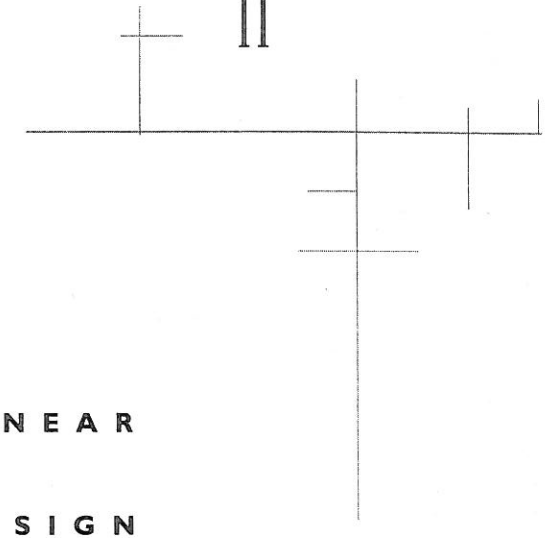
NARRATIVE DESIGN

A WRITER'S GUIDE TO STRUCTURE

MADISON SMARTT BELL

W.W. NORTON
& COMPANY
NEW YORK
LONDON

PART II



LINEAR DESIGN

SUPPOSE THAT THE ELEMENTAL INGREDIENTS of fiction may be grouped in one or another of four major categories: plot, character, tone, and form. To define these terms quickly and simply: plot is what happens in a narrative; character is who it happens to (or who makes it happen); and tone is what it sounds like. Form is the pattern of its assembly, its arrangement, structure, and design.

Form is the aspect of a story that can be abstracted from everything else and expressed in some other medium, for instance, a graph, or some other geometrical figure. Not that a recreational reader would be likely to need or want to undertake this procedure—no more than you'd want to get to know your pet cat by dissecting it. But for writers it is sometimes (not always) necessary to perform such a maneuver of abstraction so that form can be rendered in ways that yield to analysis.

Form is of primary importance, always. Ingredients of fiction from the other three groupings (regardless of appearances, which may often be to the contrary) are always subordinate to form, to design. Indeed, any or all of these ingredients can and do function as elements of design. We are accustomed to thinking of plot as what defines structure in a story. But elements from other categories—point of view, imagery, shifts and alterations of tone—may also be used structurally and often are.* In reading and writing, you must consider (consciously or unconsciously) all of these aspects of fiction in terms of their relationship to the overall design. Even the overall meaning or theme of the narrative cannot be separated from this relationship. In a properly realized work, form and function are one and inseparable.

For the writer, some sense of the final formal design of the work really ought to precede the first stages of composition. The level of prior refinement of this sense of design will vary wildly from one writer to another. It may be quite specific and detailed (though it is risky for it to become *too* specific and detailed in the early stages—lest you create a paint-by-the-numbers design whose execution will suck the life from your conception). Or it may be no more than a vague and cloudy sense of where the story is headed—where *you* are headed, across the terrain of the story.

The length of the narrative being contemplated has a good deal to do with how evolved the writer's idea of its form needs to become before it is written. Most writers can navigate their way through a short story on sheer intuition at

*For the structural use of imagery, see the discussions of "Depth Charge," p. 49, "Daisy's Valentine," p. 120, and "O Man Alive," p. 157. For the structural use of characterization, see the discussion of "The Sky Is Gray," p. 189. For the structural use of tone, see the discussion of "Red Hands," p. 284.

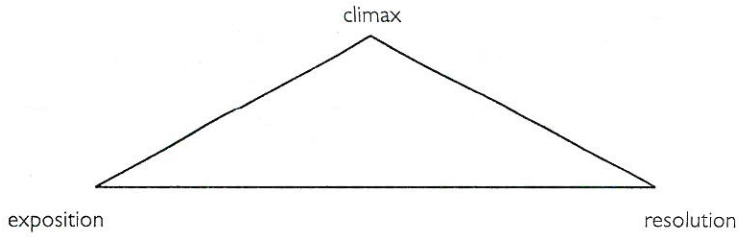
least some of the time: write a story successfully clean through without a deliberate, conscious plan—flying blind, as it were, and without frequent reference to the instruments either. In this situation, the writer discovers the *form of the story in the process of writing it*, just as the reader discovers the story's form in the process of reading it. This sense of discovery has much to do with the pleasure of reading, and for the writer who can work in the analogous way, it can truly be an ecstatic pleasure, akin to Hopkins's inscape or Joyce's epiphany. There's no better thrill, in this business, than to realize your intention at the very moment you write the last line. What makes it all possible, however, is the *unconscious* apprehension of an underlying structure. Without that, you'll become confused and lose your way.

Anyone who's ever grappled with a longer narrative, something approaching the length of a novel, say, will have discovered (quite painfully, perhaps) that sheer intuition won't carry the project all the way through. At least not successfully on the first attempt. To end up with a first draft of a novel that is structurally sound, you must do *some* structural planning in advance. Without it, you end up with an anarchic mass of material that must be arduously rewritten toward some sort of formal coherence. At the opposite extreme is the risk that excessive structural planning, prior to the actual writing, will overdetermine the work before it is realized and leach the life out of it.

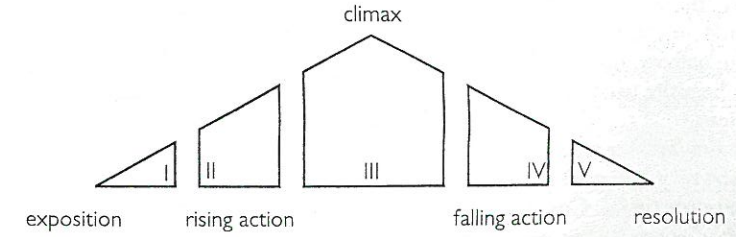
To steer a safe course between these two shoals is a demanding undertaking. In practice, most writers actually zigzag back and forth between them. Some writers can tolerate a very high level of detailed advance planning for a long work without losing their own interest and sense of discovery in actually writing it. Others are so differently constituted that they cannot tolerate any abstract advance planning at all and must proceed through novels as intuitively as they would through short stories (with the result that they suffer more and have to write a lot more drafts).

But for structural purposes, there's really no *essential* difference between a novel and a short story. The only difference is size, which means that while a short story potentially can be written in a single inspired sitting, a novel absolutely can't be. One's intuitive idea of a novel's design must be propped up with some sort of scaffolding, in order to last out a longer period of composition. But the fundamental principles of a narrative's design are apt to be much the same, regardless of scale.

THERE ARE MANY POSSIBLE STRUCTURES for a narrative, but the most common, familiar, and conventional of these is linear design. Linear stories start at the beginning, traverse some sort of middle, and stop at the end. Furthermore, all linear designs bear some relationship to what is known as the Freitag triangle.

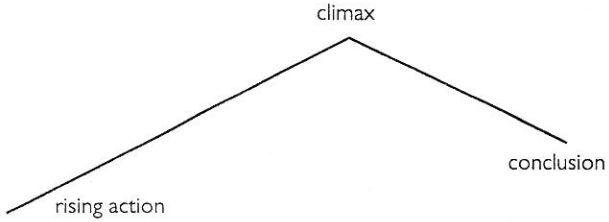


This diagram will be familiar to many students and teachers as an instrument, like a protractor, for stripping the life and interest from Shakespeare's plays. The triangle can be divided to correspond to the five acts of Elizabethan drama. According to this reading, Act I is responsible for the "exposition" (the establishment of principal characters and the situation that obtains at the opening of the narrative); Act II constitutes "rising action" (a series of complications which leads to the climax); Act III presents the "climax" itself (the moment where whatever forces have been released in the opening stages of the narrative have their definitive confrontation—i.e., the point where the conflicts of the story are ex- or imploded); Act IV contains "falling action" (a decline of the plot's movement down from the climax—thus *away* from the highest peak of interest and excitement); and Act V presents the "denouement" or "resolution," where the final outcomes are disposed of, weddings, funerals, and the like. The picture looks like this:

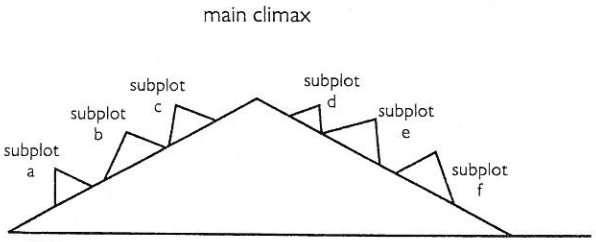


This pattern is so rigid and programmatic that even Elizabethan plays don't conform to it very strictly. Much less do fictional narratives of our own peri-

od. In modern narratives, the climax— the moment of the most important insight—will frequently be placed nearer to the end of the whole story, creating a different diagram, something like this:



As frequently, there may occur a pattern of many smaller peaks and valleys. These represent smaller subclimaxes and resolutions to subplots and subordinate conflicts over the course of a story and may be diagrammed in a figure that resembles a dragon's back:



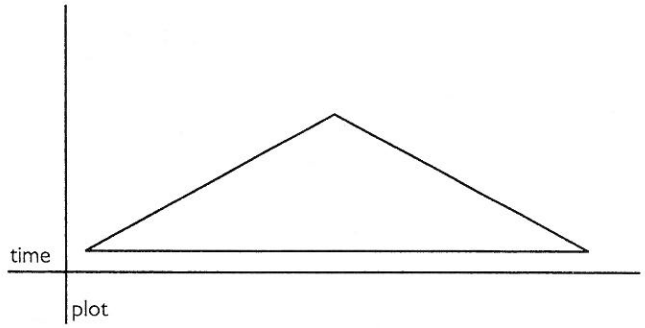
It's easy to get silly with these pictures. And indeed, most writers can get by very handily without them (or at any rate, without actually chalking them on the board) during the process of writing. The diagrams are no more than crude representations of the shape which the writer's intuition should be giving to the material as the process of composition goes forward. The Freitag triangle is a left-brain superimposition over what is for the most part a right-brain activity. But if intuition fails or goes astray, the triangle and its variants can be quite useful as diagnostic tools, perhaps even as problem-solving devices.

It is a familiar truism that all stories must always present some problem to be solved, some conflict in need of resolution— be it ever so humble, so apparently trivial. Conflict, the question which requires the story to answer it, is what generates the energy to ascend the rising slope of the triangle, toward the peak where the conflict will be, for better or worse, resolved; on the descending slope, the byproducts of the climactic fission or fusion settle back

toward a (temporarily) steady state. This, by some standards, is the very definition of what a story is; all narratives must share these qualities. There are probably just enough exceptions to this rule around to prove it.

Still, it would be difficult to think of (or write) a story in which no problem or conflict whatsoever arose at any point. At this moment, I can think of no example. The point is that all stories do bear some relationship to the structure of rising and falling action that the triangle is intended to graph. Suffice it to say that we do still expect some pattern of conflict and resolution from our narratives. To write a story with no vestige of these would be virtually impossible.

According to the Freitag triangle, the structure of a narrative is a function of plot and time.



The vertical axis represents plot; the horizontal axis represents time. The supposition here is that events will be told in chronological order, which does tend to be true in Elizabethan drama, but is not at all a fixed rule in more modern narratives. A second presumption is that plot is the primary structural element in the pattern. That presumption does not always hold for linear designs, but it is true often enough to be significant.

Because the figure represents the progress of events over time, you must consider that it is a process of motion as much as it is a fixed geometric form. That is to say, it is neither, or both. The difficulty of this distinction is reminiscent of the difficulty quantum physics encounters in distinguishing between a wave and a particle. But to the writer, this ambiguity is an advantage, because it offers two different ways of furnishing shapeliness to the work.

AS A PROCESS OF MOVEMENT, the linear narrative is timebound and sequential. The rule, proved by dazzling exceptions such as Martin Amis's *Time's Arrow* and Charles Baxter's *First Light*, is that movement in time will be for-

ward. The temporal vector runs out of the past toward the future, and the linear narrative follows it in a sequence of causes and effects, like a string of dominoes falling. Suspense, which controls the reader's desire to keep reading, is generated by the manipulation of the cause-and-effect cycle as it rolls toward the future, each effect becoming the cause of the next one. What effect will this cause produce? That's only a hifalutin version of the reader's fundamental question—*what's gonna happen?* Suspense (no narrative will hold a reader's interest without offering at least some *mild* form of suspense) comes out of a dextrous withholding of the answer to this question. With skill, you can string this withholding out almost forever—but as soon as you answer one question, you immediately have to think of another. Your task is to make the movement toward outcome, from first cause to final effect, seem inexorable. (In backward-running works such as *First Light* or *Time's Arrow,* the question is reversed: what first cause produced this final effect?)

The reader is interested in the outcome of a narrative as a scientist is interested in the outcome of an experiment. The writer of course may have other fish to fry—other aspects of the narrative may be infinitely more important to the writer. But plot and suspense are the instruments with which the reader is led (by the nose, if you like) across the passage of time that the narrative requires to take place.

Plot, suspense, causality, and time are inextricably intertwined, in linear narratives especially. But *real time,* the time in which our lives are lived second by second and hour by hour, remains a problem, sometimes an intractable problem, for the linear narrative. No story is long enough to actually express real time completely, absolutely, moment by moment. (Nicolson Baker's *Mezzanine,* a shortish novel which adheres with excruciating fidelity to the real time it takes the protagonist to traverse the mezzanine level of a department store, is the exception which proves *this* rule.) All narratives end up having to compress real time in some way or other—sometimes by summary and sometimes by skipping.

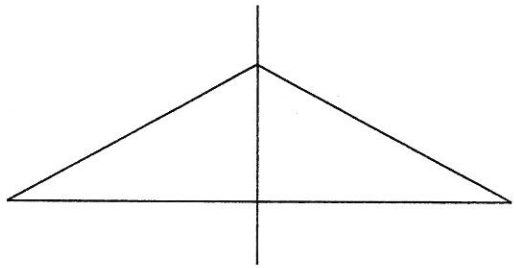
Before the twentieth century, the method for compressing real time in fiction was summary. Certain events and scenes of the plot would be given full dramatic rendering, as in theater, while the passages in between them would be summarized—most typically by the voice of an omniscient author. This voice would simply tell the reader—winningly, persuasively, and beautifully one might hope, but, failing that, at least economically—what happened between one fully dramatized scene and the next one. Exposition, the recounting of what led up to the scenes to be dramatically rendered, was the provenance of summary. In those days, successful and/or popular writers became very skilled in the writing of appealing and engaging summary, because they had to be.

All that was changed for our times by two new features of twentieth-century life: the movies and Ernest Hemingway. Perhaps Hemingway's most influential demonstration was that expository recitations of the past experiences of the characters could be eliminated entirely from stories, and that these leaner, meaner stories (his own if not those of his hosts of imitators) could be as interesting and even more energetic than their meatier, wordier predecessors. During approximately the same period, the movies were gradually teaching the audience to unconsciously accept transitions from one scene to another, across widely varying lengths of real time, with no explanation whatsoever. Look at an old silent movie, and you'll often see long title screens that summarize what happened between one scene and the next. Watch *Ordinary People* or *Jaws III* and you won't even get a voice-over to explain how one scene relates to another—and yet, somehow, you'll know.

For film editors, the term for these forward leaps across time is *jump cut.* This device—the unstated, unsummarized, unmentioned forward transition—has become increasingly popular for fiction writers also. In fiction, the jump cut is usually signaled, typographically, by the space break.

Time, however it is managed, determines the movement of the linear narrative. Plot, the all-important structural element according to the Freitag triangle, can only exist as a function of time. But other elements of fiction may at least aspire to a condition of timelessness, and these elements may also be used structurally, thus becoming design elements. Their structural use affects the triangle in its aspect not as a process of movement but as a stable geometric figure.

The classic Freitag triangle is an isosceles triangle, meaning that two of its three sides are of equal length. It is a bisymmetrical figure:



in which a line dropped to bisect the figure produces two equal halves which are mirror images of each other. The two halves are symmetrically in balance across that dividing line, as two weights might hang in balance across the vertex of a scale.

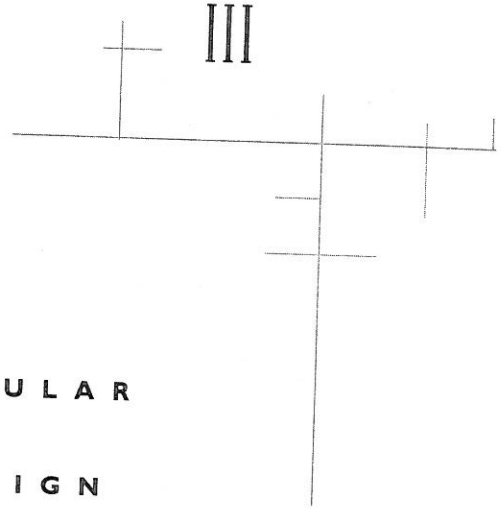
So far as the movement of a narrative goes, this symmetrical quality isn't so important. But if you think of a story more holistically, less as a temporal process and more as an integrated, unified artifact, then the issue of symmetry becomes much more significant. In linear design, this final symmetry need not be exact—no more than the two halves of your face are precisely symmet-

rical with each other, perhaps not even so much. But most narrative designs will bear some relationship to this principle of symmetry, in the same way that most plot structures will bear some relationship to the movement aspect of the Freitag triangle.

For this reason, the placement and timing of other elements of fiction—patterns of imagery, shifts of point of view or back and forth between first- and third-person narration, arrivals and departures of characters (the possibilities are very broad)— become, in fact, elements of design and are just as important to the overall design as is the plot, and sometimes more so. All these elements are to be arranged, to be used to create a sense of shapeliness, orderliness, balance, and integrity. Each must contribute to the reader's sense of the narrative as an integrated whole, for the moment when the narrative is apprehended as a whole is the moment when it is fully understood.

Narratives based on linear design do operate, on at least one level, like vectors; they are arrows fired into the future. In the final analysis, they should also be intelligible sculpturally, architecturally, as expressions of a static form. Abstracted, the triangular shape of a linear design is quite nondescript. But it is only an armature, a substructure around which you, the writer, deploy all your ingenuity, all your improvisational ability, to make the work your own. There are as many different *variations* on linear design as there are stories to be told.

PART III



MODULAR DESIGN

I SEE TWO DIFFERENT WAYS, quite distinct if not opposed, of thinking about the raw material of which narratives are made—two ways of contemplating the original experience, whether real, imagined, or some blend of the two, to which the work will give form. You may think of this primary experience as a single amorphous mass of information. Or you may conceive of it as a grab-bag of unassembled components—something like a jumble of unsnapped Legos, say.

In the first case, the task of the writer is to shape the material. The writer models a unitary mass of information in the same way that a sculptor models a blob of clay or carves a block of stone. Perhaps there is something in the mass to begin with that suggests the form the artist will give it. In that case, form is indwelling, inherent in experience, in the block. Michelangelo, at any rate, seemed to believe something of the sort when he spoke of his *discovery* of figures within masses of marble. For the writer, the discovery of innate forms underlying the superficially amorphous flux of experience is closely related to the discovery of meanings inherent in experience—if the writer is inclined to believe that the meaning of experience is intrinsic instead of being arbitrarily superimposed.

An artist with this attitude toward the raw material sees the artifact singly, as an integrated whole. Most often the work will be shaped by subtraction: only those elements which properly express the essential form are allowed to remain. That, quite literally, is the method of the woodcarver or stonecarver. Writers are less likely to think of their craft as subtractive, since in the beginning the words must always be piled up, not carved away. Still, the decisions about what to say and what not to say accomplish many exclusions, so that the mass of primary experiential material is in fact shaped by being reduced.

In narrative art, linear design provides the closest analogy to this kind of sculptural model (despite the fact that in plastic art, you really can't say that sculpture is linear). Both the carver and the writer of linear narratives are first concerned with the form of the work in its entirety. The carver thinks of the seamless smooth movement of the shape to be chiseled out. The writer thinks of the overall movement of the principal narrative vector from its start to its finish, all other issues being subordinated to this overarching concern.

In the second case, the task of the artist is not to discover the essential form of the work by whittling away the dross, but to assemble the work out of small component parts. This breed of artist is not so much a sculptor as mosaicist, assembling fragments of glass and tile to form what can be understood, at a greater distance, as a coherent, shapely image. In narrative art, this mosaic method is the basis for modular design.

If linear design can be understood as somehow subtractive, a process of

removing the less essential material so as to reveal the movement of narrative vectors more cleanly and clearly, then modular design is additive. The writer adds and arranges more and more modular units which may be attractive in themselves for all sorts of different reasons, but which also must serve the purpose of clarifying the overall design of the text as a whole. In linear design, the integrity of the finished work is obviously the first concern, since the writer is thinking of the work holistically to begin with. In the case of modular design, the writer will, at the outset, approach the raw material in a more fragmentary way. A sense of integrity in the work as a whole must be achieved by symmetrical arrangement of the modular parts. In a modular narrative design, narrative elements are balanced in symmetry as shapes are balanced in a symmetrical geometric figure, or as weights are balanced on a scale.

Of course, this symmetry need not be perfectly exact. Different kinds of symmetry are found in nature, and only the lowest organisms are radially symmetrical, so that any dividing line that passes through the center will produce two halves which reflect one another perfectly. Higher organisms, from flatworms to human beings, are *bilaterally* symmetrical, meaning that only one bisecting line will produce two precisely reflecting halves. In the case of an isosceles triangle, such as the Freitag triangle of a linear design, the symmetry will be exact. In living organisms, this bilateral symmetry will be rather more rough and approximate (as evidenced by the indubitable fact that one of my ears is even larger than the other one).

Manmade objects tend to follow similar rules. An architecture based on perfect circles or perfect squares may strangle on the constraints of its own symmetry. So the symmetry of most buildings is bilateral, not radial. This principle also extends itself to modular design. By following organic rules of order, modular design makes for a more lifelike narrative.

The units, modules, of modular design may be defined for narratives in all sorts of different ways, depending partially upon scale. At the page-by-page level, the modular unit will most probably look like a text block, separated from its fellows by space breaks. We have already seen how in linear narratives these insignificant-looking space breaks can be used to signal jump cuts from one point in a narrative forward to another, to indicate other shifts in chronology, to shift scenes or to accomplish other sorts of transitions, in a style which has been very much influenced by the movies. In modular design, this semicinematic mode of transition can be used to accomplish more complicated changes. From text block to text block, a modular design may change storyline, switch from character to character, switch between first-, second- or third-person narration in the treatment of the *same* character, make radical divergences in tone and voice, and in fact do almost anything you can imagine. All these shifts and rearrangements are held back from the brink of total anarchy by observing certain basic principles of order, symmetry, and balance.

Complex modular designs are more frequent in long fiction than short fiction. The use of modular design at the level of the short story is almost exclusively a late-twentieth-century phenomenon (with *Maldoror*, the work of the earlier French writer Lautréamont, providing at least one interesting exception to this rule). The influence of film narratives on prose narratives looks to be a likely explanation for that situation.

But at longer lengths, the modular concept may be as old as storytelling itself. Within cycles of mythology, whether Greek, Roman, or Stone Age primitive, individual tales can be and often are rearranged and reordered with respect to one another, in ways that may alter the total effect of the whole body of the narrative to which they belong. Single-author story cycles also have an ancient lineage, going back at least as far as Chaucer's *Canterbury Tales*, where the overarching conceit—twelve people telling tales to one another as they proceed on their pilgrimage—allows for considerable internal flexibility and the possibility of more than one arrangement. The case is much the same with *The Arabian Nights* and with another masterpiece of the Islamic tradition, Farid ud-din Attar's *The Conference of the Birds*.

In twentieth-century American literature, the first really powerful example of a story cycle is Sherwood Anderson's *Winesburg, Ohio*. Though Anderson wrote other novels, this modular work is his best known and best appreciated and the only book-length work of his still widely read outside of scholarly circles. The linear form of the novel was not well-suited to his gifts; modular design provided him with a better solution. Anderson's one-time apprentice Ernest Hemingway published a story cycle of sorts (*In Our Time*) as his first book; this group is not so coherently worked out as *Winesburg*, but the influence of the modular conception is quite clear. No one could claim that Faulkner was incapable of writing a linear novel, but three of his books—*Go Down, Moses; The Unvanquished*; and *As I Lay Dying*—are excellent (and quite different) examples of modular design.

More recently (from the 1970s on up) the story cycle, or episodic novel, has been a popular resort both for nonnovelists—story writers working at book length—and for novelists interested in trying a new form. Generally popular examples of this modular design–based genre include Louise Erdrich's first book, *Love Medicine*; Harriet Doerr's *Stones for Ibarra* (a winner of the National Book Award); and Carolyn Chute's *The Beans of Egypt, Maine*. Less well-known but equally admirable are Russell Banks's *Trailerpark*, Fred Chappell's *I Am One of You Forever*, and William Vollmann's *The Rainbow Stories*.

What modular design can do is liberate the writer from linear logic, those chains of cause and effect, strings of dominoes always falling forward. Modular design replaces the domino theory of narrative with other principles which have less to do with motion (the story as a process) and more to do with overall shapeliness (the story as a fixed geometric form). The geometry

of a modular design, especially one that has been well worked out in advance of composition, will be defining and confining to some degree. But the gain can be more than worth the sacrifice. The very fixity of the substructure can give the writer more latitude to improvise freely around the hidden armature with plot, character, and voice.

Of course a linear design may also show great internal diversity—in plot complication; character; shifts of point of view; shifts from one mode of narration to another; changes in voicing, tone and style. But all this diversity must be incorporated into the forward linear movement of the whole, as if the writer had taken a number of strands to make a braid. In modular design, all these ingredients can be treated less like strands and more like bricks. They can be managed as if they were discrete, particulate—capable of being assembled in more than a single way.

Time is a tyrant over all narratives: some events must always precede and others always follow. Modular design allows the writer to throw off the burden of chronology as much as is possible. It is always there, somewhere, but you may be able to proceed as if it didn't affect you. Modular design is an attractive way to show relationships between events or people or motifs or themes which are not generated by sequences of cause and effect and so are somehow atemporal, perhaps even timeless.

The sorts of problems which a modular design can solve are more likely to crop up at book length than at story length. That's partly because most writers can grasp the whole length of a story intuitively, without really thinking about it very explicitly, so that there is no real need to break it down into its components. But for most writers, a book-length work surpasses intuitive structural capacity. To try to intuit your way through a whole novel as if it were just a big story (which perhaps it is) will most likely be too overwhelming a prospect. There's a great incentive to organize it formally in advance to some degree, whether by preparing an abstract of its singular linear movement or by disassembling it into elementary component parts. And for many writers, often for the strongest and most intuitive story writers, the second, modular option will be the most appealing.

Because the composition of short stories usually doesn't require a great deal of formal advance planning, not too many writers resort to full-fledged modular designs at the length of the short story. But this rule is exceedingly well-distinguished with exceptions. The writers of the stories that follow have often adopted modular designs less for convenience than out of inevitable necessity. The modular form of such a story becomes quite inseparable from its meaning.

NARRATIVE DESIGN

A WRITER'S GUIDE TO STRUCTURE

MADISON SMARTT BELL

W.W. NORTON
& COMPANY
NEW YORK
LONDON

SUPPOSE THAT THE ELEMENTAL INGREDIENTS of fiction may be grouped in one or another of four major categories: plot, character, tone, and form. To define these terms quickly and simply: plot is what happens in a narrative; character is who it happens to (or who makes it happen); and tone is what it sounds like. Form is the pattern of its assembly, its arrangement, structure, and design.

Form is the aspect of a story that can be abstracted from everything else and expressed in some other medium, for instance, a graph, or some other geometrical figure. Not that a recreational reader would be likely to need or want to undertake this procedure—no more than you'd want to get to know your pet cat by dissecting it. But for writers it is sometimes (not always) necessary to perform such a maneuver of abstraction so that form can be rendered in ways that yield to analysis.

Form is of primary importance, always. Ingredients of fiction from the other three groupings (regardless of appearances, which may often be to the contrary) are always subordinate to form, to design. Indeed, any or all of these ingredients can and do function as elements of design. We are accustomed to thinking of plot as what defines structure in a story. But elements from other categories—point of view, imagery, shifts and alterations of tone—may also be used structurally and often are.* In reading and writing, you must consider (consciously or unconsciously) all of these aspects of fiction in terms of their relationship to the overall design. Even the overall meaning or theme of the narrative cannot be separated from this relationship. In a properly realized work, form and function are one and inseparable.

For the writer, some sense of the final formal design of the work really ought to precede the first stages of composition. The level of prior refinement of this sense of design will vary wildly from one writer to another. It may be quite specific and detailed (though it is risky for it to become *too* specific and detailed in the early stages—lest you create a paint-by-the-numbers design whose execution will suck the life from your conception). Or it may be no more than a vague and cloudy sense of where the story is headed—where *you* are headed, across the terrain of the story.

The length of the narrative being contemplated has a good deal to do with how evolved the writer's idea of its form needs to become before it is written. Most writers can navigate their way through a short story on sheer intuition at

*For the structural use of imagery, see the discussions of "Depth Charge," p. 49, "Daisy's Valentine," p. 120, and "O Man Alive," p. 157. For the structural use of characterization, see the discussion of "The Sky Is Gray," p. 189. For the structural use of tone, see the discussion of "Red Hands," p. 284.