

Three Obstacles in Empirical Research on Aesthetic and Literary Comprehension

**Arthur C. Graesser
Natalie Person
G. Scott Johnston**

Who would ever have imagined, 20 years ago, that some pressing theoretical debates about aesthetics, literary comprehension, and art could be partially answered by adopting the scientific method and collecting empirical data? Most philosophers of aesthetics would consider the idea preposterous. Scholars in literary criticism would snicker, sneer, or become profoundly irritated, but times have changed, and the contributions in this volume are a testimony to this change. We are finally emancipated from some jaded dogmas in philosophy and literary criticism.

There is now an established international society, called IGEL, that focuses on the empirical studies of literature, art, and media. This new society has attracted the attention of sci-

entists and scholars from psychology, sociology, communication, art, literature, philosophy, text linguistics, education, and several other fields. Cognitive psychologists collect verbal protocols, ratings on psychological scales, comprehension times, and other behavioral measures, whereas readers comprehend literature or appreciate a work of art. Researchers in marketing and sociology examine which films, books, or records people purchase and how the business market functions for these aesthetic products. Philosophers discuss how empirical data have relevance to traditional problems in epistemology, ontology, and aesthetics. Educators investigate the pedagogical and ethical consequences of exposing particular works of art and literature to students in the classroom. The new interdisciplinary society is growing at a healthy rate.

Nevertheless, the budding field will probably not have an easy journey in the future. There have not been a sufficient number of breakthroughs, revolutionary insights, and sophisticated theories for the field to progress at a comfortable pace, let alone to rest on its laurels. The scientists and scholars will need to continue their self-conscious creative struggle. Moreover, there are some persistent obstacles that appear to be inherent in the empirical studies of art and literature—this chapter identifies three of them. Our hope is that researchers will effectively confront them in the future.

OBSTACLE 1: THERE IS NO SINGLE ESSENCE TO THE QUALITY OF AESTHETIC OBJECTS AND TO AESTHETIC EXPERIENCES

Our research activities would be greatly simplified if there were a single essence to aesthetic quality. The life of a researcher would be easier if, for example, beautiful art objects had a set of core defining properties that were necessary and jointly sufficient to distinguish good art from bad art. However, the world of aesthetics is scruffy rather than tidy. There is no single essence that is shared by all instances of good literature, good paintings, good theater, and so on. As a consequence, there is widespread disagreement about the critical properties of aesthetic quality among writers, artists, critics, philosophers, scientists, and other experts. Disagreements are prevalent whenever experts attempt to define concepts that are inherently abstract, complex, and multifaceted.

Aesthetic quality appears to be an umbrella term for a cluster of components and mechanisms that are associated with successful aesthetic experiences. A precise specification of a successful aesthetic experience is intentionally circumvented at this point. However, we do want to emphasize that our focus is on the aesthetic *experience* (i.e., the process of a person comprehending and appreciating a work of art or literature) rather than the aesthetic *object* per se. We have identified seven aesthetic components, but we would not be surprised if more components surface in future research. Presumably, an aesthetic experience is increasingly more pleasing or enriching to the extent that more of these component mechanisms are intact during the comprehension of the aesthetic object. These seven component mechanisms are briefly described in the following sections.

Good Form

A good aesthetic object exhibits balance, symmetry, and other attributes of good form. For example, a story plot has good form when all of the good characters enjoy positive outcomes and all of the bad characters face negative outcomes at the end of the story. Circles and squares have good form compared to irregular shapes. A golden rectangle, which contains roughly an 8 to 5 ratio for sides, has a good form mathematically and is more aesthetically pleasing than rectangles with other ratios (Fechner, 1876). The central prediction of this first component is that aesthetic experiences are more pleasing (and less arousing) to the extent that the art object exhibits good form.

There have been some attempts to identify the mathematical properties of good form (Arnheim, 1971; Garner, 1970; Reed, 1972). Good form can be defined and mathematically measured on the basis of properties of the stimulus; the comprehender's background knowledge and culture have no importance. For example, one way of defining perceptual good form is that the same pattern persists (or a small number of patterns exist) when an object undergoes spatial transformations, such as rotation and reversal (Garner, 1970). That is, good form decreases as a function of the greater number of unique patterns that emerge when transformations are applied. A circle remains a circle when it is rotated and reversed, whereas a triangle generates new patterns when transformed. Therefore, a circle has

better form than a triangle. Categories of triangles would be ordered as follows on the dimension of good form: equilateral > isosceles > irregular. This computational definition of good form is formulated on the basis of the intrinsic properties of the stimulus, not the world knowledge and culture of the perceiver.

Prototypes

A prototypical concept has the typical attributes and the average values of variables when considering the set of exemplars for that concept (Rosch & Mervis, 1975). For example, if we were to take a random set of photographs of faces in a culture, the prototypical face would have the most common eye color, the average nose size, and so on. A particular face is increasingly less prototypical to the extent that it has more attributes and values that deviate from the prototype. There is some evidence that attractiveness ratings are higher for prototypical faces than for faces that deviate from the prototype (Langlois & Roggman, 1990; Light, Hollander, & Kayra-Stuart, 1981). In the context of literature, a prototypical plot would frequently appear in the narratives written in the culture, such as a story in which a hero rescues a victim from a villain. According to this prototype component, aesthetic experiences are less pleasing to the extent that the aesthetic object deviates from the relevant prototypes for that culture.

Novelty

One of the most established principles in aesthetics is that people prefer aesthetic objects that have an intermediate degree of novelty; that is, uncertainty, surprisingness, incongruity, conflict, variability, or complexity (Berlyne, 1971; Cupchik, 1986; Fechner, 1876). There are a variety of explanations of the inverted-U function that relates aesthetic preference to novelty, namely those appealing to physiological arousal or information theory. Intuitively, a perfectly mundane object or story is boring and a completely novel one is complicated. Some dimensions of novelty do not depend on the world knowledge and culture of the appreciator, such as a dramatic contrast in color or a deviation from perceptual good form (Arnheim, 1971). However, most dimensions of novelty do depend on the world knowledge and culture of the comprehender. A surprising episode in a plot,

for example, deviates from the comprehender's expectations and world knowledge. If a novel plot or art form is successful, it is copied in the culture and eventually drifts toward a prototype, thereby losing its novelty (Martindale, 1986). It would be unproductive to define novelty according to stimulus properties entirely, because a critical consideration is the relationship between the stimulus and the world knowledge of the appreciator.

An observant reader might have noticed that there is a discrepancy between the predictions of the novelty component and the previous two components. Mundane aesthetic experiences would be preferred according to the good form and prototype components, whereas objects of intermediate novelty would be preferred according to the novelty component. It is possible to resolve the apparent discrepancy in a number of ways. Perhaps the aesthetic experience is cognitively or emotionally different when the good form and prototype components dominate compared to when the novelty component dominates. For example, the experience is pleasant in the case of an object with good form or with a close match to a prototype object whereas the experience is cognitively absorbing in the case of a moderately novel object (Cupchik & Gebotys, 1990). People, no doubt, differ with respect to their preference for a pleasant experience versus an experience in which they are cognitively challenged. Researchers need to identify more precisely the scope of the good form, prototype, and novelty components.

Literary and Aesthetic Devices

Associated with each art form or literary genre is a distinctive set of aesthetic devices that enhance the aesthetic experience. In visual art, there are special uses and combinations of color, shape, texture, and so on. The devices in literature include figurative language (e.g., irony, metaphor, hyperbole), fluctuations in perspective (e.g., points of view of the narrator and the characters, flashbacks), and phonological patterns (e.g., rhyme, assonance, alliteration). The type and density of these literary devices predict the extent to which texts are regarded as literary (van Peer, 1986). A number of researchers have investigated the psychological impact of various types of literary devices on literary comprehension, particularly the different forms of figurative language (Graesser, Mio, & Millis, 1988; Kreuz & Glucksberg, 1989; Roberts, 1992; Steen, 1989). Researchers need to explain (a) why some aesthetic devices can be

combined successfully whereas other combinations fail, and (b) why particular aesthetic devices can be successfully coordinated with particular semantic content.

People sometimes need to be trained how to identify and appreciate an aesthetic device. Otherwise, the device will have little or no impact on the comprehender. For example, naive viewers of visual art are sensitive to the representational aspect of paintings and sculpture (i.e., whether they mimic objects in the world) but not the stylistic transformations of visual form (Cupchik & Gebotys, 1988). Naive readers of fiction are often preoccupied with the issue of whether the narrative depicts real life (veridicality) or a semblance of real life (verisimilitude). They miss some of the fundamental levels of literature; namely, the aesthetic devices and the concept that the "truth" of the fiction resides at a more abstract level (Sell, 1991). According to Schmidt's (1982) aesthetic convention, the reader of literature should suspend concern over whether the information in literature is factually correct or even believable. Many readers need to be trained how to do this.

Unity in Diversity

Aesthetically pleasing stimuli should strike a proper balance between complexity (i.e., multiple devices, messages, points of attack) and orderliness (Cupchik, 1986; Fechner, 1876). Stated differently, there is a coherence that connects the diverse themes and aesthetic devices. Unfortunately, there have been substantial problems in defining unity and coherence. The definition of a coherent plot, for example, has occasionally been specified computationally and formally (Hobbs, 1982; Lehnert, 1981), but most of these solutions have proven to be inadequate when submitted to rigorous tests. One of the pressing research projects for the future is to explain and computationally account for the principle of unity in diversity. We need more systematic tests of the prediction that aesthetic experiences are more successful when there is unity in diversity.

Dynamic Cognitive and Emotional Activities

When an aesthetic object is comprehended, the appreciator experiences cognitive and/or emotional activities that dynam-

cally change over time. One typical cycle builds up cognitive complexity, and then abruptly simplifies the complexity: incongruity followed by resolution, confusion followed by clarification, tension followed by release, curiosity/suspense/surprise followed by a resolution (Berlyne, 1971; Brewer & Ohtsuka, 1988; Gerrig, 1989; Graesser, Long, & Mio, 1989; Jose & Brewer, 1990; Suls, 1972). Physiological arousal patterns both reflect and are explained by these cognitive activities (Mandler, 1975). That is, arousal increases when there is an incongruity in the cognitive system and decreases when the incongruity is resolved.

There are additional types of dynamic activity other than those already discussed. According to the polyvalence assumption, a reader constructs multiple interpretations of a sentence or excerpt in a literary text, whereas there is only a single bonafide interpretation in a nonliterary text (László, 1988; Meutsch & Viehoff, 1989; Schmidt, 1982; Zwaan, 1992). Dynamic cognitive activity would occur to the extent that a single reader oscillates between or among the various interpretations of a text segment. In the case of trained appreciators of aesthetic objects, the focus of attention oscillates between (a) the representational semantic content and global message, and (b) the surface forms and aesthetic devices associated with the medium (Cupchik & Gebotys, 1988). There are, no doubt, additional patterns of cognitive and emotional activity that characterize this component of aesthetic experience.

Psychoanalytic Themes and Content

Themes that involve death, destruction, sex, revenge, betrayal, and other forms of psychoanalytic conflict sustain the interest of the comprehender and provide excellent material for intense aesthetic experiences (Freud, 1949; Holland, 1975; Schank, 1979). These themes magnify the intensity of the rise and fall of emotions compared to more subtle themes.

We have identified many of the components of aesthetic experience in order to convey the richness and complexity of the phenomena. Aesthetic quality is clearly a complex, multifaceted construct and we undoubtedly have not exhausted its components. It is not surprising, therefore, that scientists and scholars have failed to identify any single essence of aesthetic experience and encountered widespread disagreement about the meaning of aesthetic quality. Given that aesthetic quality is in-

herently a difficult construct, we believe that it is important to investigate the components of aesthetic experience in depth and to build minitheories of the functioning of these components. Hopefully, we will eventually converge on a mature theory of aesthetic comprehension after we follow this "divide and conquer" strategy for several years or decades.

OBSTACLE 2: IT IS DIFFICULT TO CONSCIOUSLY IDENTIFY AND ARTICULATE THE COMPONENTS OF AESTHETIC EXPERIENCE

It has been extremely difficult for scholars and scientists to identify the components of aesthetic experience. The components are extremely abstract. Most of the components operate at an unconscious level in the cognitive system. Just as individuals are unaware of the psychological mechanisms that are executed during the processing of sentence syntax, most aesthetic components are not accessible to consciousness. Of course, individuals can be taught theories of aesthetics and can apply this knowledge strategically and consciously to aesthetic comprehension, but that is another matter. Whenever a concept is abstract, complex, and processed unconsciously, it is difficult to identify and articulate its properties (Ericsson & Simon, 1980; Graesser & Clark, 1985).

We conducted a simple study that assessed the extent to which college students are aware of the aesthetic components discussed in the previous section. College students at The University of Memphis were instructed to write down the properties of good art and literature. Twenty of the students were classified as being "literary" because they read two or more "literary" books per month. Twenty nonliterary students reported that they read zero books per month. We prepared a distribution of the properties that the students listed for good art or literature. The list that follows includes all properties that were generated by two or more students in both groups (literary and nonliterary); the percentage of students who generated each property is also included:

1. It is interesting and/or entertaining. (30%)
2. It is understandable. (23%)
3. It evokes thoughts and ideas. (20%)
4. It is educational. (18%)

5. It is creative. (15%)
6. It evokes emotion. (13%)

The next list includes all properties that were generated by two or more of the literary students, but none of the nonliterary students:

7. It is abstract. (20%)
8. It is unique. (20%)
9. It is unpredictable. (15%)
10. It is accepted universally. (15%)

Finally, the following properties were generated by two or more nonliterary students, but none of the literary students:

11. It pleases the senses. (25%)
12. It is realistic. (20%)

These 12 properties exhaust the properties that were articulated by two or more of the 40 college students in either of the two groups.

One striking outcome of this simple study is that the students rarely identified properties of the seven aesthetic components that were discussed in the previous section. If a very loose criterion were adopted, we could spot some correspondences between properties and aesthetic components. Properties 3, 4, 8, and 9 perhaps refer to the novelty hypothesis, although the inverted-U function was not mentioned. Property 2 perhaps refers to the unity in diversity component, but that clearly would be stretching any reasonable criterion. In our judgment, these data support the general claim that college students are not consciously aware of the components of aesthetic experience. These mechanisms operate at an unconscious level, and it is virtually impossible for these typical adults to verbally articulate what these mechanisms are.

We will also go out on the limb and claim that professionals in art and literature are unaware of most of the aesthetic components. It has frequently been reported that aesthetic value judgments are not well founded when professionals evaluate particular works of art, film, literature, and other aesthetic media (van Rees, 1986; Verdaasdonk, 1981). For example, a group of literary experts would presumably be able to articulate

and use a set of sensible criteria when they evaluate the novels in a literary competition, but that does not appear to be the case. Instead, they are substantially swayed by fashion, personal taste, politics, the opinions of a small group of colleagues (a form of elitism or provincialism), and other auxiliary considerations. The creators of aesthetic works also seem to be unaware of most of the aesthetic components. While conducting some investigations of humor and wit (Graesser et al., 1989; Long & Graesser, 1988), we were particularly interested in what comedians had to say about the generation of their material. We were surprised to learn that they had few analytical and theoretical insights about the mechanisms of humor. Instead of following a formula or a systematic analytical method, they would normally produce a successful humorous excerpt haphazardly and would subsequently repeat it with appropriate improvements and modifications. They appeared to use a random, trial-and-error, generate-and-test strategy of producing the initial humorous excerpt.

Our claim is that both novices and experts are unaware of most aesthetic components, including those that were discussed in the previous section. This has made it particularly difficult to investigate art and literature empirically. Quite clearly, there is only one way of confronting this serious obstacle—researchers need to pursue theories and minitheories of aesthetic experience. Theories and minitheories furnish the deep insights about aesthetic experience that will sustain our field and prevent us from being swayed by the clouds of shallow folklore.

OBSTACLE 3: THERE IS A LACK OF SOPHISTICATED THEORIES AND MINITHEORIES

Given the two aforementioned obstacles, it is not surprising that the field has failed to develop a sophisticated general theory of aesthetic experience. We believe that a sophisticated theory would have three characteristics, which are discussed later. Our background is in psychology and text comprehension, so the emphasis here is on a psychological theory of literary comprehension.

The first characteristic of a sophisticated theory is that it is psychologically plausible, that is, the theoretical assumptions

about knowledge representation and processing should not conflict with the properties of human cognition. It is beyond the capability of the human mind, for example, for the reader to generate dozens of alternative future worlds (i.e., possible sequences of future episodes) when people comprehend a particular episode in a novel (Graesser, 1981; Johnson-Laird, 1983). The human mind can normally construct only one or two mental models when future occurrences are forecasted. Similarly, it is beyond the span of working memory to generate several alternative readings of a sentence or excerpt when people normally read literature (Just & Carpenter, in press). The limitations of the human mind also place constraints on what fictive worlds can be constructed and what literary devices can be assimilated during literary comprehension.

The second characteristic of a sophisticated theory is that it is well-specified computationally and analytically. Ideally, a computer would be capable of simulating the cognitive content and emotions of a comprehender of a literary work. Computer simulation is indeed one of the most powerful demonstrations that a researcher has pinned down all of the processing mechanisms and has not glossed over important details. In principle, the computer could simulate most aspects of comprehension: the word lexicon, syntax, propositional segmentation, the formation of coherent constituents, the mental model constructed for the plot, the point of a message, aesthetic devices, the world knowledge, memories, and cognitive style of a particular comprehender, the emotions of characters, the emotions of the comprehender, and so on. In practice, however, it is horrendously difficult to simulate most of these components on a computer. The alternative is to construct a detailed, precise, analytical model on paper that captures each level of aesthetic comprehension. Our hope is that a general model with analytical precision will eventually be developed after researchers spend a few years or decades developing minitheories. There is currently no sophisticated theory that is grounded computationally or analytically.

A third characteristic of a sophisticated theory is that it can handle actual literary text. It is important to satisfy the criterion of ecological validity, such that the texts are naturalistic and the reader genuinely wants to read the literature. It is not sufficient to conduct research exclusively on isolated sentences and short pointless stories that are written by experimenters who attempt to satisfy the constraints of research designs. The

typical research designs in experimental psychology have imposed some control over variables at shallow levels, such as word frequency and syntactic complexity, but not the variables at deeper levels, such as coherence, world knowledge, pragmatics, and aesthetics. Given that experimental psychologists have consistently failed to control variables at deeper levels of comprehension, we recommend that researchers focus primarily on naturalistic literature.

Throughout this chapter, we have advocated the development and testing of minitheories of literary comprehension. A minitheory bites off a reasonable chunk of the phenomenon rather than confronting the problem of literary comprehension in its entirety. For example, a researcher might develop a minitheory that explains how people process a particular class of metaphors. The minitheory would precisely specify the scope of the minitheory, the prerequisite knowledge of the comprehender, the elements in the text that trigger the metaphor, the mappings between levels and between information sources, and the cognitive/emotional processing components.

There is an optimal scope for a minitheory. If the minitheory is too broad, the assumptions and claims are too abstract to be of much use to researchers of literary comprehension. If it is too narrow, the researcher becomes lost in a sea of details that interest very few colleagues. So what is the ideal bite size for a minitheory? We would argue that there are three signs that a minitheory has a satisfactory scope. The first sign is that the minitheory is empirically testable—it should furnish sufficiently decisive predictions so that a researcher can collect data and test the prediction. The second sign is that the minitheory generates new insights—it should have a sufficient complexity to furnish new ideas, predictions, and perspectives. The third sign is that the assumptions and predictions of the minitheory have some teeth—there should be predictions that are counterintuitive and distinctive in the sense that they challenge the prevailing folklore, wisdom, and clichés about aesthetic experience. A claim or prediction with teeth is generated by the underlying minitheory, but not by alternative minitheories.

At this point, we elaborate on our claims about theories and minitheories by briefly presenting and critiquing a particular research project. We prefer to avoid pointing our finger at one of our colleagues, so we will select a research project that we recently conducted. This research project investigated the extent

to which reading time for excerpts in short stories could be predicted by the three major parts of a literary communication system: the author, the text, and the reader. Literary critics have varied substantially in their emphasis on each of these three components. Historical literary critics emphasized the culture, life, and motivations of the author, whereas the text reigned supreme for the New Critics, and the reader was dominant for the Reader Response theorists (Bleich, 1978; de Beaugrande, 1989; Fish, 1980; Fokkema, 1990; Ibsch, Schram, & Steen, 1991; Rosenblatt, 1985). The author, text, and reader each have their unique constraints on the literary communication process. As psychologists, we also were interested in the relative contributions of these three components on such measures as reading time, memory, aesthetic quality, and so on. The project under consideration investigated reading time.

Our method of collecting data was quite straightforward. We selected 8 short stories that were written by well-known authors: "A & P" (John Updike), "A Very Old Man with Enormous Wings" (Gabriel Garcia Marquez), "The Bass, the River, and Shiela Mant" (W. D. Wetherell), "The Necklace" (Guy de Maupassant), "Astronomer's Wife" (Kay Boyle), "The Demon Lover" (Elizabeth Bowen), "The Duchess and the Jeweller" (Virginia Woolf), and "A Hunger Artist" (Franz Kafka). The stories had roughly the same length (8–10 book pages). We segregated each story into 15 excerpts and placed a single excerpt on each page of a 15-page booklet. College students at The University of Memphis were instructed to read two of these stories on two occasions (four readings altogether). Half of the 64 subjects were instructed to read the stories for memory, whereas the other half were instructed to read for enjoyment. Each subject supplied a reading time for each of the 15 excerpts that were part of each of the four readings. This was accomplished by having subjects record the time from a clock (in second accuracy) when they began reading an excerpt and when they finished reading the excerpt. The difference between these times served as our measure of the reading time for the excerpt. It should be noted that this is an easy way to collect reading time data because subjects can be run in large groups and subjects can participate at their leisure. We had the students read the stories in groups of 3–10 in a classroom environment. We also administered a 30-minute test of literary expertise that had been developed by David Miall at the University of Alberta. In summary, the study had a factorial design in which excerpt

reading time was measured as a function of reader goals (memory versus enjoyment) and repetition (first versus second reading). We also had a measure of the literary expertise of each student.

Two expert judges supplied theoretical ratings from each of the excerpts. Given that there were 8 stories and 15 excerpts per passage, there were 120 excerpts altogether. Each excerpt was rated on five variables, with four values per variable (no, guess no, guess yes, yes):

1. *Author intent*. Is the author conveying an intention or attitude in this excerpt (see Gibbs, Kushner, & Mills, 1991; Hunt & Vipond, 1986)?
2. *Staticness of text content*. Does the excerpt convey primarily static facts (i.e., setting information) versus actions and events (Graesser, 1981)?
3. *Climax*. Does the excerpt convey the climax and major episodes in the plot, as opposed to support content?
4. *Character emotion*. Are the characters experiencing salient emotions in the excerpt (Stein & Levine, 1990)?
5. *Reader emotion*. Is the reader experiencing suspense, surprise, curiosity, or some other salient emotion while reading the excerpt (Brewer & Ohtsuka, 1988)?

In addition to these theoretically informative variables, the excerpts were measured on number of words (mean = 189) and serial position (1-15).

The results of the study uncovered a number of trends when we examined correlations between the various measures and the mean reading time per excerpt (averaging across students). In these analyses, partial correlations were computed that controlled for excerpt length (i.e., the number of words in the excerpt). Reading times significantly increased as a function of (a) the climax variable for the second reading in both the enjoyment and memory conditions, (b) the "staticness" variable for the second reading in both the enjoyment and memory conditions, and (c) the reader emotion variable for the second reading in the memory condition. The author intent and character emotion variables were not significant in any analysis. Reading times also decreased as a function of serial position in all readings and conditions. We performed follow-up analyses in order to assess whether the results were sensitive to lit-

erary expertise or to the overall reading speed of the students, but these measures of individual differences failed to reach significance.

The significant results of this study can be meaningfully compared to previous research. The finding that the theoretically interesting variables had a greater impact on the second reading than the first can be compared to previous studies that have investigated how multiple readings of literature influence reading time (Miall, 1988) and recall (Halász, Carlsson, & Marton, 1991). Brewer's structural-affect theory (Brewer & Ohtsuka, 1988) would account for the impact of the climax variable and the reader emotion variable on the reading times of the second reading. The reader emotion effect might also be explained in some fashion by the reader's reminders of personal experiences during literary comprehension (Halász et al., 1991; Larsen & Seilman, 1988). The influence of staticness and serial position on reading time replicates previous research in our lab (Graesser, Hoffman, & Clark, 1980; Haberlandt & Graesser, 1985), as well as the research of some of our colleagues. Quite clearly, we have uncovered some new findings and have integrated these findings with some previous research. Superficially, it would appear that we have successfully practiced science.

We argue, however, that the results of this study are not particularly informative from the standpoint of building a sophisticated theory of literary comprehension. Perhaps the major problem is that the theoretical variables were not specified computationally and analytically in sufficient detail to pin down what was being measured theoretically. For each theoretic variable, such as author intent, we asked the judges to read some landmark journal articles and then to supply ratings on excerpts using a 4-point scale. It was unclear what text constituents and criteria they were using when making these judgements. Each of the theoretical variables would need to be unpacked at greater depth before much light would be shed on a sophisticated theory. Consider, for example, author intent. What intentions and attitudes does the author probably have? How meaningful are these intentions and attitudes to the reader? What particular elements and constituents in the stories provide information about author intent and attitude? What knowledge does the author need in order to recognize how these elements and constituents transmit author intent? A sophisticated theory would require answers to these questions.

tively produce and comprehend literature and art. At the very least, we will take the mystery out of aesthetic genius.

REFERENCES

- Arnheim, R. (1971). *Art and visual perception*. Berkeley: University of California Press.
- Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York: McGraw-Hill.
- Bleich, D. (1978). *Subjective criticism*. Baltimore: Johns Hopkins University Press.
- Brewer, W. F., & Ohtsuka, K. (1988). Story structure, characterization, just world organization, and reader affect in American and Hungarian short stories. *Poetics*, 17, 395-415.
- Chupchik, G. C. (1986). A decade after Berlyne: New directions in experimental aesthetics. *Poetics*, 15, 345-369.
- Chupchik, G. C., & Gebotys, R. (1988). The experience of time, pleasure, and interest during aesthetic episodes. *Empirical Studies of the Arts*, 6, 1-12.
- Chupchik, G. C., & Gebotys, R. (1990). Interest and pleasure as dimensions of aesthetic response. *Empirical Studies of the Arts*, 8, 1-14.
- de Beaugrande, R. (1989). Toward the empirical study of literature: A synoptic sketch of a new "society." *Poetics*, 18, 7-27.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.
- Fechner, G. T. (1876). *Vorschule der Aesthetik*. Leipzig: Breitkopf & Hartel.
- Fish, S. (1980). *Is there a text in this class? The authority of interpretive communities*. Cambridge, MA: Harvard University Press.
- Fokkema, D. W. (1990). The study of literature. In E. Zurcher & T. Landorff (Eds.), *The humanities in the nineties: A view from the Netherlands*. Amsterdam: Swets & Zeitlinger.
- Freud, S. (1949). *An outline of psychoanalysis*. New York: Norton.
- Garner, W. R. (1970). The stimulus in information processing. *American Psychologist*, 25, 350-358.
- Gerrig, R. J. (1989). Suspense in the absence of uncertainty. *Journal of Memory and Language*, 28, 633-648.
- Gibbs, R. W., Kushner, J. M., & Mills, W. R. (1991). Authorial intentions and metaphor comprehension. *Journal of Psycholinguistic Research*, 20, 11-30.
- Graesser, A. C. (1981). *Prose and comprehension beyond the word*. New York: Springer.
- Graesser, A. C., & Clark, L. F. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Ablex.

Very little is learned when we reduce these complex phenomena to a single measured variable.

A second problem with the study was that it lacked ecological validity. We did indeed select literary texts that were written by accomplished authors. However, the study failed to satisfy ecological validity with respect to contextual variables. The readers were forced to read the stories rather than choosing to read them. Most people read a particular novel in order to satisfy personal goals: entertainment, spiritual enrichment, development, etc. Most people read in a personal setting at private times. These contextual variables were ignored in the present study; somehow the essential motivation of reading literature was entirely missed. People presumably read quite differently when they genuinely want to read versus when they have to read. For example, Graesser, Higgenbotham, Robertson, and Smith (1978) reported that college students read a sensationalist news magazine (the *National Enquirer*) quite differently under self-induced versus task-induced conditions. Under self-induced conditions, college students tended to read articles on topics with which they were familiar and extracted more active information than static information. These effects entirely disappeared under task-induced reading conditions.

Finally, a third problem with our study was that it failed to shed light on a single minitheory. Minitheories rarely evolve by collecting correlations between gross variables (i.e., reading time and reader emotion). Instead, the various components and properties of the minitheory need to be unpacked, the distinctive predictions identified, and the complex patterns of data compared to the predictions. None of the statistically significant outcomes of this study would surprise our grandmothers or would challenge any existing minitheory. Therefore, the empirical results had no teeth.

The purpose of critiquing our own study was to illustrate how a superficially reasonable study would ultimately have essentially no impact on the development of minitheories and sophisticated theories of literary comprehension. We encourage researchers in the future to develop and test minitheories that will advance the empirical studies of literature, art, and media. We promise that a sophisticated theory will eventually emerge after several years, decades, centuries, or millennia of working on our minitheories. If we are successful, then the planet will be blessed with thousands of humanistic computers that ac-

- Graesser, A. C., Higginbotham, M. W., Robertson, S. P., & Smith, W. R. (1978). A natural inquiry into the *National Enquirer*: Self-induced versus task-induced reading comprehension. *Discourse Processes, 1*, 355-372.
- Graesser, A. C., Hoffman, N. L., & Clark, L. F. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior, 19*, 135-151.
- Graesser, A. C., Long, D. L., & Mio, J. S. (1989). What are the cognitive and conceptual components of humorous text? *Poetics, 18*, 143-163.
- Graesser, A. C., Mio, J., & Millis, K. (1988). Metaphors in persuasive communication. In D. Meutsch (Ed.), *Models of literary understanding* (pp. 86-101). Amsterdam: Elsevier.
- Haberlandt, K., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General, 114*, 357-374.
- Halász, L., Carlsson, M. A., & Marton, F. (1991). Differences in recall and understanding of literary texts read several times. *SPIEL, 10*, 235-262.
- Hobbs, J. R. (1982). Toward an understanding of coherence in discourse. In W. G. Lehnert & M. H. Ringle (Eds.), *Strategies for natural language processing* (pp. 223-243). Hillsdale, NJ: Erlbaum.
- Holland, N. H. (1975). *Five readers' readings*. New Haven, CT: Yale University Press.
- Hunt, R. A., & Vipond, D. (1986). Evaluations in literary reading. *Text, 6*, 53-71.
- Ibsch, E., Schram, D. H., & Steen, G. J. (Eds.). (1991). *Empirical studies of literature: Proceedings of the second IGEL Conference*. Amsterdam: Rodopi.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Jose, P. E., & Brewer, W. F. (1990). Early grade school children's liking of script and suspense story structures. *Journal of Reading Behavior, 22*, 335-372.
- Just, M. A., & Carpenter, P. A. (in press). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*.
- Kreuz, R. J., & Glucksberg, S. (1989). How to be sarcastic: The echoic reminder theory of verbal irony. *Journal of Experimental Psychology, 118*, 374-386.
- Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 2*, 115-121.
- Larsen, S. F., & Seilman, U. (1988). Personal reminders while reading literature. *Text, 8*, 411-429.
- László, J. (1988). Readers' historical-social knowledge and their interpretation and evaluation of a short story. *Poetics, 17*, 461-481.
- Lehnert, W. G. (1981). Plot units and narrative summarization. *Cognitive Science, 4*, 293-331.
- Light, L. L., Hollander, S., & Kayra-Stuart, F. (1981). Why attractive people are harder to remember. *Personality and Social Bulletin, 7*, 269-276.
- Long, D. L., & Graesser, A. C. (1988). Wit and humor in discourse processing. *Discourse Processes, 11*, 35-60.
- Mandler, G. (1975). *Mind and emotion*. New York: Wiley.
- Martindale, C. (1986). Aesthetic evolution. *Poetics, 15*, 439-473.
- Meutsch, D., & Viehoff, R. (Eds.). (1989). *Comprehension of literary discourse: Results and problems of interdisciplinary approaches*. Berlin: De Gruyter.
- Miall, D. S. (1988). Affect and narrative. *Poetics, 17*, 259-272.
- Reed, S. K., (1972). Pattern recognition and categorization. *Cognitive Psychology, 3*, 382-407.
- Roberts, R. M. (1992). *The production of figurative language: Function and context*. Unpublished doctoral dissertation, The University of Memphis, Memphis, TN.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structures of categories. *Cognitive Psychology, 4*, 573-605.
- Rosenblatt, L. M. (1985). The transactional theory of the literary work: Implications for research. In C. Cooper (Ed.), *Researching response to literature and the teaching of literature* (pp. 33-53). Norwood, NJ: Ablex.
- Schank, R. C. (1979). Interestingness: Controlling inferences. *Artificial Intelligence, 12*, 273-297.
- Schmidt, S. J. (1982). *Foundations for the empirical study of literature: The components of a basic theory*, Vol. 1. (R. de Beau-grande, Trans.). Hamburg: Buske.
- Sell, R. D. (1991). How can literary pragmaticists develop empirical methods? The problem of modal and evaluative expression in literary texts. In E. Ibsch, D. Schram, & G. Steen (Eds.), *Empirical studies of literature: Proceedings of the second IGEL Conference* (pp. 137-145). Amsterdam: Rodopi.
- Steen, G. (1989). Metaphor and literary comprehension: Towards a discourse theory. *Poetics, 18*, 113-141.
- Stein, N. L., & Levine, J. L. (1990). Making sense out of emotion: The representation and use of goal-structured knowledge. In N. L. Stein, B. Leventhal, & T. Trabasso (Eds.), *Psychological and biological approaches to emotion* (pp. 45-73). Hillsdale, NJ: Erlbaum.
- Suls, J. M. (1977). Cognitive and disparagement theories of humor: A theoretical and empirical synthesis. In A. J. Chapman & H. C. Foot (Eds.), *It's a funny thing, humor*. Oxford: Pergamon.
- van Peer, W. (1986). *Stylistics and psychology: Investigations of foregrounding*. London: Croom Helm.

van Rees, C. J. (1986). *Literary theory and criticism: Conceptions of literature and their application*. Unpublished doctoral dissertation, University of Groningen, The Netherlands.

Verdaasdonk, H. (1981). *Literatuurbeshouwing en argumentatie*. Amsterdam: Huis aan de drie grachten.

Zwaan, R. A. (1992). *Aspects of literary comprehension: A cognitive approach*. Unpublished doctoral dissertation, University of Utrecht, The Netherlands.

2

The Strained Relationship Between the Empiricist's Notion of Validity and the Hermeneutician's Notion of Relevance

Elrud Ibsch

In 1991, the editors of the Dutch cultural journal *De Gids* (founded in 1837) asked 12 literary scholars to comment on recent developments in literary theory, and especially on the question of whether there is a new paradigm in literary scholarship and, if so, how it is being articulated. Although the small number of scholars does not promise a representative result, one has to take into consideration the relatively small quantity of Dutch universities and the fact that, among the 12, 7 are full professors (either of General and Comparative Literature or of Russian and English Literature), one is a professor emeritus.