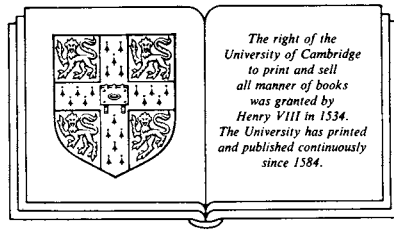


LITERATURE AND RATIONALITY

Ideas of agency in theory and fiction

PAISLEY LIVINGSTON

Professor, Department of English, McGill University



CAMBRIDGE UNIVERSITY PRESS

Cambridge
New York Port Chester
Victoria Sydney

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Victoria 3166, Australia

© Cambridge University Press 1991

First published 1991

Printed in Great Britain at the University Press, Cambridge

British Library cataloguing in publication data

Livingston, Paisley
Literature and rationality
1 Literature. Rationality.
1. Title
801

Library of Congress cataloguing in publication data

Livingston, Paisley, 1951-
Literature and rationality / Paisley Livingston.
p. cm.

Includes bibliographical references and index.

ISBN 0 521 40540 8

1. Literature – Philosophy. 2. Knowledge, theory of.
3. Literature and science. 4. Science – Philosophy. 5. Reason.
1 Title.

PN49.L545 1991

809'.93384-dc20 90-22179 CIP

ISBN 0 521 40540 8 hardback

Contents

<i>Acknowledgements</i>	<i>page</i> x
Introduction: literature and rationality	i
I THEORIES AND QUESTIONS	
1 Rationality: some basic issues	15
2 Agency, rationality, and literary knowledge	48
II TEXTUAL MODELS	
3 Naturalism and the question of agency	89
4 Agent's rationality	123
5 Plans and irrationality	150
6 Science, reason, and society	179
Coda: 'Der Bau'	205
<i>Notes</i>	210
<i>Bibliography</i>	227
<i>Index</i>	251

Introduction: literature and rationality

The idea that 'man' may be defined as the rational animal is surely an outdated philosophical fiction, a singularly Western prejudice, perhaps even *the* Western and sexist prejudice par excellence. Does not the evidence of history – or of even a moment's introspection – flatly refute the notion that human beings are rational? Rationality is the 'crystal palace' mocked by Dostoevsky's underground man; real people live in Babel. We can no longer believe in the grand old thesis that human history as a whole is animated by Reason; nor do we have any good reason to think that the preferences of even the most enlightened and lucid modern individuals correspond to the models of rationality invented by neo-classical economics, such as the subjective expected utility model and the theory of rational expectations. It would seem to follow that the concept of rationality does not embrace much of either collective or individual reality, and should therefore be abandoned.

Moreover, even if one were to grant that rationality is a notion having some restricted value in the human sciences, this would not necessarily imply that it has any great relevance to literature. Rationality may seem a particularly inappropriate concept to bring into a discussion of literature: the prevailing tendency today is to associate literature with madness, dreams, and passion, not with reason. The *homo sapiens* of the sciences, then, is contrasted to the *homo demens* of literature, particularly in the age of romanticism. Thus, although in bibliographies of philosophy and the social sciences, 'rationality' and cognate terms are headings for countless entries, in the comprehensive bibliographies published annually by the *Modern Language Association*, the subject heading 'rationality' is followed by a mere handful of items.

It seems clear, then, that the conjunction of the terms 'rationality' and 'literature' in my title is likely to evoke two kinds of

questions. One concerns the veracity and importance of the very notion of rationality; the other challenges the relevance of this concept to literature. Responding to these two issues in this introduction will allow me to set forth some of the main assumptions and goals of the present study.

A review of some of the most salient facts about human history, as well as a few moments of honest introspection, should suffice to demolish any belief that people have ever inhabited the crystal palace of perfect rationality – or that they are ever likely to do so. This is a very important point, the implications of which are anything but trivial. Yet the very same evidence also supports the idea that it would be a mistake to describe human beings as creatures who totally fail to satisfy any and all norms of rationality. People are not perfectly rational, it is true, but nor is their behaviour totally incoherent or random. Being neither omnipotent nor impotent, they sometimes behave in ways that effectively realize their goals; being neither omniscient nor totally ignorant, they sometimes achieve a measure of knowledge. This domain of possible knowledge and effective action, a domain that falls between the extremes just evoked, is explored by the various theories of bounded and moderate rationality that have been developed in the human sciences.¹ Rationality, then, is not a concept that we want to discard, even if we think it important to criticize the unrealistic assumptions that make up *homo oeconomicus* and other overly idealized social science fictions.

Arguments for the value of concepts of rationality can be based on a number of theses that are not subject to any reasonable, non-sceptical forms of doubt. For example, it is plain that in their daily affairs, women and men constantly interpret themselves and each other as persons having conscious desires, beliefs, and intentions to act. People think that it is often possible to distinguish between cases where someone does something intentionally and cases where behaviour is not under intentional control.² Moreover, these attributions of mental states to other agents are not tentative and purely speculative, but involve serious practical commitments. Every time we take a flight in an aeroplane, or step into the crosswalk before an approaching automobile when the traffic signal has just given us the right of way, we are implicitly wagering that other people's behaviour will conform to any number of complex con-

straints on perception, inference, and action. For example, we may be tacitly assuming that the other person's beliefs, desires, and actions will conform to a 'practical syllogism' of the following form: if you want a certain state of affairs to obtain and believe that a certain action will bring it about, and if you moreover have no good reasons for not doing so, then you should perform this action. In relation to the example of the pedestrian standing at the crosswalk, this schema could be applied roughly as follows: the pedestrian wants to cross the street and assumes that the driver of the approaching automobile has a certain amount of practical knowledge about traffic regulations and about how to control an automobile; the pedestrian can also ascribe to the driver a basic desire to avoid an accident, which means that the driver is most likely to want to stop and let pedestrians cross the street when they have the right of way. This ascription of beliefs and desires leads to the prediction that the driver will bring the car to a stop only a few feet away from the white line.

It is an indisputable fact that human beings are often able to predict each other's behaviour: witness the many instances where people manage to keep a rendezvous with each other, meeting at the agreed place at the right time, in spite of the complex contingencies involved. How are such things possible? It is clear that people do not manage this sort of everyday co-ordination of behaviour by thinking about each other's neurophysiological processes, neural nets, or the stimulus-response correlations postulated by behaviourist psychology. Nor do they achieve these remarkable feats of co-ordination by means of radical philosophical thinking about the status of this or that morsel of *Dasein*. They do not reckon in terms of social totalities, rhetorical systems, discursive formations, or 'desiring machines'. The way people in fact manage to understand and predict each other's behaviour is by relying on certain kinds of attributions of belief, intention, and desire. In other words, they reckon in terms of what the other person may be assumed to think and want. Moreover, as I shall argue below, this kind of intentional explanation of other people's behaviour relies on some very basic assumptions about the rationality of the agent: if we did not assume that there were any regularities or systematic connexions of some sort between a creature's behaviour and some of the intentional attitudes that we might attribute to that creature, then there would be no point in trying to

explain this being's behaviour in terms of such attitudes. Thus, assumptions about rationality come with, and are in fact presupposed by, the general framework of intentional explanation, or in other words, the language of agency, with its talk of action guided by reasons.

My argument, then, is that we have good reasons for believing in the validity of working within the very general framework of intentionalist psychology, and that assumptions about rationality are a necessary and important part of that framework. Yet it is by no means being proposed here that intentional explanations of individual action provide a satisfactory total explanatory framework for the human sciences – far from it. Some of the explanatory limits of this perspective will be discussed below. Moreover, it is not my contention here that the framework of intentional understanding gives us some kind of infallible knowledge of others. Attributing rational connexions of some sort to people's beliefs, desires, and actions may be reliable a lot of the time, but the predictions and explanations we produce in this way remain fallible, and there are many cases where they go wrong. Quite obviously, I may attribute a certain intention to another party and thereby profoundly misunderstand her action, which was in fact oriented by quite a different attitude. Or again, I may misunderstand someone's behaviour by wrongly assuming it to be linked to a particular attitude, when in fact no attitude whatsoever was responsible for this behaviour. Yet when one intentional explanation falters, we often try to make sense of what happened by continuing to think in terms of rational agency, looking for another, more viable intentional explanation. For example, if the driver sped through the intersection instead of stopping, we are likely to ask ourselves – that is, if we were lucky enough to survive – what other goal the person was trying to pursue in driving in such a manner. We look for some other intentionally meaningful connexion between beliefs, desires, and action, asking what kind of beliefs and desires could have made the behaviour seem appropriate to the agent. 'Where's the fire?' is a folk saying that is sometimes uttered on such occasions, suggesting that the motorist's criminal behaviour should be compared ironically to the fireman's legitimate and purposeful dangerous driving. It may be, then, that the conceptual framework relative to rational action – very broadly defined – is a heuristic that we keep on employing until we

encounter evidence that it simply does not apply to the case at hand. For example, we look for the driver's reasons for ignoring the traffic signal until we learn that she suffered a stroke and lost control of the vehicle as a result of purely physical causes.

In short, some very general and basic assumptions about rationality may constitute a privileged but fallible heuristic, which amounts to the operating assumption that at least some of the time we will be able to understand and explain people's behaviour in terms of their reasons – the beliefs, desires, and intentions that guided their activity. This rationality heuristic is privileged for the following reason: even if we knew for certain that human beings most often diverged from the constraints of rationality – and this even on some fairly mild definition of those constraints – it would not follow from this fact that assumptions about rationality could not be an important conceptual tool; they could still be an indispensable method of exploring human behaviour, a method we keep trying to employ as long as we grant human organisms the status of agents. The rationality heuristic is fallible because human rationality is indeed limited and imperfect: our thinking about ourselves and others is neither omnipotent nor omniscient.

I turn now to the second objection evoked above. Are the kinds of issues that surround pedestrian notions of practical reasoning particularly relevant to literature? Is it not possible to grant what has just been said about certain concepts of rationality while denying that they play any important rôle in literature and its analysis? Responding to this complex question is the object of chapter 2, where I delineate a number of different ways in which conceptions of agency and rationality are relevant to literature and to literary scholarship. Basically, I defend two theses. The first is that it is highly unreasonable to deny that at least a significant subset of literary phenomena are purposeful activities comprehensible in terms of the rationality heuristic that I have just evoked, and it follows that assumptions about rationality can play a rôle in at least some forms of literary enquiry. I also present some arguments in support of a much stronger thesis, which is that assumptions about agency and rationality are in fact essential to *all* literary phenomena and hence to all adequate literary enquiries.

Some of the broad strokes of the former contention can be conveyed in the present context by referring to a particular example,

which will also help to give the reader a better sense of what is meant by my references to a rationality heuristic. In *Egil's Saga*, an Icelandic family saga written around 1230, the modern reader encounters a description of some behaviour that is likely to be a bit puzzling.³ The protagonist Egil, we read, has learned that his father Skallagrim has passed away during the night and has been found 'sitting on the edge of the bed, dead and so stiff that they could neither lift him nor straighten him out'. What does Egil do when he arrives on the scene? Stretching his father's corpse out flat, he plugs the nostrils, ears, and mouth with wax. The text continues:

Next he asked for digging tools and broke a hole through the south wall. When that was done, he got hold of Skallagrim's head while others took his feet, and in this way they carried him from one side of the house to the other and through the hole that had been made in the wall. Without delay they next carried the body down to Naustaness, and pitched a tent over it for the night. The following morning, at high tide, Skallagrim was put into a boat and taken out to Digraness. Egil had a burial mound raised there on the tip of the headland, and inside it Skallagrim was laid with his horse, weapons and blacksmith's tools.

As is characteristic of the narrative style of the sagas, this passage gives the reader descriptions of actions, but provides no explicit information about the actor's motives, goals, intentions, or other states of mind. Egil makes a hole in the wall and carries his father's body out through it, but no reason for this behaviour is explicitly given. How are we to understand this? One way in which a reader can understand the passage is by trying to fill in the missing intentions. In other words, we try to connect the particular gestures that Egil performs to a meaningful plan or scheme of action, the assumption being that some such connexion was implicit in the narrative. We make sense of the behaviour reported to us by linking it to attitudes, such as belief and desire, that might have motivated it. Since these attitudes are not overtly stated, we can only infer their presence, and in so doing, readers must draw on their own background and contextual knowledge. In the case of the saga, modern readers may lack the background knowledge that would make it possible for them to find a plausible connexion – especially if they read an edition without notes and lack the advice of a specialist. Why does Egil break a hole in the wall of his father's house and remove the corpse through it, instead of simply carrying

him out through the door? Why does he block the corpse's nostrils, mouth, and ears? These may seem like totally inappropriate or senseless things to do. Some modern readers may be led to conclude that Egil is not at this point in the story a plausible character at all, but a marionette who has been made to move about crazily by the author, in which case they may try to infer what the author's reasons for writing the story that way could have been. And indeed there are many points in the sagas where readers are wise to give up looking for plausible psychological motivations of a character's actions and to think instead in terms of the other kinds of effects the writer may have been after. Following one well-informed hypothesis about the pragmatic context in which the Icelandic family sagas were composed and read, it is simply wrong to assume that their authors were always trying to convey undistorted truths about people and their psychology.⁴ Yet it is often possible to make some plausible inferences about the other sorts of desires and beliefs that guided the composition of a saga.

Many aspects of Egil's behaviour are readily intelligible to most readers: in all cultures people typically bury their dead and perform some kind of funeral ceremonies, and this is clearly what Egil and his helpers are doing when they erect a mound, even though the Viking trappings may seem somewhat peculiar (one may well wonder what motives there were for inhuming the corpse with a horse, weapons, and blacksmith's tools). Yet Egil's other actions are not in fact totally bizarre or senseless; they have at the very least an internal intelligibility that can be grasped once the reader possesses certain background information. Breaking the hole in the wall makes sense once we know that it was thought that ghosts returning to haunt a house could only enter it the way they came in. Although the text does not say so, we may assume that Egil plans to repair the hole in the wall once he has taken the corpse out, thereby preventing his father from getting back into the house later. The action of breaking the hole in the wall thus belongs to a larger plan that links a present intention to a decision to do something else in the future, the actions being related in a structure of means and ends. Similarly, Egil's other gestures involve planning and problem solving. A corpse's orifices were blocked because it was thought that this would prevent the dead man's spirit (*hamr*) from getting out.⁵ Thus the actions described by the text become intelligible against a background of religious beliefs. Knowledge of these

beliefs makes it possible for readers to grasp the nature of the fictional character's implicit practical reasoning in a particular situation. The religious beliefs in question form a complex network that includes such items as a distinction between two ways of being dead – one in which the deceased party undertakes the journey to *Hel*, the other in which he comes back from the grave to pester the living. The network of beliefs includes various ideas about how one may effectively deal with a ghost (*draugr*) – such as the method adopted by Egil. One of the most striking features of the network of beliefs in question is the fact that in the Old Norse judiciary system, it made perfect sense to file a legal suit against a corpse.

The example shows how a very basic conception of rationality can play a rôle in the intelligibility of narrative: we understand crucial aspects of the narrative by thinking in terms of the character's practical reasoning. More specifically, we fill in meaningful relations between the character's actions and such intentional attitudes as beliefs, desires, emotions, and intentions. In the example at hand, the desire is a matter of Egil not wanting his father's ghost to come back and haunt him; the belief is the idea that carrying the body out through the window and fixing the hole in the wall afterwards will prevent the ghost from returning; Egil forms an intention to perform a certain action as part of a plan, and acts on it – all of which follows appropriately from Egil's desires and beliefs. Thus, what may seem like a strange and irrational superstitious practice can be understood as manifesting a basic syllogism of practical rationality: very generally, a certain course of action follows from the actor's beliefs, desires, and intentions. And the practical syllogism that is implicit in the action is also implicit in our understanding of the description of the action: given the gestures, we infer the plan (or means–end schema) of which they are a part.⁶ It would seem, then, that understanding narrative depictions of action requires that we apply a basic rationality heuristic, meaning simply that the reader makes inferences about coherent and meaningful connexions between an agent's actions and intentional attitudes.

A number of contemporary literary critics and theorists do not share the view that the language of rational agency is directly pertinent to our comprehension of literary phenomena. Roland Barthes expressed the contrary point of view when he wrote that 'in narrative, what deliberates over action is not the character, but the

discourse'.⁷ Barthes grants that what he calls the 'code' of action and psychological verisimilitude is an important part of literary discourse, but he insists that the determinations of this discourse are none the less not intentional, that is, not to be understood in terms of the practical reasoning of any agent or agents. Yet such claims are dubious. One may very well doubt that a discourse can really do any deliberating, and I return to this objection below. Yet suppose we accept the idea that what deliberates over action is the discourse and not the character. It would still be the case that in order to understand the discourse's deliberation, readers would have to grasp what this discourse says about the character's deliberations, which means they would need to apply a rationality heuristic. For example, if we are to understand Oedipus as a mere pawn or function of the *muthos* of tragic discourse, we must first of all think about him as an agent of purposeful action and try to understand him in such terms. Only then is the reader in a position to measure the distance that is created in the discourse between the intentions and the real consequences and causes of the tragic hero's actions. To eliminate any and all reference to Oedipus's plans, beliefs, and desires would make it impossible to identify tragedy's *dissoi logoi*, which contrast the character's attitudes and efforts to the other conditions and forces that determine what happens to him. More generally, there is good reason to believe that in all discourses where there is any form of psychological verisimilitude (which means all or almost all literary works, and certainly all genuine narratives), making sense of the discourse requires, at the very least, that the reader take some fairly long detours through an understanding of what the characters are doing *qua* intentional agents (albeit fictional ones). Developing intentional understandings of the characters' actions, then, is a necessary (but not a sufficient) condition of the comprehension of most literary works. If assumptions about rationality are a necessary feature of all intentional understandings – a point to be defended in chapter 1 – then it follows that these same assumptions about rationality play an important rôle in the understanding of literature.

In chapter 2 I also advance arguments supporting a stronger thesis, namely, that concepts of agency are necessary to all forms of literary enquiry. Briefly, I contend that there are good reasons to think that it is impossible to understand and explain literary phenomena adequately without relying on the assumption that

intentional agency is not reducible to textual processes and codifications. In this regard it is significant to note that in spite of Barthes's explicit proclamations in favour of a thoroughgoing textualist stance, there is still a psychology subtending the theory of discourse and codes that he advances in *S/Z* and other texts – Lacanianism, which is a psychology that tries to reduce agency to language. It is anything but self-evident that this attempt is successful. In chapter 2 I argue that it is a mistake to assume that texts, discourses, and works of art are autonomous entities or functional systems capable of engaging in effective practical interventions in the world; it follows that it is literally impossible for a discourse to 'deliberate over action' or to engage in any other sort of purposive activity. I also explore arguments to the effect that the very identification of symbolic artefacts as such depends on a number of important assumptions about the activity of sentient agents, beginning with the writers whose efforts are indispensable to the existence of literary works. Critics' ability to individuate literary works of art may require reference to the action of relevant agents and not merely to features of texts. One of the central theses of the theory of minimal rationality is 'no rationality, no agent', and I propose and defend a related dictum: 'no agency, no textuality'.

The two lines of argument that I have just evoked jointly suggest that critics can hardly afford to dispense with the rationality heuristic: concepts of agency and rationality may very well play a fundamental rôle in all understandings and explanations of literary phenomena, and they *certainly* play such a rôle in many of them. A point that should be stressed here is that my claim in what follows is not that applying a rationality heuristic to the actions depicted in texts is all that there is to our understanding and appreciation of literary and other narratives – for indeed it is not. What is more, it is not my assumption that literary enquiry should be limited to the appreciation and explication of literary works, for even when these goals are successfully achieved, other aspects of literary history remain to be explained. In what follows I shall argue that concepts of rationality can play a valuable rôle in a number of distinct varieties of literary enquiry, no one of which should be taken as the sole end of the literary disciplines. What is more, although I shall be arguing that the rationality heuristic has a privileged status in relation to literary-historical enquiry, this status does not entail the erroneous view that everything in literary history can be explained

in terms of the conscious intentions and practical reasoning of individual authors. My stress on concepts of agency and rationality does not amount to a rehabilitation of what is known as the 'intentional fallacy'.

There is another way in which I motivate my conjunction of the terms 'literature' and 'rationality' in the present study. Not only can we ask how concepts of rationality can be applied to the understanding of literature, but it is also possible to ask how literary works illustrate, challenge, and complexify various conceptions of rationality advanced in the social sciences. What contributions can the critical analysis of literary works make to the formation of hypotheses about the patterns of rationality and irrationality in human action and interaction? In what ways can the reading of literary texts lead to the refinement of contemporary hypotheses about rationality, irrationality, and the framework of intentional explanations? Such questions lead us to explore the value of the analysis of literature within the context of a research problematic of central importance to the human sciences. In such an approach, literature is not an 'object' to which a social scientific concept is applied in cookie-cutter fashion. Instead, 'rationality' and cognate terms are better understood as standing for an important constellation of issues and hypotheses, the value of which is to open up lines of enquiry. One of my central goals in what follows is to analyse aspects of literary texts with the aim of contributing to our understanding of the kinds of questions raised by rationality, irrationality, and intentional explanations. I am seeking, in this manner, to draw on the heuristic value of the critical analysis of literary works.

In the second part of this book, then, I interpret a number of literary works with the aim of working on questions pertaining to hypothetical models of rationality and agency, a particular point of focus being motivational concepts. I begin in chapter 3 with a discussion of basic questions about agency and rationality raised in the naturalist fiction of Theodore Dreiser, arguing that his documents clarify the shortcomings of reductive views of motivation. In spite of Dreiser's proclamations to the contrary, the models of agency in his narratives grant a significant rôle to complex intentional attitudes in the generation of motivational states. Chapter 4 mobilizes a reading of aspects of Dreiser's *Sister Carrie* in relation to a distinction between atomistic and agential models of rationality. The same text makes possible a clarification of the question of the

rationality of desire, and allows me to trace some of the limits of intentionalist explanations. In chapter 5, I discuss Émile Zola's *La Joie de vivre* in terms of the place of long-term planning in a theory of agent's rationality. Reference to the complex episodes presented in this narrative makes it possible to propose a hypothesis about a source of irrational planning, namely, a species of social influence in which the agent adopts an unreasonable imitative strategy. Finally, chapter 6 discusses Stanislaw Lem's *His Master's Voice* in relation to the issue of the rationality of social situations, focusing in particular on the text's depiction of the epistemic and practical rationality of a scientist engaged in a collective research project.

In proposing these particular literary readings, it is not my ambition to exhaust the meanings of the works involved or to survey all possible literary texts that are potentially relevant to the theoretical questions being raised. Even less do I claim to have written comprehensive author studies or a detailed and realistic specimen of literary history. In my view these are complementary avenues of enquiry that move on different tracks. Rather, my aim is to explore the heuristic value of a number of texts in relation to problems raised by a few central concepts pertaining to rational agency. If my choice of texts and topics manifests any form of rationality, it is of the bounded and minimal variety.