Margaret Boden has argued that creative thought can be explained "in terms of the mapping, exploration, and transformation of structured conceptual spaces." Her definition of conceptual spaces is vague; she describes them as a "style of thinking"—in music, sculpture, choreography, chemistry, etc. In spite of this vagueness, the idea of transforming conceptual spaces is intuitively appealing. It has a close association with the earlier notions of Koestler, who has described creative thinking as arising from the sudden interlocking of two previously unrelated skills or matrices of thought, and the contemporary ideas of Perkoff, who uses the terminology of "klondike spaces" and argues that these are often systematically explored in the process of creative thinking. In this regard, while creative thinking is clearly part of ordinary thinking, and not something restricted to "geniuses," we can nevertheless see the potential for how particularly creative thoughts may arise from quite unusual transformations of conceptual spaces undertaken by particular individuals in particular circumstances.

What has occurred, then, is a convergence toward the essential idea that there is a single mental operation involved in creativity in a number of different domains. Whereas Aristotle and others picked out some interesting features of a few strikingly creative examples, and Koestler proposed that there is a special operation that underlies all these striking cases, contemporary creativity theorists have argued that this operation is not reserved for geniuses or for extraordinary acts of creation. We will show that this operation is indeed fundamental to all activities of the human mind, and we will try to lay out a precise and explicit theoretical framework in which to study its nature.

We begin where Koestler left off, with the case of "The Buddhist Monk."

---

**Three**

**THE ELEMENTS OF BLENDING**

A Buddhist Monk begins at dawn one day walking up a mountain, reaches the top at sunset, meditates at the top for several days until one dawn when he begins to walk back to the foot of the mountain, which he reaches at sunset. Make no assumptions about his starting or stopping or about his pace during the trips. Riddle: Is there a place on the path that the monk occupies at the same hour of the day on the two separate journeys?

This is the amazing riddle that Arthur Koestler presents in *The Act of Creation*. What we have to say about this Buddhist Monk will be more effective if you close the book for a moment and try to solve the riddle without any hints.

Now that you have found your place again, try this: Rather than envisioning the Buddhist Monk strolling up one day and strolling down several days later, imagine that he is taking both walks on the same day. There must be a place where he meets himself, and that place is the one we are looking for. Its existence solves the riddle. We don't know where this place is, but we do know that, whatever its location, the monk must be there at the same time of day on his two separate journeys. For many people, this is a compelling solution to the riddle.

But solving this little riddle only poses a much bigger scientific one: How are we able to arrive at the solution, and why should we be persuaded that it is correct? Is it impossible for the monk to travel both up and down? He cannot "meet himself." Yet this impossible imaginative creation gives us the truth we are looking for. We plainly don't care whether it is impossible or not—that's irrelevant to our reasoning. But the scenario of two people meeting each other is not only possible but also commonplace. Using that scenario is crucial to seeing the solution, even though it is nowhere in the original riddle, which describes just one person doing different things on different days.
The imaginative conception of the monk's meeting himself blends the journey to the summit and the journey back down, and it has the emergent structure of an "encounter," which is not an aspect of the separate journeys. This emergent structure makes the solution apparent.

The Network Model

The Buddhist Monk example reveals the central principles of the network model of conceptual integration. We will lay them out here.

Mental Spaces

Mental spaces are small conceptual packets constructed as we think and talk, for purposes of local understanding and action. In the Buddhist Monk network, we have a mental space for the ascent and another mental space for the descent. Mental spaces are connected to long-term schematic knowledge called "frames," such as the frame of walking along a path, and to long-term specific knowledge, such as a memory of the time you climbed Mount Rainier in 2001. The mental space that includes you, Mount Rainier, the year 2001, and your climbing the mountain can be activated in many different ways and for many different purposes. "You climbed Mount Rainier in 2001" sets up the mental space in order to report a past event. "If you had climbed Mount Rainier in 2001" sets up the same mental space in order to examine a counterfactual situation and its consequences. "Max believes that you climbed Mount Rainier in 2001" sets it up again, but now for the purpose of stating what Max believes. "Here is a picture of you climbing Mount Rainier in 2001" evokes the same mental space in order to talk about the content of the picture. "This novel has you climbing Mount Rainier in 2001" reports the author's inclusion of a possibly fictional scene in a novel. Mental spaces are very partial. They contain elements and are typically structured by frames. They are interconnected, and can be modified as thoughts and discourse unfold. Mental spaces can be used generally to model dynamic mappings in thought and language.

At various times along the way, we will use diagrams to talk about mental spaces and blends. In these diagrams, mental spaces are represented by circles, elements, by points (or icons) in the circles, and connections between elements in different spaces, by lines. In the neural interpretation of these cognitive processes, mental spaces are sets of activated neuronal assemblies, and the lines between elements correspond to coactivation bindings of a certain kind; in addition, the frame structure recruited to the mental space is represented as either outside in a rectangle or iconically inside the circle.

Input Spaces. In the Buddhist Monk network, there are two input mental spaces. As shown in Figure 3.1, each is a partial structure corresponding to one of the two journeys. The day of the upward journey is $d_a$, the day of the downward journey is $d_b$, the monk going up is $s_a$, and the monk going down is $s_b$.

Cross-Space Mapping. A partial cross-space mapping connects counterparts in the input mental spaces (see Figure 3.2). It connects mountain, moving individual, day of travel, and motion in one mental space to mountain, moving individual, day, and motion in the other mental space.

Generic Space. A generic mental space maps onto each of the inputs and contains what the inputs have in common: a moving individual and his position, a path linking foot and summit of the mountain, a day of travel, and motion in an unspecified direction (represented in Figure 3.3 by a double-headed arrow).

Blend. There is a fourth mental space, the blended space, that we will often call "the blend" (see Figure 3.4). Each of the mountain slopes in the two input mental spaces is projected to the same single mountain slope in the blended
space. The two days of travel, $d_1$ and $d_2$, are mapped onto a single day $d'$ and are thus fused. But the moving individuals and their positions are mapped according to the time of day, with direction of motion preserved, and therefore cannot be fused. Input Space 1 represents dynamically the entire upward journey, while Input Space 2 represents the entire downward journey. The projection into the blended space preserves times and positions. The blended space, which has time $t$ and day $d'$ contains a counterpart of $a_i$ at the position occupied by $a_i$ at time $t$ of day $d_i$ as well as a counterpart of $a_i$ at the position occupied by $a_i$ at time $t$ of day $d_i$.

**Emergent Structure**

The blend develops emergent structure that is not in the inputs. First, composition of elements from the inputs makes relations available in the blend that do not exist in the separate inputs. In the blend but in neither of the inputs, there are two moving individuals instead of one. They are moving in opposite directions, starting from opposite ends of the path, and their positions can be compared at any time of the trip, since they are traveling on the same day, $d'$. Second, composition brings additional structure to the blend. This structure of two people moving on the path can itself be viewed as a salient part of a familiar background frame: two people starting a journey at the same time from opposite ends of a path. Third, by means of composition, this familiar structure is recruited into the blended space. At this point, the blend is integrated: it is an instance of a particular familiar frame, the frame of two people walking on a path in opposite
directions. By virtue of that frame, we can now run the scenario dynamically: In the blend, the two people move along the path. This "running of the blend" is called elaboration. Running of the blend modifies it imaginatively, delivering the actual encounter of the two people. This is new structure: There is no encounter in either of the input mental spaces, even if we run them dynamically. But those two people in the blend are projected back to the "same" monk in the two input mental spaces. The meeting place projects back to the "same" location on the path in each of the inputs, and, of course, the time of day when they meet in the blend is the same as the time of day in the input spaces when the monk is at that location. The mapping back to the input spaces yields the configuration suggested by Figure 3.5.

As we run the blend, the links to the inputs are constantly maintained, so that all these "sameness" connections across spaces seem to pop out automatically, yielding a flash of comprehension, Koestler's magical "act of creation." But for this flash to occur, counterpart links must be unconsciously maintained even as they change dynamically across four mental spaces. In particular, there are geometrical regularities across these spaces. Given the way we have built the blend, we know that any point on the path in the blend projects to counterparts in the input spaces. More generally, anything fused in the blend projects back to counterparts in the input spaces. But this "geometric" knowledge of correlations among time, position of the monk, and location on the path in the different spaces is completely unconscious. What comes into consciousness is the flash of comprehension. And it seems magical precisely because the elaborate imaginative work is all unconscious.

**WHAT HAVE WE SEEN?**

Blending in the middle of the Buddhist Monk has features that turn out to be universal for conceptual integration.

Building an integration network involves setting up mental spaces, matching across spaces, projecting selectively to a blend, locating shared structures, projecting backward to inputs, recruiting new structure to the inputs or the blend, and running various operations in the blend itself. We will talk about these operations in sequence, but it is crucial to keep in mind that any of them can run at any time and that they can run simultaneously. The integration network is trying to achieve equilibrium. In a manner of speaking, there is a place where the network 'works' is happy. Context will typically specify conditions of the equilibrium, as when we are instructed to find a solution to the riddle of the Buddhist Monk. The network will achieve equilibrium if structure comes up in the blend that projects back automatically to the inputs to yield the existence of the special point on the path. More generally, what counts as an equilibrium for the network will depend on its purpose, but also on various internal constraints on constraint dynamics.

The Basic Diagram in Figure 3.6 illustrates the central features of conceptual integration: The circles represent mental spaces, the solid lines indicate the matching and cross-space mapping between the inputs, the dotted lines indicate connections between inputs and either generic or blended spaces, and the solid square in the blended space represents emergent structure.
Conceputal integration network. Blends arise in networks of mental spaces. In the network illustrated in the Basic Diagram, there are four mental spaces: the two inputs, the generic space, and the blend. This is a minimal network. Conceptual integration networks can have several input spaces and even multiple blended spaces.

Matching and counterpart connections. In conceptual integration, there is partial matching between input spaces. The solid lines in the Basic Diagram represent counterpart connections produced by matching. Such counterpart connections are of many kinds: connections between frames and roles in frames, connections of identity or transformation or representation, analogical connections, metaphoric connections, and, more generally, "vital relations" mappings (as explained in Chapter 6). In the Skiing Waiter case, for example, ski poles are counterparts of a tray. When matches are created between two spaces, we say that there is a cross-space mapping between them.

Generic space. At any moment in the construction of the network, the structure that inputs seem to share is captured in a generic space, which, in turn, maps onto each of the inputs. A given element in the generic space maps onto paired counterparts in the two input spaces. In the Iron Lady case, the generic space is something like "Western democracy with labor unions and voters." "Labor unions in the generic space maps onto American labor unions in one input and British labor unions in the other, which are accordingly counterparts. In the Skiing Waiter case, the generic has a moving individual carrying something in his hands. The carried object in the generic space maps onto the ski poles in one input and onto the tray in the other. They, too, are accordingly counterparts.

Blending. In blending, structure from two input mental spaces is projected to a new space, the blend. Generic spaces and blended spaces are related: Blends contain generic structure captured in the generic space but also contain more specific structure, and they can contain structure that is impossible for the inputs, such as two monks who are the same monk.

Selective projection. Not all elements and relations from the inputs are projected to the blend. The calendrical time of the journey in the Buddhist Monk case is not projected to the blend. In the Skiing Waiter case, neither walking nor the customer nor the price of champagne is projected from the waiter input. Sometimes two counterparts are both projected (both paths, both monks), sometimes only one (in the Iron Lady example, only American voters are projected, not British voters), sometimes none (in the Buddhist Monk example, calendrical dates). Sometimes counterparts in the input spaces are fused in the blend (the two paths), but often not (the two monks). And, finally, sometimes an element in one input without a
counterpart in the other gets projected to the blend (skin in the Skiing Waiter case).

- **Emergent structure.** Emergent structure arises in the blend that is not copied there directly from any input. It is generated in three ways: through composition of projections from the inputs, through completion based on independently recruited frames and scenarios, and through elaboration ("running the blend").

- **Composition.** Blending can compose elements from the input spaces to provide relations that do not exist in the separate inputs. In the Buddhist Monk example, composition yields two travelers making two journeys at the same time on the same path, even though each input has only one traveler making one journey. Counterpart elements can be composed by being included separately in the blend, as when the monks from the inputs are brought into the blend separately; yielding two monks; or by being projected onto the same element in the blend, as when the two days in the two inputs are projected onto the same day in the blend. We refer to this kind of projection as "fusion."

- **Completion.** We rarely realize the extent of background knowledge and structure that we bring into a blend unconsciously. Blends recruit great ranges of such background meaning. Pattern completion is the most basic kind of recruitment: We see some parts of a familiar frame of meaning, and much more of the frame is recruited silently but effectively to the blend. Figure 3.7 demonstrates this well-known psychological phenomenon, where we see two line segments and a rectangle and, through pattern completion, infer that there is a straight line running "behind" the rectangle. A minimal completion in the blend is often automatically interpreted as being a richer pattern. In the Buddhist Monk case, the composition of two monks on the path is completed so automatically by the scenario of two people journeying toward each other that it takes some thinking to see that the "journeying toward each other" scenario is much richer than the "two monks" composition.

- **Elaboration.** We elaborate blends by treating them as simulations and running them imaginatively according to the principles that have been established for the blend. Some of these principles for running the blend will have been brought to the blend by completion. We run the Buddhist Monk blend to get the "encounter" in the blend that provides the solution to the riddle. We are able to run the blend because we know the dynamics of the scenario of two people making opposite journeys along a path, which was brought in by pattern completion. That scenario gives us principles having to do with the passage of time, the possibilities of self-motion and so on. Part of the power of blending is that there are always many different possible lines of elaboration, and elaboration can go on indefinitely. We can run the blend as much and as long and in as many alternative directions as we choose. For example, the two monks might meet each other and have a philosophical discussion about the concept of identity. That particular elaboration would divert us from the purpose of solving the riddle, but it could also lead to something interesting and useful. The creative possibilities of blending stem from the open-ended nature of completion and elaboration. They recruit and develop new structures for the blend in ways that are principled but effectively unlimited. Blending operates over the entire richness of our physical and mental worlds.

Composition, completion, and elaboration lead to emergent structure in the blend; the blend contains structure that is not copied from the inputs. Note that in the Basic Diagram (Figure 3.6), the square inside the blend represents emergent structure.

- **Modification.** Any space can be modified at any moment in the construction of the integration network. For example, the inputs can be modified by reverse mapping from the blend, as in the Buddhist Monk case where we add to the inputs the existence of the location asked for in the riddle by backward projection of the spot of "encounter" from the blend.

- **Entrenchment.** Blends are often novel and generated on the fly, as in the Buddhist Monk case, but they remain entrenched mappings and frames. Blends themselves can also become entrenched, as in the Complex Numbers blend, giving rise to conceptual and formal structures shared throughout the community.

- **Event integration.** Blends are a basic instrument for achieving event integration. In the Skiing Waiter and Image Club examples, the event integration is the purpose of the imaginative construction. But in the Buddhist Monk case, it is only a means for solving the riddle about the existence of a location with certain properties.

- **Wide application.** Though uniform in their dynamics, integration networks can serve many different goals. To use the examples we have seen so far, these goals include transfer of emotions (Image Club) and inferences
pervasive and usually completely unnoticed aspect of the universe, schoolbooks introduce the subject with exotic and memorable illustrations, such as Benjamin Franklin's drawing down lightning by flying a kite in a storm. Indeed, it is quite reasonable that we do not become aware of a general, pervasive process until we see a case that looks exceptional.

The blend in the Iron Lady counterfactual is completely fantastic but, in context, is not noticed as unusual. We will see this repeated: Expressions that seem completely normal and literal to people (e.g., "This beach is safe") turn out to involve intricate blending, which stands out once we have analyzed it. Here are some other completely normal examples that have fantastic blends similar to the Buddhist Monk but go unnoticed in context: "This chapter is written itself," "You're getting ahead of yourself," "Normal Mailer loves to read himself," "I can't keep up with the schedule," and "My heart is disagreeing with my head."

The Buddhist Monk blend, rather amusingly, shows up in real life. Ed Hutchins studies the fascinating mental models set up by Micronesian navigators to sail across the Pacific. In such models, it is the islands that move, and virtual islanders serve as reference points. Hutchins reports a conversation between Micronesian and Western navigators who have trouble understanding each other's conceptualizations. As described by David Lewis, the Micronesian navigator Beirong comes to understand a Western diagram of intersecting bearings in the following way:

He eventually succeeded in achieving the mental tour de force of visualizing himself sailing simultaneously from Orolok to Ponoapo and from Ponoapo to Orolok and picturing the ETAK bearings to Ngrik as the start of both voyages. In this way he managed to comprehend the diagram and confirmed that it showed the island's position correctly. (The ETAK is the virtual island, and Ngrik is the island to be located.)

It is not just the blends that can be fantastic. Real life, from a certain point of view, can be fantastic, too. The Micronesian navigator's tour de force is especially fantastic for the anthropologist, Lewis, who knows the standard Western techniques of navigation, but what is fantastic for the Micronesian, and at first completely unintelligible to him, is the Western chart with the intersection of bearings. The spectacular blend invented by the Micronesian navigator Beirong exists only to make sense of the bizarre Western conceptualization. Ed Hutchins and Geoffrey Hinton, in their article "Why the Islands Move," explain how the seemingly mysterious Micronesian system actually works and reveal its great sophistication and power.

One might object that this case, even if real, is quite exceptional. But even so, it is no less representative of human cognition than more ordinary cases. Cognitive scientists take all such examples as revealing the hidden aspects of cognition: The
it is yesterday, and the automobile makes the journey. Then we are told that inevitability, yesterday, the person would have been passed by the automobile. This is a second blend, again nonfantastic.

By comparing this variant solution with that given by Koestler, we can see three general points.

First, impossibility need not be crucial to the blend: Koestler's blend is impossible, but the more complicated double blending in the Spanish story gives us two nonfantastic blends.

Second, the schemata used for pattern completion and emergent structure in the blend are experiential. In Koestler's solution, the experiential frame recruited is two people meeting each other on a path. The Spanish version recruits a richer frame: an oncoming car passing a pedestrian, where the car is moving much faster than the pedestrian. The encounter schema looked almost trivial in Koestler's solution, but finding the right encounter schema looks like a much more imaginative achievement in the Spanish version. In fact, the story frames the person who finds this solution as exceptionally intelligent for being able to deliver a blend that is familiar and immediately intelligible. Because the Spanish version has the descent taking much less time than the ascent, finding the automobile schema gives us a way to exploit our everyday knowledge about differential speeds. Even though the Spanish riddle is "mathematically" more complex, given the differences in the time it takes to ascend and descend and the corresponding differences of average speed, some people find the solution easier to accept because the final blend is a straightforward instance of a familiar situation: The oncoming car slows past the pedestrian. It is worth noting that in both the original and Spanish solutions, if "the same place" were interpreted literally, we would have to imagine two objects inhabiting the same space. In the Spanish solution, the car would run over the pedestrian. But everyone is willing in both cases to exploit, unconsciously, the convention that if two moving bodies pass each other closely, they are momentarily in "the same place."

Third, both solutions begin from the same inputs (ascend and descend of one person) and deliver the same topological solutions, but the imaginative networks they put together are different. So the inputs do not determine the integration network.

THE SCIENCE OF BLENDING

We have seen that blending is not deterministic.

Questions:

• Does the fact that we can't predict the integration network from the inputs mean that just any projections from the inputs are OK?
The Elements of Blending

More generally, explanation does not come from taking simplified accounts of simplified observations and extending them by adding some bells and whistles to give better explanations. Science moves in the other direction: A general scientific account explains the supposedly simple examples as special cases of the general account. For instance, it is an intuitive primitive for us that if we drop something, it falls to our feet. But that intuitive primitive is really only a very special case of the general theory of gravity, which also explains planetary motion and the parabolic trajectory of projectiles. We cannot explain scientifically how a dropped object falls to our feet without the general theory of gravity. There isn't a part of the theory of gravity that applies only to apples. Similarly, what we need for blending is a general theory that explains both the everyday and exotic examples.

Falsifiability

We have so far given analyses of blends, but we have not framed our analyses in terms of prediction and confirmation.

Questions:
- Does science involve making falsifiable predictions?
- What falsifiable predictions come from the theory of blending?

Our answer:
Actually, sciences like evolutionary biology are not about making falsifiable predictions regarding future events. Given the nature of the mental operation of blending, it would be nonsense to predict that from two inputs a certain blend must result or that a specific blend must arise at such-and-such a place and time. Human beings do not think that way. Nonetheless, in the strong sense, we hope to make many falsifiable predictions, including predictions about types of blending, what counts as a good or bad blend, how the formation of a blend depends on the local purpose, how forms prompt for blending, what possibilities there are for composing mappings, what possibilities there are for creating successive blends, how other cognitive operations (such as metonymy) are exploited during blending, and how categories are extended.

In fact, we take it that we have already falsified existing accounts of counterfactuals by showing the centrality of counterfactuals like the Iron Lady, which such theories are on principle unable to handle. We will do more of the same as we go along, by showing that, to the extent that they do not have a place for blending, existing accounts of metaphor, analogy, or grammar are in certain respects "false." Sociologically and psychologically, theory tends to define data, so it is hard to use data to falsify a theory while staying conceptually inside that.

First things first?
We have begun with striking examples so the blending can be easily seen.

Question:
- Wouldn't it be in better scientific taste to start by explaining the meaning of "The cat is on the mat" rather than trying to tackle the science fantasy of the Buddhist monk who can be in two places at the same time? First things first!

Our answer:
The scientific study of meaning would be easier if there were simple theories for simple meanings, which could be supplemented by more complicated theories for the more complicated meanings. But the simple meanings, it turns out, are also complicated, and as we will show repeatedly, you must have all of the operation of blending in order to put them together. Blending is not special in this way; it takes all of our cognitive powers to have common sense.
theory. A logician whose specialty is counterfactuals might honestly believe that an example like the Iron Lady falls outside the range of relevant data for logic and is only an anomaly of natural language, a "way of speaking," not a real example of counterfactual thinking. Philosophers like Donald Davidson who claim to study meaning have decreed in a similar vein that the study of metaphor is not part of the study of meaning.

**Visibility of the Blending**

We have tried to make consciously visible some of what happens during blending.

**Questions:**

- How can meaning be consciously apprehended at all if it is constructed unconsciously?
- What parts of blending are visible to consciousness, and when do invisible aspects rise to the surface?
- The Buddhist Monk example gives us a "Eureka" or "aha!" effect in which the whole solution pops out in a way that we seem to understand all the way through. How does that happen, and how is it different from moving step by step through a long argument to reach an earned conclusion? Is there a difference between meanings that we apprehend all at once and those we build up step by step?

**Our answers:**

Consciousness—like identity, sameness, and difference—appears to the conscious mind to be a primitive. For millennia, the face of consciousness and unconsciousness has been recognized and thought to be interesting, but theorists have only recently begun to be able to ask why people should have consciousness at all. Previously, this seemed a nonsensical question. Now it is recognized as a central and extremely difficult question for cognitive neuroscience, and several hypotheses have been proposed. We refer readers to the *Journal of Consciousness Studies* and *Brain and Behavioral Sciences* for the rich and short history of this controversy.

The question "How can meaning be consciously apprehended at all if it is constructed unconsciously?" makes the false assumption that what is apprehended consciously must be the output of a conscious process. But as we saw with the perception of a coffee cup, it is in the nature of consciousness that it gives us effects we can act on, and these effects are related to the unconscious processes. In the case of meaning, the apprehension of these effects typically induces us to refer meaning. Consciousness sees an effect and relies the effects to provide a cause: I see a cup, and in the folk theory, the reason I do that there is a cup that causes me to see a cup. In the same way, I hear a sentence and