

The Intentional Fallacy Reconsidered

By Lin Dong Zhang (though
I'm not completely sure)

TODAY 'the intentional fallacy' has apparently become an established critical term, for we can find it in almost all books of literary terms. Its meaning, however, has often been misunderstood since W. K. Wimsatt and Monroe C. Beardsley first introduced it in their famous essay bearing the same name as its title. In fact, there seems to be more and more people getting confused about its usage. And many fallacious ideas about this particular 'fallacy' have poured into the present-day 'trade market' of literary criticism.

Evidence of this term's confusing usage can be found in the various ways it is defined or explicated in some glossarial books. For instance, in M. H. Abrams' *A Glossary of Literary Terms*, it is simply stated that the term is 'sometimes applied to what is claimed to be the error of using the biographical condition and expressed intention of the author in analyzing or explaining a work'. In C. Hugh Holman's *A Handbook to Literature*, it is similarly said that in contemporary criticism the term is 'used to describe the error of judging the success and the meaning of a work of art by the author's expressed or ostensible intention in producing it'. But it is also noted therein that 'Wimsatt and Beardsley say, 'The author must be admitted as a witness to the meaning of his work.' It is merely that they would subject his testimony to rigorous scrutiny in the light of the work itself'. Under the entry of intentional fallacy in J. A. Cuddon's *A Dictionary of Literary Terms*, we read: 'The error of criticizing and judging a work of literature by attempting to assess what the writer's intention was and whether or not he has fulfilled it rather than concentrating on the work itself'. When the same entry appears in Northrop Frye and others' *The Harper Handbook to Literature* we are told that it refers to 'the idea that the meaning of a work can be explained by considering the author's intention, a fallacy according to the New Criticism.' And we are told that critics who emphasize the intentional fallacy 'are attempting to minimize the effect of too much reliance on Alexander Pope's advice, long standard in criticism: 'In every work regard the writer's end, / Since none can compass more than they intend'.

With these different explanations we are really confused as to when 'the intentional fallacy' may occur. Does it occur when we use biographical data to analyze or explain a work, or when we use the author's expressed intention to judge the success or meanings of his work, or when we attempt to assess what the author's intention was, or when we believe we cannot disregard the author's intention in any critical process?

In effect, it should not be too difficult to understand Wimsatt and Beardsley's own intentions in writing the essay and attacking the 'intentional fallacy.' From their

biographical data and the essay itself we can clearly see that they are posing as 'New Critics,' are positing the 'objective theory' that a literary work has an independent public existence, are encouraging 'intrinsic studies' while discouraging 'extrinsic studies' of literature, are trying to replace a system of values (covering the ideas of sincerity, fidelity, spontaneity, authenticity, genuineness, originality, etc.) with another system (integrity, relevance, unity, function, maturity, subtlety, adequacy, etc.), and are disputing the Romantic view of the author as an important source of meaning for works and they are doing all these by arguing that the author's intentions are not the proper concern of the critic.

However, not all critics agree that Wimsatt and Beardsley's intentions can rest at that. There are those who carry their propositions to extremes and assert that poems are strictly autonomous or autotelic, that works are discontinuous from language and each other, that any external evidence is critically inadmissible, that we cannot talk about an author's intention even in terms of the internal evidence of his work, etc. In brief, in attacking 'the intentional fallacy,' Wimsatt and Beardsley have begot, ironically, a number of fallacies about their intentions.

Many fail to comprehend Wimsatt and Beardsley's real intentions, I believe, because they never bother either to investigate the authors' life in question nor to examine their work closely. (They are neither traditional critics nor 'New Critics.')

Some of them may interpret the term merely by surmise or by free association of the term. Others may prefer imposing their own ideas on the term, thus asserting the authority not of the author, not of the work, but of the reader. (They are Humpty Dumpty's, who want to be masters of words.)

We know Wimsatt and Beardsley attack not only 'the intentional fallacy' but also 'the affective fallacy'; they disapprove of any attempt to derive the standard of criticism from the psychological effects as well as causes of the poem. They preach for a critical objectivism stemming from looking at the work itself while opposing any relativism coming from considering the author's intentions or the reader's impressions. This New Critical stance, I think, is basically sound and firm in that it stresses the priority of the work as the basis of critical judgment. No one, I believe, would deny that one's reading of 'To His Coy Mistress' should be done first and foremost on the text itself, and Andrew Marvell's politics, religion, and career as well as the responses or reactions of other readers to the work can only serve as an accessory guide to one's interpretation.

Nevertheless, the New Critical stance is not without its weak aspect. It has been pointed out that art does not exist in a vacuum. Any artifact is a creation by someone at some time in history. Many literary classics are admittedly autobiographical, propagandistic, or topical. Hence it would be dangerous to assume that a work of art must always be judged or looked at or taught as if it were disembodied from all

experience except the strictly aesthetic.

But my aim here is not to reiterate this frequently-pinpointed weakness of the New Criticism. My intention is to point out that the New Critical idea of 'intentional fallacy' is itself a fallacy in that it is a sort of separatism like the idea of those who are accused of 'the intentional fallacy' or 'the affective fallacy.' In my mind a literary work is hardly separable from the intention of the author who creates it nor from the intention of the reader who reads it. It is only when it is viewed as a pure physical object, a mere construct of black spots with blanks on paper entirely detached from author and audience alike, that we can say it has an 'impersonal' or 'objective' existence. Otherwise, we must admit that the work, the author, and the reader are a trinity, bound each to each with a common 'intention'; therefore, no critical effort can manage to separate them without committing a sort of 'intentional fallacy,' and to assert the absolute authority of any one of them is impractical if not impossible.

This idea of mine bears largely on the meaning of the word 'intention.' I think we must first understand that any object that comes into one's mind is an intended object. If a perceived object has any meaning at all, the meaning is never intentionless. Actually, to intend is to 'in-tend,' to turn an external object into an internal object by our mental activity. Hence an 'intention' is just an internalized version of an external object. Before an author writes a text, he usually has to 'intend' a lot of things. (In our common language, we say he has to have a lot of life experience.) The things he 'intends' include natural objects (mountains, waters flora and fauna, etc.) and man-made objects (all practical inventions and devices as well as artifacts). After he 'intends' these objects for some time, some developed 'intentions' will arise in his mind to direct his outward action. (The developed 'intentions' are commonly called 'ideas.')

If he is an ordinary man, the 'intentions' may cause him to live an ordinary man's life. If he is a wit (in the neo-classical sense) or a genius (in the romantic sense), however, his 'intentions' may lead him to create artistic works. When creating an artistic work, he is then an author. And in the process of artistic creation, he is in fact turning his 'intentions' again into an external object by the use of [a] tool (language is a writer's tool). If we can coin the word 'extention' as the antonym of 'intention' and make it mean the external object resulting from externalizing one's 'intention,' then we can hold that to read is to 'in-tend,' to form 'intentions' in the mental world while to write is to 'ex-tend,' to form 'extentions' in the physical world.

It follows then that our most important problems are whether or not an author's 'intention' is identical with his 'extention' and whether or not the author's 'extention' is identical with his reader's 'intention' of it. As we know, in the long past very few critics ever doubted the identity of an author's 'intention' with his 'extention.' That is why people could comfortably resort to studies of authors' lives and freely connect their discovered

authorial intentions with works, thus committing the so-called ‘intentional fallacy.’ After the Anglo-American New Criticism, however, people seem to become gradually aware of the discrepancy between an author’s ‘intention’ and his ‘extention.’ In Rène Wellek and Austin Warren’s *Theory of Literature*, for instance, it is said that

‘Intentions’ of the author are always ‘rationalizations,’ commentaries which certainly must be taken into account but also must be criticized in the light of the finished work of art. The ‘intentions’ of an author may go far beyond the finished work of art: they may be merely pronouncements of plans and ideals, while the performance may be either far below or far aside the mark. If we could have interviewed Shakespeare he probably would have expressed his intentions in writing Hamlet in a way which we should find most unsatisfactory. We would still quite rightly insist on finding meanings in Hamlet (and not merely inventing them) which were probably far from clearly formulated in Shakespeare’s conscious mind.

What Wellek and Warren mean by ‘intentions’ here does not much accord, of course, with my definition. Nevertheless, the above quotation makes it clear that they do think an author’s ‘extention’ can become very different from his ‘intention.’ And I think they are right in suggesting that. For a work certainly can be either above or below or even far aside the mark because of conscious or unconscious factors on the part of the author. Consequently the New Critics have reason to warn us not to rely on the author’s expressed intentions for judgment of his work. (If we should stubbornly do so, we will be laboring under the ‘intentional fallacy,’ according to their belief.)

I think the discrepancy between ‘intention’ and ‘extention’ can be best clarified with the idea of *différance*. As we know, when Jacques Derrida coined the word, he was playing on two meanings of the French word *différer*: difference — between signs as the basis of signification, and deferment — of presence by the sign which always refers to another sign, not to the thing itself. Now we can say an author’s ‘extention’ (i.e. written text or work) is a version (or copy, or transcription, or expression, or code, or record, or embodiment, or whatever else you think fit to use) of his ‘intention’ (i.e. idea). Between these two terms, there is also a semiotic relationship: the author’s ‘extention’ is the signifier and his ‘intention’ is the signified. So if here we can apply the Derridean idea of *différance*, we can say there is always difference between an author’s ‘extention’ and his ‘intention,’ though the difference may be hard to specify.

As a matter of fact, an author’s entire creative process includes both the stage of reading and the stage of writing. No one can write anything without reading something. And for an author, to read is to ‘read’ life, which includes the experience of reading books and other experiences. When an author ‘reads,’ he is building up his ‘intention’; when he writes, he is then turning his ‘intention’ into ‘extention,’ which as an external entity can be further read by others, whom we call readers. If one of the readers becomes a critic,

that is, becomes one who expresses his idea in oral or written language about an author's work, then he will indeed undergo the process of turning his own 'intention' of the author's 'extention' into his own 'extention,' which is again readable by others. (Critics' criticism can be criticized again just as a translation can serve as the basis for another translation.) Thus, if we think of the universe as a composite of things each of which is a text, then the universe is full of texts which are created, that is 'read' and 'written' all the time by various authors including God and man. When we make 'textual analysis,' we are interpreting, which often involves 'the hermeneutic circle' of repeatedly beginning with 'reading' and ending with 'writing.' And, we must remember, to 'read' is to 'in-tend'; to 'write' is to 'ex-tend.' Yet, in this hermeneutico-semiotic system, no single 'reading' ('intention') or 'writing' ('extention') has an absolute determinate 'presence,' though it is always supplementarily present in one form or another in the mental or physical world. The definite 'meaning' we seek in any text is always deferred by the alternate acts of 'reading' and 'writing,' or by the constant interchange between 'intention' and 'extention.' If we understand this, can we accuse any reader of having 'the intentional fallacy'? Can we say someone has a wrong idea when we know no idea, be it the reader's or the author's, ever exists as absolute truth or determinate presence?

A moot question of modern criticism is whether or not an objective interpretation of the text is possible. In approaching that problem, E. D. Hirsch, Jr., says that there is no objectivity unless meaning itself is unchanging. And for him the one underlying meaning of the work which does not change is the author's willed meaning, that is, his intention. For he believes that the meaning of a text 'is determined once and for all by the character of the speaker's intention'. If Hirsch's position is right, then any mode of reading is but a way of approaching the authorial intention; any study of the text, be it intrinsic or extrinsic, is but an attempt to reconstruct the unchanging intention of the author. Thus, the 'intentional fallacy' as the Anglo-American New Critics conceive it is out of the question with Hirsch.

We know Hirsch's position has been devastatingly criticized by David Couzens Hoy. In the latter's *The Critical Circle*, Hirsch is said to have committed the Cinderella fallacy (a fallacy which grows out of the dogmatic belief that if we think a thing must be there, then it is in fact there, even if it can never be seen), because he 'begins by noting that there cannot be reproducibility without determinate meaning and goes on to assert that since there is reproducibility, it follows that there must be determinate meaning'. I agree that Hirsch has committed the Cinderella fallacy in doing that logical reasoning. Nevertheless, I still think Hirsch is right in postulating the idea of a determinate meaning which is tied up closely with the authorial intention. Although it is theoretically true that there is no dependable glass slipper we can use as a test, since the old slipper will no longer fit the new Cinderella, yet, we can suppose that a short lapse of time should not bring about

change appreciable enough in Cinderella's feet to render impossible our recognition of the true girl. Paradoxically, our senses are not keen enough to perceive any minute change in objects so as to hinder seriously our sense of identity. Theoretically, the idea of *différance* is right: the author's 'intention' is never identical with his 'extention,' and his 'extention' is never identical with the reader's 'intention' of it. Yet in practice the author often so succeeds in making his 'extention' accord with his 'intention' that we can say what a text means and what its author intends it to mean are identical. And the reader often so succeeds in 'in-tending' the author's 'extention' that he can feel his 'intention' of it is equivalent to, if not identical with, the author's 'intention.' Theoretically it is true that every interpretation is a misinterpretation. Yet, in practice every interpreter believes he has made a right interpretation and his belief is often justifiable in terms of our common understanding or consensus. The truth is, understanding literature is like understanding life: no one can claim his understanding is the only true understanding, but all can agree on an understanding as the valid understanding within a certain space and time. In other words, reading (and writing as well) is a social behavior. And any social behavior is a matter of common agreement, not a matter of scientific truth.

In his 'The Deconstructive Angel,' M. H. Abrams grants that Jacques Derrida's and J. Hillis Miller's conclusions are right, but he still believes in the three premises of traditional inquiries in the human sciences.

The basic materials of history are written texts and the authors who wrote these texts (with some off-center exceptions) exploited the possibilities and norms of their inherited language to say something determinate, and assumed that competent readers, insofar as these shared their own linguistic skills, would be able to understand what they said.

The historian is indeed for the most part able to interpret not only what the passages that he cites might mean now, but also what their writers meant when they wrote them. Typically, the historian puts his interpretation in language which is partly his author's and partly his own; if it is sound, this interpretation approximates, closely enough for the purpose at hand, what the author meant.

The historian presents his interpretation to the public in the expectation that the expert reader's interpretation of a passage will approximate his own and so confirm the 'objectivity' of his interpretation. The worldly-wise author expects that some of his interpretations will turn out to be mistaken, but such errors, if limited in scope, will not seriously affect the soundness of his overall history. If, however, the bulk of his interpretations are misreading, his book is not to be accounted a history but an historical fiction. For me these premises are correct and I think their rationale can hardly be challenged. Furthermore, I think the premises can dovetail perfectly into the three conclusions reached by Stanley Fish, who as we know is one of the 'New readers' Abrams accused of being apostles of indeterminacy and undecidability. In the end of his *Is There*

a Text in This Class? Fish says:

We see then that

1) communication does occur, despite the absence of an independent and context-free system of meanings, that

2) those who participate in this communication do so confidently rather than provisionally (they are not relativists), and that

3) while their confidence has its source in a set of beliefs, those beliefs are not individual-specific or idiosyncratic but communal and conventional (they are not solipsists).

And then he adds that ‘the condition required for someone to be a solipsist or relativist, the condition of being independent of institutional assumptions and free to originate one’s own purposes and goals, could never be realized, and therefore there is no point in trying to guard against it.’ And so he thinks it unnecessary for Abrams, Hirsch, and company to ‘spend a great deal of time in a search for the ways to limit and constrain interpretation’.

Here we have arrived at the point where we can summarize the ‘intentional fallacies’ we have discussed so far. The traditional critics commit ‘the intentional fallacy’ because they rely too much on the author’s intention (especially the expressed intention) for interpretation or judgment of his work. The New Critics try to rectify this ‘fallacy’ by calling our attention to the fact that the work is the primary and ultimate ground on which we can base our interpretation or criticism. They forget, however, that no work can be really detached from either the author’s intention or the reader’s. They do not admit that a poem in fact cannot have a really independent public existence. To think that one can ignore the author’s or the reader’s intention in criticizing works is itself an ‘intentional fallacy.’ More recent critical theories have tried again to modify and correct the New Critical position. From that effort, however, two new types of ‘intentional fallacy’ have arisen. On one hand, we have such critics as Hirsch, who try to bridge the author and the work by locating the work’s meaning again in the author’s intention (both explicit and implicit). These new author-oriented critics are right in pointing out a determinate ground for readings. But they overlook the fact that the meaning of a work is decided not by the author’s intention alone; it is equally if not chiefly decided by the reader’s intention as well. So they also have some ‘intentional fallacy’ of their own. On the other hand, we have such reader-oriented critics as Fish, who in attacking ‘the affective fallacy fallacy’ and trying to see literature only in the perspective of the reader, have themselves committed the ‘intentional fallacy’ of neglecting the authorial intention. As I have suggested above, the author, the work, and the reader are a trinity hardly separable. The work is only an ‘extension’ where the author’s intention and the reader’s meet. We cannot assert that the author’s intention is completely identical with the reader’s intention

through the 'extention.' Still we must agree that insofar as communication is possible, there should be a considerable amount of sameness remaining between the author's intention and the reader's through the medium of the 'extention.' To emphasize the difference (as Derrida and others do) may be logically or metaphysically correct; it is, however, impractical as Abrams has feared. The deconstructionists, in fact, have committed an 'intentional fallacy' as well: they turned intended objects into intentionless objects by reducing all things to signs and accounting for an intentional process in terms of non-intentional semiotic relationship.

In their 'Against Theory,' Steven Knapp and Walter Benn Michaels conclude that the theoretical impulse always involves the attempt to separate things that should not be separated: on the ontological side, meaning from intention, language from speech acts; on the epistemological side, knowledge from true belief. Our point has been that the separated terms are in fact inseparable. ... Meaning is just another name for expressed intention, knowledge just another name for true belief, but theory is not just another name for practice. It is the name for all the ways people have tried to stand outside practice in order to govern from without. Our thesis has been that no one can reach a position outside practice, that theorists should stop trying, and that the theoretical enterprise should therefore come to an end.

I think these conclusions are true in their own right, especially when, as we know, they are directed against such 'objectivists' as Hirsch on one hand and such proponents of indeterminacy as Paul de Man on the other. However, as people cannot stop 'in-tending,' it is of no avail to try to stop them from theorizing. For a 'wise' man, to 'in-tend' is to theorize.

Some twenty years before the publication of 'Against Theory,' Susan Sontag brought forth her essay 'Against Interpretation.' In it she proposes that the function of criticism 'should be to show how it is what it is, even that it is what it is, rather than to show what it means.' And her famous conclusion is: 'In place of a hermeneutics we need an erotics of art'. I think Sontag, in making the statements, has probably understood, like an existentialist, that we are forever living in a meaningless absurd world where to interpret (i.e. to seek meaning) is always vain. Yet, I presume she has not understood that from time immemorial wise men have been interpreting and theorizing (to interpret is also to theorize); they think only vulgar men can be content with living an 'erotic' life. To forbid people to interpret is tantamount to forbidding them to think, to 'in-tend.' And that is morally the biggest 'intentional fallacy.'

In his *The Critical Path*, Northrop Frye wishes we could avoid 'two uncritical extremes': the 'centrifugal fallacy' which feels that 'literature lacks a social reference unless its structure is ignored and its content associated with something non-literary,' and the 'centripetal fallacy' where we 'fail to separate criticism from the pre-critical direct

experience of literature'. In the critical path, in fact, we are perpetually making fallacies of all kinds. We are so many blind men feeling the same elephant. Our insight indeed comes from our blindness. But all fallacies can be reduced and traced to a single fallacy, namely, 'the intentional fallacy,' which is the necessary result of 'in-tending' anything. When we understand this, we may be willing, then, to forgive anyone who fails to understand others' 'intentions,' including those who we know have distorted Wimsatt and Beardsley's idea of 'the intentional fallacy.' For 'to err [through intention] is human; to forgive [with intention], divine.'