1998) who, on an operant level, found that readers and spectators show task emotions, which contribute significantly to the appreciation.
Finally, the imaginary domain is also needed to recompose, for example, fantasy-figures and additional story-settings where description is minimal.

Furthermore, Cupchik’s (1995; 1996, 189-197) notions of ‘deeper processing’ when readers or spectators encounter an art-element that causes suspense, novelty, inconsistency or uncertainty, are specified by different cognitive processes, such as encoding, comparison, modification and performance. These findings support assumptions from cognitive psychology, that different cognitive processes may be distinguished in human reasoning (Meili, 1981; Sternberg, 1985).

Van het Reve (1979) outlined a gloomy future for literary studies. The scientific prerequisite that the description of a thing only describes that one thing and no other, was not so obvious in the literary studies. However, the QCPR meets Van het Reve’s demand by describing books, without using judgmental terms such as ‘good’, ‘bad’ or similar words, but in terms of cognitive activity.
Psychology of Entertainment

REFERENCES


Chapter 3: Cognition of Literature


Psychology of Entertainment


