1998

Intentionalism in aesthetics

Paisley LIVINGSTON
University of Aarhus, pl@ln.edu.hk

Recommended Citation

Livingston, P. (1998). Intentionalism in aesthetics. New Literary History, 29(4), 831-846. doi: 10.1353/nlh.1998.0042

Intentionalism in Aesthetics

Paisley Livingston

INTENTIONALISM IN AESTHETICS IS, quite generally, the thesis that the artist's or artists' intentions have a decisive role in the creation of a work of art, and that knowledge of such intentions is a necessary component of at least some adequate interpretive and evaluative claims.[1] In this paper I develop and defend this thesis. I begin with a discussion of some anti-intentionalist arguments. Surveying a range of intentionalist responses to them, I briefly introduce and criticize a fictionalist version of intentionalism before moving on to an approach I call moderate intentionalism. I consider a salient alternative known as hypothetical intentionalism and try to show why moderate intentionalism should be preferred to it.

Saying what, precisely, intentions are is no small problem, and disputes in aesthetics often hinge on rival assumptions about the nature and function of intentions in general. I shall assume, in what follows, that intentions are mental states having semantic contents, various psychological functions, and practical consequences—but not always the targetted results.[2] I shall not take up any of the more global challenges to intentionalist psychology, such as eliminative materialism or macro-sociological and historicist critiques.[3] I assume, then, that agents sometimes intend to perform an action, such as writing a poem, and that they occasionally succeed in realizing such aims, thereby intentionally doing such things as writing poems.

I. Extreme Intentionalism and Anti-Intentionalism

In an extreme version, intentionalism holds that a work's meanings and its maker's intentions are logically equivalent. Such a thesis still has its defenders, yet it is hard to see how it can be reconciled with the fact that intentions are not always successfully realized. A theory of interpretation based on Humpty Dumpty's semantics does not seem promising.[4] An extreme version of anti-intentionalism also has its advocates, who confront the intentionalist with the following dilemma: either the artist's intentions are successfully realized in the text or structure produced by the artist, in which case the interpreter need not refer to them; or, the

New Literary History, 1998, 29: 831–846

artist's intentions are not successfully realized, in which case reference to them is insufficient to justify a related claim about the work's meanings. Any viable form of intentionalism must find a way out of this dilemma.

A premise of many anti-intentionalist arguments—including the dilemma just mentioned—is that if a work has determinate meanings and value, they must be immanent in the artistic text or structure. This sort of empiricism in aesthetics is vulnerable to some powerful criticisms. Not all of the artistically or aesthetically relevant features of a work of art are intrinsic properties of the text; some are relational and can only be known when the text or structure is cognized correctly in the context of its creation. In making this point, a number of philosophers, such as Arthur Danto, David Davies, Jerrold Levinson, and Gregory Currie, have evoked versions of Jorge Luis Borges's fictional example of Pierre Menard: tokens of the same text-type, created in different contexts, manifest different, artistically relevant relational features; to know which features are those of one work as opposed to another work, one must interpret the text in its context of creation.[5]

Once attention has been drawn to the constitutive status of a work's relational properties, cogent responses to the anti-intentionalist dilemma can be formulated. The intentionalist can argue that some successfully realized intentions are not simply redundant with regard to the text's intrinsic features. An example is the intention that a certain meaning be unstated in the text yet implicitly expressed by the work. Even when the intentions are successfully realized, such relations are not immanent in the final artistic structure or text and cannot be simply read off from the latter. Intentionalists also contend that whenever our goal is to evaluate a work as a certain kind of achievement, the artist's intentions, including unsuccessfully executed ones, are always relevant, because *part* of what we want to do is take note of the manner and extent of the artist's realization of the relevant aims. Although it is not the case that success at realizing one's intentions entails success at creating a valuable work, success or failure in realizing intentions does have implications for the kind of value a work possesses, if only because there is a significant and relevant difference between lucky and skillful creative activities. Noël Carroll contends, for example, that Ed Wood's *Plan 9 From Outer Space* would have been a better film had the director been trying to make a parody. What one sees and hears at a screening of this movie is logically compatible with such an intention, but our knowledge that the director in fact had no such intention is decisive, and we cannot justifiably praise the film as a clever parody.[6] So the dilemma can be avoided: knowledge of the relation between an artist's intentions and the resulting structure is necessary to at least some interpretive and evaluative claims.

One way in which anti-intentionalists challenge even modest versions of intentionalism is to evoke epistemological worries about the difficulty or impossibility of obtaining reliable knowledge of intentions. Usually this amounts to an unjustified demand for a kind of infallible justification or proof that is unattainable in any empirical domain, and no reason is given why such high standards should be imposed on claims about intentions and other mental states. The inconsistency is flagrant when the theorist who voices such skepticism about intentions makes all sorts of bold claims about such complex topics as the nature of textuality and the ways in which readers construct authors. If we can know how readers construct things, why can we not know how artists do so?

A more reasonable worry voiced by anti-intentionalists is that intentionalist strictures make the focus of appreciation shift radically away from the text or structure and toward the life and mind of the artist, the ultimate result being a ratification of the most narrow variety of biographical criticism. Although the latter has its rewards, it does not yield the only or the most valuable interpretations of an artist's works. For example, George Painter's biographical study of Marcel Proust certainly sheds some light on Proust's achievement, but interpretations by Vincent Descombes, René Girard, and many others are also valuable, partly because they elucidate Proust's texts in ways that the writer himself would never have done so.[7] So if we are to defend some form of intentionalism, it should not be one that prescribes only biographical approaches or one that prohibits novel and creative interpretations, including those that explore the significance that a work takes on outside the context of its creation. Part of the solution, then, is to observe that interpretations can manifest many different sorts of value. Elucidating the meanings and artistic value a work had in its original context is one sort of valuable goal, but finding clever, new, and even anachronistic ways of using a text can also be worthwhile. It is not clear that these two sorts of interpretive projects are always in competition with each other. So a premise of any tenable intentionalist theory of interpretation is that a valued *but not exclusive* goal of interpretation is the epistemic one of knowing something about the work of art *qua* work of art, and this in its original context of production. The intentionalist theses I discuss in the rest of this paper all share this premise.

II. Fictionalist Intentionalism

One approach that appeals to many critics is to maintain some sort of intentionalist framework while adopting an instrumentalist stance with regard to authorship. Looking at the textual evidence relevant to an artistic corpus, the interpreter seeks, then, not to build the most realistic

possible portrait of the life and works in combination, but to yield an interesting and rewarding interpretation of the works, viewed as the product of a postulated or fictional author. Such an approach has been defended by Alexander Nehamas, for whom the target of interpretation is the attitudes of a non-existent, make-believe authorial figure.[8]

The sort of *fictionalist* intentionalism is awkward, I think, because there is an unresolved tension between two tendencies. On the one hand, the interpreter is supposed to be building a picture of an agent whose actions and attitudes *explain* the genesis of the work, and such a process is presumably governed by constraints having to do with psychological plausibility. If we really are interested in how works are made, we should be interested in the desires, beliefs, intentions, and other relevant attitudes of the actual makers, and we should also deem it relevant to know whether an attitude has been intentionally expressed or has only accidentally been made manifest in the work. Yet in spite of its apparent emphasis on artistic agency, fictionalist intentionalism describes an interpretive process that is not really aimed at forming a hypothesis, not even a selective one, about the actual writer; instead, interpreters only imagine or make-believe that the fictional entity they describe was responsible for creating the text. As a result, significant differences between possible creative histories are effaced.

Consider, for example, cases where an interpreter holds, solely on the basis of evidence that is in principle accessible to members of the appropriate audience, that there are artistically significant *implicit* patterns and meanings in a novel. The history of the work's production could have been a matter of three different kinds of processes: (1) the implicit meanings could have been intended by the artist, who wanted them to be implicitly conveyed by the work; (2) the author may have had no such intentions, but ended up writing a text compatible with such a reading; or (3) the author may have had such relations in mind, intending to make them *explicit* in the text, but failed to realize this intention. The fictionalist intentionalist who is attuned to the implicit pattern cannot speak of the differences between these three cases, and can only describe the implicit meanings as expressing the intentions of a fictionalized author. The actual author's intentions, when successfully acted on, do not have any constitutive or evidentiary role in making these implicit relations part of the work's artistic content.

Textual appearances can be deceptive: a text that emerges from a chaotic and uncoordinated process of multiple authorship involving the efforts of various individuals could look like it resulted from this kind of messy history; but on the other hand, it could look like it had been intentionally produced by a single author or group of authors acting on a reasonably well-conceived and executed scheme.[9] Similarly, a text

intentionally crafted by a single (or collective) author acting on a well-conceived and well-executed scheme could look like something emerging from an uncontrolled process of multiple, uncoordinated contributions; but it could also look like some author's controlled, intentional doing. The interpreter who is oblivious to the intentional or unintentional nature of the actual creative process, as opposed to the text's appearances, is not in a position to distinguish between these four different kinds of cases. Working with a default assumption in favor of intentional creation and expression, the fictionalist runs the risk of mistaking accidentally coherent textual meanings for intentionally expressed attitudes, attributing the latter to a nonexistent creator.

III. Moderate Intentionalism

Moderate intentionalism is the thesis that the *actual* maker(s)' attitudes and doings are responsible for some of a work's content, and as such are a legitimate target of interpretive claims; more specifically, knowledge of some, but not all intentions is necessary to some, but not all valuable interpretive insights because such intentions are sometimes constitutive of the work's features or content. Moderate intentionalism recognizes that the artist's intentions do not always constitute the work's meaning. The contention, rather, is that when intentions are compatible with the text, they can be constitutive of a work's implicit meanings. Just as hinting and insinuation are part of the pragmatics of everyday conversational exchange, so do artists sometimes enhance the value of their works by expressing attitudes in an implicit and indirect manner. In many artistic contexts, subtlety is a valuable feature, and bluntness a failing.

Moderate intentionalism's claims about the implicit meanings of a work can be articulated within a *broadly* Gricean framework where the notion of conversational implicature has been adapted so as to develop a conception of what could be dubbed "artistic" implicature.[10] A key claim, then, is that appropriate inferences made within the artist/interpreter relation are guided by assumptions analogous—but certainly not identical—to the maxims proposed by Paul Grice with reference to everyday conversation. Artistic implicatures, then, are inferences to implicit content based on the explicit content of a text or artifact, as well as on assumptions shared by artists and their audiences, including contextual beliefs and beliefs about the nature of the artist/interpreter interaction. For example, authors and interpreters are guided by the hypothesis of a "thin" authorial rationality: if an author intends to express p implicitly, the author will try to adopt expressive means that

are likely to make *p* manifest to interpreters who are reasonably competent at assessing textual and contextual evidence. To that end, the author intends to write a text that does not contain *p* as part of its explicit content, a text, however, which will make it possible (if not highly likely) for the members of the audience to infer the implicit content by relying on both the text and contextual assumptions. What is more, when authors try to communicate something implicitly, they sometimes intend for their success in realizing this aim to depend on the audience's recognition of that intention. A schematic illustration of how moderate intentionalist principles may be exemplified in the interpretation of a work of fiction is provided below in section V.

IV. Hypothetical Intentionalism

Moderate intentionalism is, I think, the right way to go, but it is important to see how one may defend it against certain challenges. One objection that appears frequently in the literature takes the following form. Take some literary text or artistic structure that is well known and that is generally recognized as having valuable and complex meanings, and imagine that we were to discover that the artist in question produced the work while acting on only some very limited semantic intentions. Does not moderate intentionalism then have the crippling consequence of requiring us to limit our understanding of the work's original, artistic meanings to the ones intended by the artist? And why should we want to do this when the interest and value of the work would appear to suffer as a result? In Jerrold Levinson's version of this challenge, we are asked to imagine that we discover that Franz Kafka's intentions with regard to "*Ein Landarzt*" were simply a matter of critiquing rural medical practices. Should we not reject any hermeneutic principle that would have the deflationary consequence of forcing us to ignore the rich symbolic dimensions of Kafka's story?[11]

The idea behind this sort of challenge and the examples that are used to illustrate it is that some artist's intentions can detract from the value or interest of a work, and that a theory of interpretation should provide a principled way of ruling them out. An approach along these lines is ably defended by Levinson, who follows William Tolhurst in speaking of hypothetical intentionalism.[12] Crucial to this approach is a distinction between two main kinds of intentions, labeled "semantic" and "categorial" by Levinson. Speaking quickly, Levinson proposes that categorial intentions can determine a work's features and therefore have a constitutive status, while semantic ones cannot and are at best suggestive of a work's meanings. Whenever heeding someone's semantic intentions would

make the interpretation less interesting and the work less valuable, we should overrule them in favor of a superior interpretation that is compatible with the textual and contextual data. In the case of the Kafka example, Levinson's strictures would have us rule out Kafka's inferior semantic intentions while retaining our crucial knowledge of his larger categorial aims.

Such an approach obviously hinges on the distinction between semantic and categorial intentions, which is drawn, first of all, on the basis of the contents of intentions, and involves, more specifically, the aspects of the work of art that the artist has in mind. As Levinson puts it, categorial intentions "govern not what a work is to mean but how it is to be fundamentally conceived or approached" (I 188). In one of Levinson's examples, the intention to make a sculpture and have it be taken as such is categorial, while the intention to express rage with this work of art is semantic.[13]

Why should categorial intentions have a different status in a theory of interpretation? In what follows, I survey various potential reasons, contending that on closer inspection, they do not in fact justify the use made of the distinction in hypothetical intentionalism.

One potential reason for a difference in status has to do with reliability. Perhaps the two kinds of intentions have significantly different functions in the creative process, in which case interpreters who follow the principles of hypothetical intentionalism are attuned to an important difference. Levinson writes that categorial intentions are *decisive* or *determinative* of a work's features in a way that semantic intentions are not. He points out that semantic intentions often fail—as a result, say, of clumsiness or mistaken beliefs. He then adds: "But if the writer intends his text *as a poem*—as opposed to a short story, a dramatic monologue, a piece of calligraphic visual art, or a mere diary entry—then that intention is of a different sort and of a different order, and virtually cannot fail—so long as the text in question at least allows of being taken, among other things, as a poem" (I 188). In the same context, Levinson goes on to say that semantic intentions do not "determine" meaning, while categorial intentions "*do* in general determine how a text is to be conceptualized and approached on a fundamental level and thus indirectly affect what it will resultingly say or express" (I 189). And that, presumably, is a reason, perhaps even a sufficient reason, why semantic intentions should have only a *suggestive* role in the construction of interpretive hypotheses, while categorial ones have an "*evidential* role."

Levinson says semantic intentions can fail; categorial ones *virtually* cannot fail. Does this phrase mean that they do, sometimes, fail? So it would seem. Levinson allows that both categorial and semantic intentions are fallible, so the reliability of the former is no reason for granting

the two a logically distinct status in our theory of interpretation. Perhaps Levinson's point in this regard is that it is in general easier to realize categorial intentions, and that semantic intentions are more likely to misfire. Yet even this more modest thesis is not so obvious. Some categorial intentions may, in some contexts, be very hard to realize; and some semantic intentions are easy to pull off. Degree of difficulty and likelihood of success do not in any case correspond in any simple way to whether knowledge of someone's aims has constitutive or merely suggestive value with regard to their actual achievement.

A version of the intentional fallacy pertains to categorial intentions just as much as it does to semantic ones. We cannot infer from someone's having a categorial intention that it has been successfully realized in the work, even if the agent is known to have acted on that intention. Nor can we automatically infer back from features of a realized text or artistic structure to the relevant categorial intentions. A writer shows us a sonnet he has authored. Can we conclude, therefore, that the author categorially intended to write a sonnet and intentionally did so? The argument is invalid, even if we are willing to set aside cases of wayward causality. The author could have been trying to realize a specific categorial intention incompatible with the poem's actually being a sonnet.

Levinson allows that categorial intentions are decisive or determinant only if the text "allows of being taken" that way. The same sort of constraint can be placed on our use of facts about semantic intentions: a semantic intention is to be deemed decisive of a work's content only if the text "allows of being taken" that way. Semantic intentions do not, indeed, succeed "by fiat," but neither do categorial ones. In both cases, recognition of the artist's constitutive role is constrained by facts about what the writer has managed to do in producing a text. Moderate intentionalism can make use of the same insight, holding that intentions of any stripe are decisive only if they are textually or structurally compatible, that is, if they are consistent with the features of the work's text or artistic structure. Successful realization of intention is, in both cases, a matter of intentionally producing something compatible with the content of the intention.[14]

Reliability, or degree thereof, turns out not to be the key to any important difference in status between the two kinds of intention, and therefore not a decisive reason for preferring hypothetical intentionalism over moderate intentionalism. Are there other reasons? One candidate to consider has to do with accessibility or epistemic access: perhaps categorial intentions are more readily known, while semantic ones are elusive. Yet once we are in the business of making claims about the respective contents of intentions, we are in no position to say that

semantic intentions are inscrutable creatures of the mentalistic night, while categorial ones are solid and scrutable features of objective behavior. We manage to know both sorts of intention—when, that is, we do manage to know them at all—in the same way. Some categorial intentions are, in any case, very hard to fathom. What precisely was the categorial intention of Apuleius when he wrote the last part of *The Golden Ass*? To write a parody, or something else entirely? What, exactly, were Virginia Woolf's categorial intentions when she wrote *Orlando*? Which of the passages in Franz Kafka's notebooks were meant to be just diary entries, and which parts were intended to be works of literary fiction? What were his categorial intentions when he interrupted his writing to draw sketches in these notebooks? Did these three writers successfully realize their categorial intentions, whatever they were? Puzzling questions, as Baudelaire sometimes said.

Another unjustified asymmetry in the treatment of the two kinds of intentions concerns their implications for a work's value. If a semantic intention would lead to a lower estimate of the work's value, it should be disregarded. Yet the hypothetical intentionalist does not think that unfortunate or unsuccessful categorial intentions should be similarly replaced. For example, the hypothetical intentionalist does not agree to replace Ed Wood's serious categorial intent with a more appropriate parodic one, thereby improving on his film. Why not? Perhaps because it is important to recognize that it is the artist who creates the work and its artistic value, as opposed to the critic or theorist. It seems plausible to think that the latter premise should apply to semantic intentions as well.

Another problem with hypothetical intentionalism is that the distinction is not sharp enough to be used the way this theory of interpretation prescribes. Can we in practice sort out the contents of various intentions involved in the making of works of art, deciding that some are categorial and others not, and giving constitutive status only to the former? It looks like there are some clear-cut cases where the rule can be applied fairly easily, but a lot of cases where it cannot.

The intention to make a work of art as opposed to something else may be an example of a clear-cut case, assuming we have a successful analysis of the art/nonart distinction and of the place of intentions therein. And, with regard to at least some of the more specific art forms, the same could be true. It seems plausible, for example, to say that an artist can have the intention of making a sculpture, even a representational one, without having any idea what the sculpture will represent. He intends to figure that out once he has his hands in the clay. We may even want to say that he categorially intends to develop his semantic intentions later, in which case the categorial intention is a "second-order" intention quite unlike a first-order semantic one.[15]

Are there semantic intentions devoid of categorial contents? Perhaps. Stirred by a powerful emotion of rage, the artist intends to express it, but has no other intentions with regard to how this will be done. Like Pier Paolo Pasolini, he is very versatile. Perhaps he will make a movie; perhaps he will write a political poem, or maybe he will make a sculpture depicting some angry workers. It could be, though, that the intention to make a work expressing rage must involve at least a negative categorial component: intending to communicate a sense of outrage, the artist cannot intend to make a purely decorative or minimalist work. A sonata for dog whistle probably will not do.

Even if we allow that some intentions can be sorted nearly, the friend of hypothetical intentionalism must help us deal with the messy cases, and I suspect there are a lot of them. Consider the intention to make a work of fiction (as opposed, say, to some sort of nonfictional work). On some prominent accounts, the intention to have members of an audience recognize that a text or structure has a certain propositional content is a necessary component of the larger fiction-making intention.[16] If such analyses are correct, the relevant intention is a mixed affair, having categorial and semantic components. Or consider a writer's intention to write a novel belonging to a trilogy, where the intention, more specifically, is to create various meaningful, implicit relations between the characters in the three novels. The writer intends for the readers to think about the successive protagonists "as if" they were continuations of a single type of person.[17] It seems hard to separate the categorial and semantic aspects of the content of such an intention, or cluster of interrelated intentions. Can we truly isolate the artist's goal of making works that are part of a trilogy from the meanings that in his mind constitute the links between the stories related in the three works?

Given such cases, how are we to apply an interpretive principle that instructs us to exclude "any fact about the author's actual mental state or attitude during composition, in particular what I have called his semantic intentions for a text" (I 206)? Perhaps any intention having a categorial component should be recognized, even if it is "contaminated" by semantic elements. Or should we, on the contrary, decide that a categorial intention involving a semantic component is merely "suggestive" and not constitutive? Both solutions seem arbitrary. The latter sacrifices important categorial intentions in order to screen out the semantic intentions; the former violates the clause about excluding semantic intentions. Moderate intentionalism, which places no great weight on the distinction between categorial and semantic intentions, does not face such a problem. Yet moderate intentionalism is attuned to other distinctions between intentions, beginning with the difference between intentions that are never acted on (such as an artist's aban-

doned musings about what she plans to do at some point in the future), and those that are acted on and actually orient the artist's work in a medium. In some cases, the action involved is a mental one, and what the artist does is to make a decision, thereby constituting the work in a particular manner.[18] Such intentions cannot countervene facts about the actual text or structure that has been produced, but when no such conflict obtains, they can decisively inform both the artist's and critic's thinking about the work. Such, at least, is the key claim of moderate intentionalism.

V. A Test Case

As a result of his commitment to the uptake of the actual artist's categorical intentions, Levinson's hypothetical intentionalism differs from the kind of fictionalist intentionalism that I discussed earlier, but it is not clear that his approach avoids all of its problems. We can see this if we focus on a case where the hypothetical intentionalist disregards the actual author's semantic intentions. Suppose we have a work of prose fiction where the words and sentences in the text, standardly and literally interpreted in a holistic way, are compatible with at least two significant and incompatible interpretations, each of which appears to provide an excellent, if not optimal, reading of the work. In one reading, *R1*, the governess who is the narrator of the work's embedded tale is really quite mad; it is true in the story that she wrongly believes and sincerely narrates that there are ghosts in the manor. In the other, rival reading, *R2*, of which the text also allows, this narrator is distraught, but not deluded; she detects and reports the presence of supernatural beings, and it is true in the story that these malevolent entities exist. Now, suppose as well that the author is known to have intended the latter reading. He aimed at creating a ghost story in which the horrors would be more terrible because presented only indirectly through the report of an observer. As he put it, "prodigies, when they come straight, come with an effect imperilled; they keep all their character, on the other hand, by looming through some other history."[19] That is why he penned a text in which we only hear about the ghosts through the account of the governess; the reading of that account is, in any case, clearly framed as part of an exchange of ghost stories. And it just so happens that the author's contemporaries, including such astute readers as Virginia Woolf, knew all this and never hesitated to call the work a ghost story.[20]

With regard to such a case, moderate, actual intentionalism rules that *R2* alone is right; this is a ghost story, albeit one where the ghosts are

only implicitly presented. This reading is supported by reference to the text and to both the writer's semantic intentions and the related, categorial intentions.

What does a hypothetical intentionalist say about such a case? A first problem is that of sorting out the intentions. Is the author's intention to make ghosts part of the story a purely semantic intention? What is its relation to the categorial intention of writing a certain type of story, namely, a ghost story? How could the author have framed the latter intention without also having some sort of semantic intention relevant to the presence of ghosts in the tale? It looks like a case of a single, mixed intention, not two significantly distinct aims. In that case, an interpreter working with hypothetical intentionalism has to decide whether the semantic aspect disqualifies the intention. I have not been able to find a principled basis for making such a decision (given, of course, the text's compatibility with the relevant intentions).

Suppose the hypothetical intentionalist determines that the author's intention to write a ghost story is a separate, purely categorial one. In that case he or she should agree that *R2* is correct. But let us stipulate that the facts in the case are different, and that the author had only semantic intentions about there being ghosts in the story (ghosts, that is, that would loom through some other history). We stipulate, then, that the author had no related, categorial intentions about the work being a ghost story of some kind.[21] The example is, we suppose, logically and psychologically possible: let the author be a super-Crocean having the deluded belief that his semantic and other artistic aims are unique. Now, in that case, the interpreter who follows the strictures of hypothetical intentionalism should not let the fact about the semantic intention tip the scales in favor of *R2*. The work is, instead, seen to be ambiguous between *R1* and *R2*; its meaning is their exclusive disjunction.

At this point another feature of Levinson's hypothetical intentionalism should be noted. The meanings discovered are now attributed to the actual author—only hypothetically so. Being intended, the *hésitation* has a different status than it would have had had one decided that it was an ambiguity resulting from authorial failure or ineptitude. The work is better when we abandon the latter hypothesis in favor of the former, and we are right to appreciate it as such.[22] So although I know that the actual author wrote the story with semantic intentions, I hypothesize—but do not feign or imagine—an author who did not have those intentions, but who intended for the work to be ambiguous. I do this because actual semantic intentions are not constitutive, and because the work has greater artistic value when I work with this hypothesis about the author's intentions.

So in this instance hypothetical and moderate intentionalism do not

converge; instead, they present us with a choice between two rival interpretive conclusions:

(MI) the author meant to imply the presence of ghosts in the story, and since the text can be squared with that intention, the author successfully did so, and *R2* is correct;

(HI) the hypothetical actual author's intention is a story ambiguous between the narrator's madness and sanity; the exclusive disjunction (*R1* or *R2*) and not (*R1* and *R2*), is correct.

What reasons can we given to justify a choice between the principles yielding these two options? Basically, the accounts differ with regard to what they claim about the relation betweeen the work's causal history and its meanings. Hypothetical intentionalism is selective about what facts and evidence relevant to that causal history can be decisive in an account of meaning. As the actual author's semantic intentions are not decisive (given the text's ambiguity), the hypothetical actual author is determined to be someone who intended to write an ambiguous tale. Ambiguity cannot, it would seem, be an *unintended* feature of the work's content.

Why should the possibility of recognizing unintended semantic ambiguity matter? Why, more generally, should we care about the relation between the work's causal history and its meanings? The answers rest on assumptions that are, I think, shared by friends of hypothetical and moderate intentionalism. Conclusions about meanings are often relevant to judgments concerning the artist's achievement. Someone who tries to write a straightforward, unambiguous story, but ends up writing something that everyone reads as involving a complex rhetoric of unreliable narration, may have written something fascinating to read; but this person's work should not be prized as the artistic achievement of devising an unreliable narration. We want an interpretive theory that is attuned to the difference between glorious serendipity and unfortunate failures, as well as the difference between the skillful realization of valuable and difficult aims and the routine realization of lowly or mediocre goals.[23] Ambiguity, or lack thereof, is one such relevant aim. Our appreciation of artists' achievements in this respect depends on our working with an explanation of the work's genesis that is as accurate as the evidence allows, an ideal to which the selective overruling of semantic intentions is inimical. To echo Levinson's own phrases, artists should not be allowed to make a work mean, by fiat, whatever they want them to mean; yet critics should not, by selective weighing of evidence, convert unintended meanings into intended ones, not even "by hypothesis,"

because this sort of move obscures the historical role of semantic intentions, including unsuccessful ones, in the making of the work.

Conclusion

The moderate intentionalist holds that the theory of appreciation and interpretation should be attuned to the artist's constitutive role in the making of works. It is the artist who, within "natural and logical limits," makes the work, and choosing and settling on categories and meanings is part of that creative process. We ought to reject the criticism-promoting idea that it is the reader who invents the story; we prefer, instead, a communicative model in which the reader attempts to discover the nature of the story as told, acknowledging that it is the storyteller who, within limits and contingent on his or her ability, decides what happens in the story he or she is going to tell, including events that need not, for various reasons, be related directly in the text. Hypothetical intentionalism suffers from the problem that we do not have any systematic way to separate the categorial wheat from the semantic chaff. What is more, it is not even obvious that we have any good grounds for trying to do so. Some intentions are inextricably semantic and categorial; some chaff is categorial, and there is semantic wheat to be harvested. If the works of art that actual authors have created are the prime target of an interpretive hypothesis, then we should let all of the available evidence about the causal history of the artistic structure have the same, initial status. Part of that history is a matter of the semantic intentions on which the artist has successfully or unsuccessfully acted.[24] Sometimes the author's semantic intentions are less limited than the meanings a reader may be able to dream up on the basis of the text and other background evidence. Sometimes interviews and diaries open up all sorts of wonderful undiscovered meanings. We can indeed imagine a Kafka whose diary reveals stupid semantic intentions, but we can also actually read the remarkable diaries of the real Franz Kafka. What is harder to imagine is why critics should be required to refrain from allowing their interpretations of Kafka's works to be in any way guided by an interpretation of these fascinating diaries and other evidence relevant to the actual author's thoughts and experience. Recognizing that in some cases limited or boring semantic intentions are decisive of a work's features is the price we pay for an interpretive principle that allows us, on other, happier occasions, to recognize that the artist's laudable and complex aims were decisive.

ÅRHUS UNIVERSITY

NOTES

1 For background, see *Intention and Interpretation*, ed. Gary Iseminger (Philadelphia, 1992); *Understanding the Arts*, ed. Jeanette Emt and Göran Hermerén (Lund, 1992); and Berys Gant, "Interpreting the Arts: The Patchwork Theory," *The Journal of Aesthetics and Art Criticism*, 51 (1993), 597–610.

2 For more background on intentions, see Paisley Livingston and Alfred R. Mele, "Intention and Literature," *Stanford French Review*, 16 (1992), 173–96; and Alfred R. Mele, *Springs of Action* (New York, 1992).

3 For valuable background, see Mette Hjort, *The Strategy of Letters* (Cambridge, Mass., 1993), and Lynne Rudder Baker, *Explaining Attitudes: A Practical Approach to the Mind* (Cambridge, 1995).

4 See George M. Wilson, "Again, Theory: On Speaker's Meaning, Linguistic Meaning, and the Meaning of a Text," in *Rules and Conventions: Literature, Philosophy, Social Theory*, ed. Mette Hjort (Baltimore, 1992), pp. 1–31.

5 The work by Jorge Luis Borges is translated by Anthony Bonner as "Pierre Menard, Author of Don Quixote," in *Ficciones* (New York, 1962), pp. 45–55. For some of the philosophical extrapolations, see Arthur Danto, *The Transfiguration of the Commonplace* (Cambridge, Mass., 1981); David Davies, "Text, Context, and Character: Goodman on the Literary Artwork," *Canadian Journal of Philosophy*, 21 (1991), 331–[illegible] Gregory Currie, "Work and Text," *Mind*, 100 (1991), 325–40; and Jerrold Levinson, "What a Musical Work Is," in his *Music, Art, and Metaphysics* (Ithaca, 1990), pp. 63–88. A similar point was made much earlier by Carl Lange in a little-known treatise on aesthetics, *Bidrag til Nydelsernes Fysiologie som Grundlag for en rationel Æstetik* (Copenhagen, 1899), p. 125. Lange argued that two physically identical paintings could have very different artistic values. The one produced first has the virtue of novelty and can thereby offer the special pleasure of admiration made possible by our awareness of this feature of the work, which is lacking in any subsequent copies.

6 Noël Carroll, "Art, Intention, and Conversation," in *Intention & Interpretation*, pp. 97–131.

7 George D. Painter, *Marcel Proust* (Harmondsworth, 1982); Vincent Descombes, *Proust: Philosophie du roman* (Paris, 1987); René Girard, *Mensonge romantique et vérité romanesque* (Paris, 1961).

8 Alexander Nehamas, "Writer, Text, Work, Author," in *Literature and the Question of Philosophy*, ed. Anthony J. Cascardi (Baltimore, 1987), pp. 267–91, and "The Postulated Author: Critical Monism as a Regulative Ideal," *Critical Inquiry*, 8 (1981), 131–49. For insightful criticisms, see Robert Stecker, "Apparent, Implied, and Postulated Authors," *Philosophy and Literature*, 11 (1987), 258–71.

9 The importance of multiple authorship in literary history is ably demonstrated by Jack Stillinger in his *Multiple Authorship and the Myth of Solitary Genius* (New York, 1991); Stillinger does not, however, discuss cases where several persons successfully engage in collective authorship following a shared plan or intention.

10 I say "broadly" because I intend to remain neutral on many of the controversies in contemporary pragmatics. For background, see Paul Grice, *Studies in the Way of Words* (Cambridge, Mass., 1989); Stephen C. Levinson, *Pragmatics* (Cambridge, 1983), pp. 97–166; and François Recanati, *Meaning and Force: The Pragmatics of Performative Utterances* (Cambridge, 1987), pp. 118–21.

11 See Jerrold Levinson, "Intention and Interpretation in Literature," in his *The Pleasures of Aesthetics: Philosophical Essays* (Ithaca, 1996), pp. 175–213, hereafter cited in text as I; for the discussion of the Kafka example, see pp. 184–86. A slightly different version of this essay appeared as "Intention and Interpretation: A Last Look," in *Intention & Interpretation*.

12 William E. Tolhurst, "On What a Text Is and How It Means," *British Journal of Aesthetics*, 19 (1979), 3–14.

13 See Jerrold Levinson, "Extending Art Historically," in *The Pleasures of Aesthetics*, pp. 150–71; the example is given on p. 158, n. 8.

14 For background, → Alfred R. Mele and Paul K. Moser, "Intentional Action," *Noûs*, 28 (1994), 39–68.

15 See Levinson, "Extending Art Historically," p. 158, n. 7 for this sort of emphasis on first-order and second-order intentions.

16 I have in mind the account proposed by Gregory Currie in *The Nature of Fiction* (Cambridge, 1990), ch. 1. It stands to reason that if making fiction is communicatively intending that something be imagined, the intended something has to be the content of a semantic intention. "Semantic," after all, is not the sharpest of terms, and the best we have is the old "relations between signs and that for which they stand" idea that comes to us from Charles Sanders Peirce via Charles Morris. Attempts to give a sharper demarcation of semantic/nonsemantic in terms of such notions as meaning and reference only shift the burden onto another pair of frail shoulders.

17 An example is the so-called first trilogy by the Japanese writer Natsume Soseki. For background, see Norma Moore Field, "Afterword," *And Then: Natsume Soseki's Novel* Sorekara (Rowland, Vt., 1988), pp. 258–78.

18 For example, the artist's second-order decision that her own creative work on a given piece is completed is crucial to the commonly applied distinction between unfinished and finished works of art, where the latter category includes such purposefully "incomplete" items as romantic fragments and ruins. For more on this topic, see my "Counting Fragments, and Frenhofer's Paradox" (forthcoming in *British Journal of Aesthetics*).

19 Henry James, "From a Preface," in *The Turn of the Screw*, ed. Robert Kimbrough (New York, 1966), p. 103.

20 Virginia Woolf, "Henry James's Ghosts," in *The Turn of the Screw*, ed. Kimbrough, pp. 179–80.

21 Here we probably depart significantly from the actual case of Henry James's writing of *The Turn of the Screw*. As my aim in evoking the example is primarily a matter of conceptual clarification, this does not matter to the argument.

22 Note, however, that this is not a reading whereby the governess is known to be an unreliable narrator, which would require that, in spite of some misleading evidence to the contrary, we know that she is deluded. For a valuable clarification of concepts of ambiguity and unreliability, see Gregory Currie, *Image and Mind: Film, Philosophy, and Cognitive Science* (Cambridge, 1996), ch. 9. Currie would say that the intentions in question are those of the implied author (an interpreter's construct), not the actual author.

23 To risk an analogy, in some games we keep score by judging whether the shot made was the shot called; in others, the scoring system does not filter out lucky shots. The appreciation of art is more like the former than the latter. One reason why categorial intentions are always relevant, and sometimes decisive, to interpretation is that artistic appreciation is attuned to the relation between aims and achievements. My claim is that the same reasoning holds with regard to semantic intentions.

24 Thanks to Jerry Levinson, David Davies, and Al Mele for helpful comments on a draft of this paper, a version of which was presented at the Nordic Society for Aesthetics meeting in Oslo in May 1997. Financial support for this research was provided by the Social Sciences and Humanities Research Council of Canada and by the *Fonds pour la Formation des Chercheurs et l'Aide à la Recherche* of Quebec.