CHAPTER 7

Naïve Analogies

Three Anecdotes

Timothy is watching his father shave one morning. He observes his father moistening his face, spreading some shaving cream across it, using the blade, and rinsing. From his four-year-old vantage point, Timothy categorizes the scene as best he can. The idea that a razor, so different in appearance from a pair of scissors or a knife, might be able to cut something does not cross his mind. On the other hand, Timothy knows very well that certain substances dissolve in other substances, such as sugar in hot water. He is therefore absolutely convinced that the shaving cream dissolves his father’s stubble, and that the razor’s sole purpose is to wipe away the shaving cream once it has done its job.

Janet is on a local mailing list, and one day she received a message from a chatty neighbor sharing this news: “This morning I enjoyed watching a bunch of titmice feasting on the insects on the branchlets of our tulip tree. There’s such an abundant crop of insects this year that I think it will attract a large population of insect-eaters like titmice and chickadees.” Janet was puzzled by the image of teeny mice scurrying about on the branches of a tree, since she had never witnessed any such thing. When she came to the phrase “titmice and chickadees” she was puzzled yet more, since it seemed odd that mice and birds would happily coexist on the branches of a tree. All at once it hit her that titmice are not in fact teeny mice, but are birds, just like chickadees.

Professor Alexander is bidding good-bye to a younger colleague who is leaving for Germany for a month. He says, “When you arrive, please send me your email address, won’t you?” Seeing the look of perplexity on his colleague’s face, Professor Alexander bangs his hand against his forehead. “What am I saying? Obviously, your email address isn’t going to change at all!”
The age difference between Timothy and Janet is about the same as between Janet and Professor Alexander, but despite these large gaps, the same cognitive phenomenon is at work in the young child, the young adult, and the older adult. All three were taken in by tempting analogies, which, just like the categorical blinders discussed in Chapter 5, led them into error. In the cognitive-psychology literature, one finds all sorts of terms for this phenomenon, including “preconceived notion”, “spontaneous reasoning”, “naïve reasoning”, “naïve theory”, “naïve conception”, “tacit model”, “conceptual metaphor”, “misconception”, and “alternative conception”. Although these terms are not all interchangeable, they do have in common one key thing, and we will call that core notion “naïve analogy”.

The idea is that an unfamiliar concept (such as shaving cream, titmouse, or email address, in the three anecdotes) is apprehended plausibly, although inaccurately, through a natural-seeming analogy with a prior piece of knowledge (here, knowledge about hot liquids, mice, and postal addresses). Such analogies allow a person to make at least some sense of the new situation by likening it to something familiar, and yet it is all done in a spontaneous, unconscious, automatic way, without the person’s least awareness of making an analogy.

This stands in stark contrast to the standard image of analogy-making as a process of deliberate construction of mappings between situations. Naïve analogies lead directly to conclusions without there being any consideration of other options, and without any uncertainties or doubts arising. Thus the shaving cream is taken for granted as a dissolving substance by Timothy, the hungry titmice as a type of tree-born rodent by Janet, and the email address as a place-specific address by Professor Alexander. The presence of these analogies is never felt explicitly, however.

Just like other acts of categorization, naïve analogies lead one to a perfectly reasonable (and thus self-consistent) interpretation of a situation, but they unconsciously assume that one is dealing with a typical member of the selected category. However, the situation may well involve an atypical member or even a non-member of the chosen category, in which case the conclusions reached will be irrelevant and useless. Thus if an email address were a postal address (the most familiar type of address to Professor Alexander), then the question he asked would have been totally reasonable, because the colleague was indeed going far away. Similarly, if a titmouse were indeed a very small rodent, then it would have been reasonable to be surprised by an image of such animals scampering about on tree branches, and Janet’s confusion at this image would have been perfectly comprehensible.

As for young Timothy’s categorization of his father’s shaving scenario, that too, was very reasonable, given his prior knowledge. For an adult, the category shaving presumes that there is a blade that will cut some hair and that there is a lotion whose purpose is to make the cutting easier and to reduce the blade’s chances of nicking the skin. For an adult, it’s obvious and unquestioned that what’s doing the cutting is the razor blade. But Timothy saw it quite differently. He was quite right in thinking that if the shaving cream dissolved the small hairs, then some sort of spatula might be useful in getting the shaving cream off his father’s face. Even if this interpretation might make adults smile,
there was nothing particularly childish about the thought process. No matter how wrong it might have been, it is not an iota less self-consistent than the adult’s vision of the process. One can even think of it as a rather ingenious invention, for after all, if such a marvelous hair-dissolving cream did exist, then all our razors would soon be museum pieces.

Generally speaking, naïve analogies have a certain limited domain in which they are correct, and which justifies their existence and their likelihood of survival over years or possibly even decades. This domain of validity can be narrow or broad. This is the case, for example, for children who personify animals. A grasshopper has much in common with a person: it is alive, breathes, moves about, reproduces itself, is mortal, and can be wounded; up to that point, the naïve analogy is perfectly useful. However, unlike what six-year-old children generally think, a grasshopper will not be sad if the person who is taking care of it disappears, and this illustrates one limitation (among many) of the analogy.

As for the naïve analogies at the chapter’s start, their strengths and limitations are easy to see. For example, the analogy between a postal address and an email address is valid in a number of ways: both are pieces of data that are structured hierarchically, moving from local to global information (an email address starts with something like a personal name; then comes an at-sign; then something that corresponds roughly to a street address; then a dot and finally something vaguely akin to a state name or country name), and which can be given to certain people and kept secret from others; both are associated with “boxes” where mail accumulates and can be accessed; both are subject to occasional changes; and so forth. For this reason, the analogy is shared by nearly everyone, and of course it is implicit in the shared term “address”.

But the analogy has its limitations, too. To send something electronically, the sender needs to have an email address, whereas no such thing is needed in order to send something by post. When one sends an email, usually a copy is automatically kept by the sender, in contrast to postal mailings. An electronic message arrives almost instantly, while a postal shipment may take a few days. If one moves, one gets a new postal address but one can keep one’s email address. And so forth. It is therefore understandable that it might fleetingly occur to a person to ask about the new email address of a friend who is moving to a new place (though this is far more likely to happen to a novice email user than to a seasoned one). This is where the limit of validity of the analogy becomes clear.

The naïve analogies made by Timothy and Janet are far more idiosyncratic than the one made by Professor Alexander. To figure out just how common Timothy’s naïve analogy is, one would have to make a careful study of how children understand the process of shaving. As for Janet’s confusion, it probably strikes you as rather quaint that an adult might envision a titmouse as being a very small mouse that scampers about on tree limbs, but if such naïveté makes you smile, keep in mind that we all live in glass houses, for we have all fallen into traps of the same sort from time to time, by making overly rapid and inappropriate categorizations. Let’s take a look at an example from classical popular literature.
Many readers will be familiar with Æsop’s fable “The Ant and the Grasshopper”, and some will know the seventeenth-century French poet Jean de La Fontaine’s rhyming version thereof, called “La cigale et la fourmi” (literally, “The Cicada and the Ant”), of which the opening lines run as follows (in our own translation):

\[
\begin{align*}
All \ summer \ long, \ without \ a \ care, \\
Cicada \ sang \ a \ merry \ air, \\
But \ when \ harsh \ winter \ winds \ arrived, \\
Of \ food \ it \ found \ itself \ deprived: \\
It \ had \ no \ wherewithal \ for \ stew: \\
No \ worm \ or \ fly \ on \ which \ to \ chew. \\
\end{align*}
\]

The last two lines are unlikely to give most readers pause, but in them, in fact, there lurks a mistaken assumption. To bring this out into the open, let’s explore a small variant of them. Suppose La Fontaine had instead written, “It had no wherewithal for stew: / No horse or cow on which to chew”. In that case, readers would almost certainly be thrown by the incongruous image of a mere insect having failed to build up a stock of barnyard animals on which to feed. And readers would have been even more disoriented had La Fontaine written, “It had no wherewithal for stew: / No shark or whale on which to chew” or else “It had no wherewithal for stew: / No stick or stone on which to chew.” Such lines would have instantly aroused suspicion and bafflement.

Although the closing lines of La Fontaine’s actual poem appear to lack any such incongruity, that impression is wrong, since it turns out that cicadas are not carnivorous, and so they have no use for flies or worms. The innocuous-seeming assumption that cicadas feed on small creatures of about their own size is simply erroneous. The famous French biologist and science writer Jean-Henri Fabre, in his autobiographical memoirs entitled “Entomological Souvenirs”, observed that cicadas have only a sucking tube with which to nourish themselves. In their larval stage, they get their nourishment from the sap of roots, and when they reach maturity, they suck sap from the branches of various trees and bushes. The plausible-seeming image of what cicadas eat is simply based on a naïve analogy with people or farm animals or other types of insects. La Fontaine’s poem would have been far more faithful to the true nature of cicadas if it had run this way:

\[
\begin{align*}
All \ summer \ long, \ without \ a \ care, \\
Cicada \ sang \ a \ merry \ air, \\
But \ when \ harsh \ winter \ winds \ arrived, \\
Of \ food \ it \ found \ itself \ deprived: \\
It \ hadn’t \ stocked \ one \ sip \ to \ lap \\
Of \ what \ cicadas \ crave: \ thick \ sap. \\
\end{align*}
\]

If the great La Fontaine was so naive, perhaps we should not judge Timothy, Janet, or Professor Alexander too harshly.
Naïve Analogies, Formal Structures, and Education

As people move into higher realms of abstraction, naïve analogies, with all their strengths and weaknesses, inevitably become trusted guides. The strengths of such analogies derive from their easy availability in long-term memory, in the form of efficient and ready-to-use mental structures. And their weaknesses stem from the fact that in certain contexts they are misleading. Naïve analogies are like skiers who sail with grace down well-groomed slopes but who are utterly lost in powder. In sum, naïve analogies work well in many situations, but in other situations they can lead to absurd conclusions or complete dead ends.

What the study of naïve analogies tells us about the human mind is of paramount importance for education, and this chapter is therefore oriented to some extent towards the educational payoffs connected with our ideas. A certain number of entrenched ideas about what it means to learn and to know will be called into question, and some new directions for education will be suggested. Below we list three key ideas that we will discuss in this chapter and the following one.

First of all, ideas that are presented in school classrooms are understood via naïve analogies; that is, children unconsciously make analogies to simple and familiar events and ideas, and these unconscious analogies will control how they will incorporate new concepts.

Secondly, naïve analogies are in general not eliminated by schooling. When teaching has an effect on a student, it usually just fine-tunes the set of contexts in which the student is inclined to apply a naïve analogy. The naïve notion does not displace the new concept being taught, but coexists with it. Both types of knowledge can then be exploited by a learner, but they will be useful in different contexts. And this is fortunate, since banishing naïve analogies from people’s minds would be extremely harmful. For example, looking at the world from the point of view of a professional physicist in everyday situations would often be hopelessly shackling. A physicist who sees a glass start to fall floorwards doesn’t need to wheel out Newton’s laws of motion and his universal law of gravitation in order to figure out what’s about to happen. Reaching out to catch the glass that’s about to be shattered is a straightforward consequence of pre-Newtonian, non-technical world knowledge. And if two astronomy students are walking hand in hand on the beach admiring the pink-and-orange sunset, they most likely are not in the least thinking about the fact that it’s the earth that’s turning rather than the sun that is descending, and it’s most probable that they are enjoying the beautiful colors and the romantic feelings in much the same way as any other couple would.

Finally, a formal description of a given subject matter does not reflect the type of knowledge that allows one to feel comfortable in thinking about the domain. Humans do not generally feel comfortable manipulating formal structures; when faced with a new situation, they favor non-formal approaches. Learning is thus the building-up not of logical structures but of well-organized repertoires of categories that themselves are under continual refinement.
Chapter 7

Familiarity and Entrenchment

Familiarity is crucial in analogy-making for the simple reason that, in order to deal with an unknown situation, one intuitively feels more secure about extrapolating what one knows well than what one barely knows. We don’t mean that such choices are made consciously; they take place below one’s level of awareness. Thus unconscious analogical processes dominate the way we interact with our environment, forming the very basis of our understanding of the world and the situations we find ourselves in.

Quite obviously, we are not equally familiar with all the things that surround us. Certain notions seem totally natural to us, and others very little so. The notions of addition, equality, adjective, verb, continent, and planet strike us as familiar, while the notions of partial differential equation, Fourier series, topological space, spinor, lepton, electrophoresis, and nucleotide synthesis exude considerable strangeness for most of us. This is the case not only for abstract notions of that sort, but also for concrete objects. For most of us, rockets are less familiar than cars, and household robots are less familiar than computers. A category’s familiarity has to do with how much one has been exposed to it, with the amount of knowledge one has of it, and with the degree of confidence one has in one’s knowledge about it. One feels more comfortable with cars than rockets because one has seen many more of them, because one knows much more about them, and because one has greater clarity about how to get into and out of them, about what their control devices (steering wheel, brake, etc.) will do, and so forth. One is also, for similar reasons, far more familiar with gravity than with electromagnetism.

Familiarity has been studied in certain psychological experiments that explore the effect of the degree of entrenchment of a category. For instance, cognitive psychologist Lance Rips showed that a new piece of knowledge is more readily transferred from a typical member of a category to an atypical member of the same category than the reverse. For example, if participants in an experiment are told that robins are susceptible to a certain disease, they are likely to conclude that hawks, too, might catch this disease, whereas in the reverse situation, in which they are told that hawks are often afflicted with a certain disease, the chance is much smaller that they will infer that robins will suffer from it as well. The greater familiarity of the category robin than that of hawk is the source of this asymmetry. This means that the more entrenched a concept is (here, robin being more entrenched than hawk), the more likely it is to act as the source of an analogy.

A set of experiments conducted by Susan Carey, a developmental psychologist, on children of various ages and on adults as well, led to similar conclusions. Participants were told that all the members of a certain category — dogs, for instance — have a certain property, such as possessing an internal organ called an “omentum”. They were then asked if members of other categories, such as humans or bees, were also likely to have this internal organ. For very young children, but not for adults, the existence of such an organ was more easily transferred from person to dog than from dog to person. The explanation for this is that in the minds of young children, the concept of person is more entrenched than the concept of dog. Adults were perfectly happy extending
Naïve Analogies

possession of such an organ to any species of animal whatsoever, as long as they had been told that two very different kinds of animals (such as dogs and bees) both possessed it. Young children’s thought processes were different; they would decide whether or not to extend a property to a new kind of animal depending on that animal’s perceived proximity to either of the two species (i.e., depending on the strength of the analogy in question), and they did not use the superordinate category animal to make the analogy, because that abstract category was not sufficiently familiar to them.

Everyday Concepts Versus Scientific Concepts

A narrow vision of learning that we would be happy to see fully eliminated from educational programs presumes that knowledge acquired in school is independent of everyday knowledge. This philosophy would encourage the teaching of new ideas without any reference to everyday concepts, except for the most rudimentary notions, which are impossible to separate from everyday experience. Such a limited view of learning is based on a conception of the mind according to which we would keep facts and ideas that we pick up in school in a separate mental compartment from facts and ideas that we pick up in daily life, and year after year we would increase our school-based knowledge much as we would build a brick house: each brick added would be supported by previously installed bricks, and would in turn support new bricks.

Such a vision is appealing in some ways, since it would simplify the design of school curricula. It would make teaching easier, treating each discipline as an island disconnected from other disciplines, and trying, within any given discipline, to decompose its notions in a logical fashion and to arrange them in a strict, natural order. Teaching any complex idea would thus involve teaching a set of simpler ideas in the same domain, each of which would in turn be taught through yet simpler ideas in the same domain, and so forth.

It is a long-lived myth in the world of education that there is a watertight boundary between two types of knowledge: everyday knowledge, presumed to grow on its own with no need for formal teaching, and formal knowledge, conveyed in schools and presumed to be communicable independently of everyday knowledge. In this naïve view, school is seen as a magical shortcut that allows ideas arduously developed by humanity over thousands of years to be transmitted in just a few years to a random human being.

Another belief that exerts considerable influence on school curricula, particularly in science, is that scientific knowledge is best conveyed in precise, formal terms — especially through mathematical formulas — and that precise formalisms correspond to the way in which knowledge ought to be absorbed by beginners and in which it is manipulated by experts. It is presumed that the initially large gap between a printed formalism and one’s internal mental representation gradually approaches zero as one comes closer and closer to expertise.

While we’re on the subject of distances approaching zero, the following rather prickly formula expresses the idea that the function $f(x)$ is continuous at the point $x_0$:
∀ε ∃δ ∀x \mid x - x_0 \mid < \delta \implies \mid f(x) - f(x_0) \mid < \varepsilon

(The upside-down “A” and the backwards “E” are shorthand notations for the words “for all” and “there exists”, and the arrow “⇒” can be read as “if… then…”.) Should one presume that all people who are comfortable with mathematics think of continuity in exactly these terms? Do people fluent in calculus really always imagine and mentally manipulate Greek letters, and is this what continuity means for them? Are all their prior intuitive notions of continuity totally dispensed with after they have absorbed this formula? If so, then in order to educate experts, or even to convey technical notions of this sort to ordinary students, the right way would be to transmit just such formulas, since they presumably embody the most distilled and precise essence of the notions. In particular, one would want the notion of continuity to be seen by students as synonymous with the above-displayed formula.

However, this view of education unfortunately conflates the actual way that experts think with the use of dense formalisms designed to capture subtle notions as rigorously and unambiguously as possible. It is the result of a long-standing philosophical assumption, reinforced by the broader culture, which is that logical thinking is superior to analogical thinking. More specifically, this view comes from misguided stereotypes of analogical thinking, which maintain that dependence on analogy, although possibly useful when one is just starting out in a field, is basically childish, and that analogies should rapidly be shed, like crutches or training wheels, when one gets down to brass tacks and starts seriously thinking in the domain. A related half-baked stereotype of analogical thinking is that it is like an untamable wild horse, so unpredictable and unreliable that it must be shunned, even if it might once in a while provide a spark of true insight; thus analogies belong not to the realm of reason but to that of “intuitions”, which, being irrational, cannot and should not be taught.

The fluid way that a scientific notion is realized and “lives” in the minds of people who deeply understand it is a very different thing from a formal and rigid symbolic notation, which has been carefully devised to be as concise and rigorous as possible, and the two should not be conflated. As we have just seen, the notion of continuous function has a precise formal definition, and yet continuous functions form a mental category having blurry boundaries and members with different degrees of typicality, and such judgments will vary from one mathematician to another. A famous historical case illustrates this fact. Toward the end of the nineteenth century, functions continuous at every irrational point and discontinuous at every rational point were discovered, and this behavior was so unexpected that some renowned mathematicians labeled them “pathological” and strove to banish this “scourge” from mathematics, whereas other mathematicians were extremely excited by the revelation of such a rich new area for research. This episode hardly jibes with the image of mathematics as a frozen body of precise and absolutely objective knowledge. In the following chapter we’ll see that even a category as familiar as number has blurry boundaries and members with varying degrees of typicality, and seasoned mathematicians can hold different opinions and can argue vehemently about what is and what is not a member of the category.
The unfortunate but widespread conflation of austere formal definitions with the psychological reality of concepts in human minds has had worrisome repercussions in education. One consequence is that it tends to make the educational establishment lose sight of what ought to be its primary goal — the construction of useful, reliable categories. Another consequence is that it tends to favor teaching methods based on prickly formalisms and rigorous deductions rather than methods that use analogies to build up suitable families of categories in intuitive ways.

Our view is very different from one in which logic is seen as central. Indeed, as we stressed in Chapter 4, expertise builds up as categories are acquired and organized. Rather than depending on formal perceptions of situations, people have the ability to treat novel situations as if they were familiar, thanks to categorization. To acquire knowledge in a domain is to build up relevant categories. Analogical thinking is the key to understanding new situations and to building up new concepts, and this holds at all levels, ranging from the shakiest beginner to the most fluent expert. The difference between those two is not their style of thinking — logical for the expert and analogical for the beginner — but the repertoire of categories that they have at their disposition, and the way those categories are organized.

For example, in the case of continuity of a function at a given point $x_0$, do experts really think in terms of Greek letters and flipped and rotated roman letters? Hardly. Instead, they have vivid visual imagery about the concepts concerned. The epsilons and deltas in the formula are merely helpers in translating that imagery. Thus if one wants the value of $f(x)$ to be very close to that of $f(x_0)$ — in particular, if one wants their discrepancy not to exceed the tiny number $\varepsilon$ (this desideratum is expressed by the notation $\mid f(x) - f(x_0) \mid < \varepsilon$) — then continuity requires that there should exist some little zone centered on the point $x_0$ and whose two edges are at a distance of $\delta$ from the center (this zone is expressed by the notation $\mid x - x_0 \mid < \delta$), and throughout this zone the desideratum should be the case.

Thus if one has a mathematically trained mind and is thinking about what continuity means, one envisions a small rectangular box centered on the point $(x_0, f(x_0))$ in the plane. Continuity then will mean that no matter how close one wishes $f(x)$ to be to $f(x_0)$ (this desire translates into the image of a box whose height is very tiny, so that all $y$-values in it are vertically close), one can always make the box so narrow that the desideratum will hold everywhere inside it. The phrase “no matter how close one wishes” is the visual translation of “$\forall \varepsilon$” (pronounced “for all epsilon”), while the phrase “one can always make the box so narrow” is the visual translation of “$\exists \delta$” (pronounced “there exists some delta”). In the end, it all comes down to the idea of zooming in on the graph of the function at arbitrarily small scales, and thus it has to do with such familiar things as magnifying glasses and microscopes and walking towards objects so that what is blurry at a distance comes into focus. These are the kinds of everyday experiences in which a mathematical sophisticate’s understanding of continuity is rooted. Thus we see that to understand the rich concept of continuity is to have built up this kind of imagery, and the epsilons and deltas are merely tools, used briefly and discarded quickly, which help one to realize that goal.
In sum, notions taught in schools and colleges are internalized by students not formally, but by means of naïve analogies, which is to say, by means of analogies with familiar categories. These categories can come from any domain, and they tend to come from domains that are not covered in school. Thus the presumption that a given area of knowledge is self-contained is erroneous, because students will inevitably connect every new notion in that area with experiences they’ve had in other areas of life. Moreover, a naïve analogy, once it has taken root in a beginner’s mind, will not go away as time passes; it will stick in the mind as school goes on, year after year, because everyone has a deep need for simple, basic intuitions. Indeed, some naïve analogies are so persistent that one might begin to doubt whether schooling can have any effect at all in certain areas. The goal of transmitting to average pupils subtle ideas that humanity as a whole took thousands of years to discover and absorb is admirable, but it is not something that just happens by itself.

We will see that certain notions that are usually thought of as extremely simple and that are taught in elementary school are, unfortunately, understood through naïve analogies having quite limited realms of validity, and these often-misleading analogies remain entrenched through middle school, high school, and college. Their subliminal effect lives on, long after education should theoretically have thoroughly drummed them out. Thus naïve analogies, thanks to their great robustness, lie behind the thought processes of well-educated adults, even experts; this casts doubt on the idea that education gets rid of them. The stereotypical vision, by contrast, is that people who have deeply absorbed scientific ideas swim in a soup of formally defined notions that obey logical laws; analogies play no part in this vision. That, however, is just wishful thinking. To be sure, scientific education does convey ideas that go beyond everyday experience, and in particular they tend to be more abstract, but scientific ideas are not learned any differently from other ideas — they, too, are rooted in naïve analogies, and this is a universal fact, whether one is talking of beginners, intermediates, or experts.

Where Novelty and Familiarity Walk Hand in Hand

There is no question that the development of computers and related technologies has given rise to a major revolution on our planet. Those who remember the day when there were no computers in homes or businesses, and when there was no World-Wide Web to which to connect such devices, have children who are filled with wonderment at the idea that people could get along at all in such a primitive world. It thus makes sense to suggest that computers were the great innovation of the twentieth century. It also might seem that when something so new and so completely unprecedented takes over so widely, it would involve such a radical break with prior concepts that completely new categories would have to be introduced, categories not grounded in any kind of analogy at all.

If computer technologies are so hugely innovative, don’t they have to do much more than merely use various analogies to clothe familiar old notions in fancy new wrappings? Actually, the key question is the exact opposite — namely: how could
anyone imagine a great burst of creativity taking place without its pioneers building left and right on familiar notions? And how could the lay public possibly absorb the tsunami of radically new ideas without basing their new skills on analogous skills involving notions that were already deeply rooted? The simple truth is that only through homey analogies based on familiar, everyday categories can an average person relate to all the revolutionary new technologies; such analogies are anchors keeping people from getting totally lost in a huge sea of technical terms. The following memo is a good case in point.

Please find my “Recent stuff” folder and make a copy of the document in it, then put it in my “Paid bills” folder, and also send it to my personal address. Then could you clean up my messy desktop? Would you also look in my “Miscellaneous” folder and toss out anything in it that’s outdated or irrelevant? When you’re done, please empty the trash, close all the windows, and turn everything off. Thanks a lot!

This note, had it been written in the 1970s, would have referred to a wooden desk, some cardboard folders, documents and sheets of paper, a metal wastebasket, glass windows, and a postal address, and there couldn’t possibly have been any confusion as to what was being talked about. Today, however, the paragraph is highly ambiguous, because the supplies and furniture to which it refers have both real and immaterial embodiments, which are so closely analogous to each other that it’s impossible to tell which type of stuff is being referred to.

We thus see that computer technology — the domain that represents the greatest break between this century and the previous one — relies on scads of analogies with old-fashioned things, and these analogies are so tight that a perfectly realistic little scenario can be concocted in which there is no way to tell which century’s technology is being referred to. Actions that can be performed on physical objects have their counterparts in the virtual world. The desk, the files, the folders, the trash all exist in both worlds and just about anything that can be done electronically can also be done in the old-fashioned way. Even if people today read documents that cannot be crumpled, placing them in folders that cannot be torn, stacking those on desktops that are not made of wood, sending them to addresses that have no particular geographical identity, opening and closing windows that have no handles, and filling and emptying the trash with just a few flicks of their fingers, they are always doing so by analogy with the most familiar situations they know.

The reason that computers have revolutionized our society but not our vocabulary is that these very powerful devices have all been grafted onto familiar categories and have borrowed their verbal labels wholesale. These days, everyone speaks about virtual technologies using terms that wouldn’t in the least have disoriented our grandparents. Just as in the good old days, we open our mailbox to get our mail, we send messages to addresses, we send and receive documents, we visit sites, we chat with friends, we do searches for things we want to find, we consult pages, we make new links, we browse, we toss useless files into the trash, we open and close folders and windows, and so forth.
To be sure, there are certain new expressions that would have completely thrown anyone who heard them thirty years ago, such as “hosting a site on a server”, “surfing the Web with one’s browser”, and “installing a firewall against hackers”, but even so, these droll expressions, too, bear witness, in their own way, to the fact that all of our electronic activities are still cast in terms of the physical world around us. The humorous juxtapositions audible in these phrases reveal that as a species, we humans are totally dependent on familiar categories in order to adjust to brand-new realities. If hackers are using the Web to transmit viruses hidden inside spam, then who’s to say we shouldn’t use firewalls to protect ourselves from such shenanigans?

A more systematic exploration of the lexicon that has grown up around the Web and electronic technologies only serves to confirm our thesis that extremely familiar, everyday physical categories are overwhelmingly the most standard and relied-upon source of analogies for brand-new phenomena. Here, for instance, are roughly one hundred terms all of which could be found in dictionaries published long before computers played any role in society, but which, if read today, give the sense of being technological:

account, address, address book, animation, application, archive, attach, bit, bookmark, bootstrap, browser, bug, burn, bus, button, capture, card, chat, chip, clean up, click, close, compress, connection, copy, crash, cut, delete, desktop, disk, dock, document, drag, dump, entry, erase, figure, file, find, firewall, folder, font, forum, gateway, hacker, highlight, history, home, host, icon, image, input, install, junk, key, keyboard, library, like, link, mail, mailbox, map, match, memory, menu, mouse, move, navigate, network, open, output, pad, page, palette, paste, peripheral, point, pop up, port, preference, preview, print, process, program, quit, read, reader, record, save, screen, scroll, search, select, send, server, sheet, shopping cart, shortcut, site, sleep, style, surf, tab, thumbnail, tool, trash, turn on/off, virus, wall, web, workplace, worm, write, zoom…

One might at first be inclined to suppose that many of these terms are deceptive — that is, that they are not really based on helpful analogies to older ideas, and that their new and old meanings are no more intimately related than are the two different meanings of “race” (“a running competition” and “a subspecies”) or the two different meanings of “number” (“a quantity” and “more anesthetized”). But this idea is far from the case: the semantic connections linking the new technological meanings to the meanings that originated in far older domains are in every case quite straightforward. Aside from a couple of terms such as “mouse” and “chip”, where the analogy between the everyday notion and the technical one is little more than a visual resemblance, what indisputably lies behind each of these familiar terms is an abstraction linking the new and the old uses.

Sometimes one finds much charm in definitions that were drawn up in a long-gone era when today’s technologies could not have been imagined by anyone. For example, the 1932 Funk & Wagnalls New Standard Dictionary, previously quoted in Chapter 4, defines the words “browse”, “folder”, and “hacker” as follows:
browse. trans. verb To feed upon, as twigs, grass, etc.; nibble off; also, sometimes, to graze; as, the goat browsed the hedge.

The fields between
Are dewy-fresh, browsed by deep-udder’d kine.

TENNYSON The Gardener’s Daughter, stanza 3.

intrans. verb To eat the twigs, etc. of growing vegetation; graze.

Wild beasts there browse, and make their food
Her grapes and tender shoots.

MILTON Psalm XXX, stanza 13.

folder. 1. One who or that which folds. Specif.: (1) A flat knife-shaped instrument for folding paper. (2) A map, time-table, or other printed paper so made that it may be readily spread out. (3) A leaf, as one containing a map, larger than the other leaves of a book into which it is folded and secured. (4) A folding-machine. (5) A folding sight on a firearm. (6) An addition to a sewing-machine which folds the material before it is sewn.

hacker. A tool for making an incision in a tree to permit the flow of sap.

It’s rather astonishing to think that less than a decade before the invention of the digital computer, in an era when radios and movies, cars and airplanes, telephones and record players already existed and were even commonplace, the three concepts above were conceived of in ways that today strike us as incredibly concrete, limited, and quaint; indeed, if someone from back then could visit us today, it would take us a little while to explain to them the modern technological senses of these words. There is no doubt that the old and new meanings are cousins, and yet it would take a nontrivial crash course and some major leaps of the imagination to spell out some of the analogies that bridge the gap, but once one sees the common abstract core, it is very clear.

Consider, for instance, this possible contemporary definition of “hacker”: “A person who invades a data base to permit the illicit flow of information”. The analogy with the 1932 definition has deliberately been made salient, and yet the conceptual gap is still a huge mental stretch, somewhat reminiscent of the vast conceptual gap between California’s rural and picturesque Santa Clara Valley in the post-Depression years, dotted with scenic orchards and small farms, and what it became just a few decades later: the ultra-modern bustling metropolitan area known as Silicon Valley, jam-packed with high-tech firms, criss-crossing freeways, upscale housing developments, and more Thai, Indian, and Chinese restaurants than you can shake a stick at.

Your Trip Has Been Placed in Your Shopping Cart

The emergence of modern computer-oriented meanings for many words not only has given rise to new categories, but has helped to make the abstract essences behind
the terms in the above list become clearer (as was discussed in Chapter 4). As we already saw in the case of the word “desk”, the original meaning of a term gives rise to an analogical extension, allowing the old concept to apply to virtual (or “software”) entities; this extension in turn simplifies and refines the old category, in the sense that some of its once-central qualities are now seen as superficial and thus merely optional. For instance, the virtual concept of address has made us understand that physical location, standardly tied tightly to a postal address, is not crucial to this concept, but that what matters, in fact, is accessibility through a symbolic label. When a message is sent to an electronic address, what matters is simply that it should reach the person concerned, and the geographical whereabouts of said person are of as little import as whether a desktop is made of wood or pixels or whether it has drawers or not. In short, the recent appearance of a digital version of the category address has made us see that geography is not relevant to the term, even though up till then, geography had been absolutely indispensable to the notion.

Certain curious phrases that crop up now and then on Web sites, such as “Your trip has been placed in your shopping cart”, reflect this tendency of categories’ essences to be revealed ever more clearly as categories get analogically extended. Thus an online shopping cart shares with a physical shopping cart the property of accumulating potential purchases, about which one can change one’s mind at any moment until one comes to the “checkout stand”. In real life (i.e., in “the good old days” before computers became a central reality in our lives), it would seem absurd, to say the least, to speak of “placing a trip in one’s shopping cart”, but what would be absurd for the old and narrow categories of trip and cart need not be so for the new and broader categories, because once a trip has become a potential purchase, and once a shopping cart has become a repository for potential purchases, then the phrase “to place a trip in a shopping cart” seems perfectly innocuous — indeed, quite sensible.

Finding a new category that gracefully combines the pre-technological and the technological versions of a concept often requires one to jump to a relatively high level of abstraction, as is the case, for example, for firewall, hacker, peripheral, and port. Thus in the on-line world, a firewall is a protection against hackers, whereas in a more traditional context, the phrase “a firewall against hackers” has little or no meaning. Similarly, in a computer context, it makes perfect sense to speak of “plugging a peripheral into a port”, whereas in the pre-computer world, such a phrase would merely have sounded like meaningless gibberish.

And yet there are other categories, such as address book, keyboard, move, screen, delete, and send, which scarcely seem to have been extended at all to incorporate their new aspects. One might think that these categories have emerged unscathed from all the technological upheavals, that they have withstood the earthquake without being touched in the least. However, the truth of the matter is a bit subtler than that. Thus, for instance, an electronic address book, unlike its paper forebear, is not a concrete object. One “writes” entries in it not with a pen but by typing (or possibly by speaking aloud!), and it allows one to do certain things in a flash that would be very laborious with an old-fashioned address book, such as finding a person given their telephone number.
Asking whether a virtual address book belongs to the category called “address book” or simply is analogous to an old-style address book is a question that needn’t be answered, because the dichotomy it presumes is a false one. Belonging to a familiar category and being analogous to a familiar thing are not black-and-white matters, and should not be thought of as opposites, or as excluding each other; both have sliding scales (or shades of gray) that depend on both perceiver and context; indeed, strength of category membership comes down to nothing but strength of analogousness. The gradual and natural extension of technological terms provides an excellent illustration of the fact that analogy and categorization are just two sides of the same coin.

What about the concepts expressed by verbs such as “move”, “erase”, and “send”? It might seem that their meanings haven’t budged an inch as a result of the computer revolution. And yet, would you say that “moving” a file from one folder (on a hard disk) to another is exactly the same thing as moving a paper file from one cardboard folder (in a wooden drawer) to another? Or would you say that highlighting (on a screen) the set of pixels forming an alphabetic character and then hitting the “delete” key is exactly the same thing as quickly and forcibly rubbing the pink end of a pencil back and forth across some marks (on a piece of paper) until they are barely visible any longer? And would you say that sending an electronic message is exactly the same thing as sending a letter via “ordinary” mail? It’s easy to forget that when one sends a material letter, one has to relinquish it physically, whereas when one sends a message electronically, the original remains on one’s hard disk. And so we are reminded once again that even when it comes to terms that, on first blush, seem completely unextended by their computer versions, the truth is that the categories in question have indeed been broadened, and it is only thanks to analogy-making with previously familiar everyday categories that these extensions could take place.

The Best Interface is No Interface at all

Do the hundred-odd terms given a few pages back, all resulting from analogical extensions, coexist with all sorts of other new terms that did not come from analogical extensions? The fact is that whenever a new technology comes along, the standard way of devising a new set of terms that work naturally with it is to borrow pre-technological terms and to rely on the predictable naïve analogies that most people tend to make. Anyone who doubts this should just listen to computer people talking and take note of terms that didn’t exist fifty years ago. They will discover that everyday down-to-earth words are ubiquitous, while terms that are unique to the new technologies and that are nowhere to be found in old dictionaries, such as “motherboard” or “pixel”, are not all that frequent. To be sure, if one were to transcribe a technical conversation between two computer specialists, one would find a rich harvest of acronyms and other narrow technological terms, just as one would for any specialized discipline, but there can be no doubt that ordinary people, in speaking about their computers, are constantly exploiting terms that hark back to a day long before anyone could have dreamed of the abstract uses to which those terms would be put, decades or centuries hence.
Thus concepts from the world of computers now permeate our daily lives because our down-to-earth concepts, through hundreds of naïve analogies, have permeated the technology itself. These naïve analogies, building as they do on extremely familiar categories rooted in mundane daily activities, allow us to endow complex and mysterious technological entities with all sorts of simple and unmysterious properties, and they do so at minimal cognitive cost. Terms based on naïve analogies catch on easily because they naturally bring out qualities that otherwise would be highly elusive.

Thus thanks to our naïve analogies, we all speak with ease and accuracy of “placing objects on the desktop”, of “inserting documents into a folder”, of “opening or closing a window”, of “moving, copying, or throwing away a file”, and so forth — and we didn’t need a course or an instruction manual to gain this skill. Moreover, this down-to-earth lexicon for describing the behavior of abstract technological entities (such as desks, files, windows, and documents) is not just a rough-and-ready set of linguistic tools, serving solely to help novices get their feet wet but then to be summarily dropped; to the contrary, even the most technically savvy people speak in this concrete manner. They, too, open and close windows, files, and folders — and when using such phrases, they feel they are expressing themselves perfectly straightforwardly and non-metaphorically.

Experts in the field of human–machine interface design have stated that the best possible interface should be invisible — indeed, that “the best interface is no interface at all”. Such assertions, stressing the importance of natural and intuitive interfaces, mean essentially that designers should always try to exploit analogies to familiar things. Only if this is done will the interface become “transparent”, which means that users will feel as if they are manipulating everyday objects, a feeling that frees them up to concentrate on their main goals. Interfaces designed in this felicitous manner convert the computer into an easy-to-use tool for accomplishing a particular type of task. Instead of working up a sweat figuring out how to use the tool, one simply concentrates on the task itself.

Interfaces carefully based on analogies to familiar activities do not suffer from the great awkwardness of poorly-designed interfaces. Donald Norman, who is not only a distinguished cognitive psychologist and error-collector, -classifier, and -modeler, but also a pioneer in human–machine interface design, has stated this idea succinctly: “The real problem with the interface is that it is an interface. Interfaces get in the way. I don’t want to focus my energies on an interface. I want to focus on the job... I don’t want to think of myself as using a computer, I want to think of myself as doing my job.”

The Naïve Side of Naïve Analogies

Although indispensable, naïve analogies that help us to relate to new technologies have their limitations, because members of the new category will sometimes behave differently from those of the older, more familiar category. In such cases, the naïve analogy is likely to lead one down a garden path. After all, when one depends on a naïve analogy, one does so, by definition, naively — that is, lock, stock, and barrel. For better or for worse, the naïve analogy is one’s only guide — and on occasion it will mislead. In a word, we are back again in the land of categorical blinders.
Naïve Analogies

When the differences between a virtual category and its old-fashioned analogue do not involve the categories’ most central aspects, there generally is no problem. For example, it’s obvious that virtual desktops, files, and folders have no volume or mass, cannot get dirty, and aren’t subject to wear and tear. This is because categories such as virtual file and virtual folder are immaterial, just as is the category virtual desk (or desk2).

It’s another matter, though, when the discrepancies involve central aspects of the familiar category; the naïve analogy can then give rise to serious confusions. For instance, at one time Apple systems required users who wished to eject a disk to drag its icon into the trash. Many users balked at doing so: it struck them that there was a fair chance that taking such an action would delete all the data on the disk. The analogy behind this reaction was so natural and irrepressible that even experienced users couldn’t help but feel a twinge of uncertainty when dragging their disk into the trash, as if this operation, no matter how many times they’d done it before, was still just a tiny bit risky: “Uh-oh… Could it be that this time all my files will get erased and will be lost forever?” It’s as if they were imagining that the computer itself was thinking by analogy, and that it might get confused like a human and, by error, throw all the disk’s contents away. (“Oops! Sorry about that! I got a little distracted and when I saw you’d dragged the disk’s icon into the trash, I just tossed everything on it out. Silly me! I’d forgotten that for disks I’m not supposed to do that, but should just eject them.”)

In light of this common fear, should one infer that Apple’s user-interface designers suffered from a fleeting “senior moment” when they decided that dragging the disk icon into the trash was the natural way to say, “Please eject the disk”? Not really. They were simply presuming that users would easily jump to a higher level of abstraction than obliteration when they dragged the disk icon into the trash — say, to a concept such as getting rid of something no longer relevant. However, as it turned out, this presumption overestimated the typical user’s mental fluidity. When the designers finally realized this, they removed the ambiguity in subsequent versions. Nowadays, whenever the icon for a disk is brought near the wastebasket icon, the latter magically mutates into a different icon denoting ejection rather than destruction.

Another striking example of the failure of a naïve analogy involves the virtual desktop. Usually, the hard disk is represented as sitting on the desktop (or possibly in the workspace, which is itself on the desktop). But the fact of the matter is that all the data in memory are stored on the hard disk, and this includes the entire desktop. There is thus an apparent paradox, with the disk being on the desktop and the desktop being inside the disk. What sense does it make for A to be on top of B, while B is at the same time inside A? This shows that at times naïve analogies cannot fully do the job they were intended to do. To be sure, we humans can live with small inconsistencies on our computers just as we do in life in general, but sometimes this little paradox does cause genuine confusion, as when a user wants to locate the desktop in the computer’s memory. Interface designers, after recognizing the possible confusion due to this naïve analogy, eventually made a patch. Today the hard disk is no longer shown as being located on the desktop; instead, it is simply accessible through it. Still, it took designers some twenty years to take care of this small problem.
The World of Computers Yields Analogy Sources for Itself

Not all naïve analogies designed to help people use computers more easily are rooted in pre-computer experiences, because today computers are familiar enough that some of their best-known properties can themselves be exploited as sources for naïve analogies. Indeed, something that was once understood only by analogy can eventually become familiar enough that it can act in turn as the source for new analogies. This happens not only with technological devices, but in all aspects of life. For example, sound waves, which were first understood by analogy to water waves, became in turn, many centuries later, the analogical basis for understanding light waves.

The world of computers is thus starting to yield sources for its own analogies. For example, the notion of a floppy disk, which for many years was the standard device on which one saved all one’s files, was originally understood by analogy to a vinyl record. But once floppydisks had become familiar to all computer users, thanks to their widespread use in the 1980s and 1990s, they became the source for new analogies. And thus even today, the icon that stands for saving a file is frequently a stylized picture of a floppy. This is ironic, since writing a file onto a floppy disk is almost unheard of today; floppydisks have long since been supplanted by internal and external hard disks, CDs and DVDs, flash drives, and so forth (each of these, after a brief day in the sun, gave way to new technologies). Although the floppy-disk icon is still found in some software, it is a remnant of a bygone era — a bit like the stylized pictures of ancient bicycles with huge front wheels that were sometimes used, not too long ago, to indicate bike lanes in the U.S. — and it’s a safe bet that this icon will soon go the way of floppydisks themselves. As the objects themselves are no longer around, the concept of floppy disk is approaching extinction. Children who see the square icon don’t know where it comes from, and the recent tendency is to make the icon for file-saving look like a hard disk instead.

When the Virtual World Helps Us to Understand the Real One

Because of their constant and increasing presence in our lives, computers and related technologies have recently turned into a rich source of categories that, through their great familiarity to us, can serve as rich new sources for analogies. This is a curious twist, since computers, for most of their brief existence, have standardly been explained through analogies to phenomena in the physical world, but today, the reverse is happening: that is, physical things are coming to be described through analogies to phenomena in the world of digital technology. For example, a recent television ad for an SUV crowed, “Think of it as a search engine helping you to browse the real world!” Who would have predicted such a reversal of roles? This tendency will surely increase, ushering in unpredictable changes in society.

Take the concept of multitasking. This was a clever invention of the 1960s allowing a computer to execute several distinct processes concurrently by breaking each process into tiny steps and doing a single step of process #1, then a step of process #2, and so on, thus interleaving the various processes so finely that, to all appearances, they are all
being carried out simultaneously. But in our lives today, the concept of *multitasking* is routinely applied to human beings and their activities. Thus, sentences like “As a single mom, believe me, I’m constantly multitasking!” and “Every morning on my way to work, I sip my mocha, yak on the phone, savor the scenery, and listen to music, all while driving my car—I’m such a multitasker!” are standard parlance. Indeed, the computer origin of the term has begun to fade out of view. Here are definitions of it taken from two dictionaries a couple of decades apart:

*Webster’s New World Dictionary* (1988):

*Computer science*: the execution by a single central processing unit of two or more programs at once, either by simultaneous operation or by rapid alternation between the programs.


1. The concurrent operation by one central processing unit of two or more processes.
2. The engaging in more than one activity at the same time or serially, switching one’s attention back and forth from one activity to another.

As this comparison shows, the term was purely technical in the 1980s, whereas today many people use it fluently for everyday activities, having no idea that it came from computer science or even has *any* technical application! There was a transition period where the analogical extension was explicitly felt by people who knew they were stretching the concept, but after a while, the stretching had been accomplished, and the term, stripped of its original technical connotations, entered the public lexicon.

Another computer term that has been imported into daily life is “to interface”, which originally meant adapting two pieces of hardware or software so that they would work together seamlessly, but which is now used in such nontechnical phrases as “The gay community needs to interface much more with the black community.”

The term “core dump” was used for decades to mean a printout of the entire contents of a computer’s main memory (once called “core”); this was a somewhat desperate measure that could help in pinpointing very recalcitrant bugs. But it was analogically extended to daily life, with the result that today nontechnical people say things like, “Sorry for going on and on so long—I didn’t mean to give you a brain dump!” What is preserved is the abstract idea of visibly or audibly outputting a huge amount of information that normally is invisibly stored in some kind of memory device.

Another computer concept that has recently enjoyed considerable popularity as the source of casual conversational analogies is *cut-and-paste*. Thus, a television newscaster describes a political candidate’s speeches as being “cut-and-pasted from her previous speeches”, a newspaper describes attempts to cut-and-paste Silicon Valley into various European countries, and a book reviewer criticizes a new book by saying, “This book is just a cut-and-paste of other books on the same subject; I learned nothing new from it.”

The notion of *debugging a computer program* is yet another fertile source of imagery for everyday life. Thus a salsa dancer says, “I’m working on debugging my Latin hip
motion — my hips always move in the wrong direction”, while a Chinese teacher says, “You really have got to debug your tones — they’re all mixed up!”

And finally, a few miscellaneous examples of the insidious manner whereby computer terminology worms its way into everyday discourse. One business executive confides to another: “I’d really like to have your input on this matter.” A medical-school student sighs: “Every so often, I just need to disconnect from this crazy routine.” An advertisement exhorts: “Reboot your brain with a caffeine nap!”

Taken together, these examples provide another pillar of support for the thesis of the first two chapters that word choices are made via analogies, and that word-choice analogies are usually (though not always) made unconsciously. Take the computer-science term “multitasking”, for instance. In the first few years of its existence, the idea of applying it to some kind of human behavior was a choice available only to computer-savvy people, and for them the computer-to-human analogical bridge was built with deliberateness and delight. Pushing a word’s range outwards is fun (because making inventive analogies is fun). Eventually, however, this extended sense of the term leaked, by osmosis, into the much broader community of non-computer-savvy people, and at that point, using the term didn’t require building an analogical bridge linking a human behavior to an arcane computational trick (after all, these speakers knew nothing of that trick); the analogical bridge was simply to human behaviors that had been labeled by the catchy new term. But no matter which set of speakers we focus on — technically savvy or technically non-savvy — we see that it’s always analogy in the driver’s seat, always analogy that is handing words to speakers, and often handing them words on such convenient silver platters that, to them, their word choices seem to have taken place instantly, naturally, effortlessly, and without any help from analogy-making. That silver platter was just magically there. Of course, that’s an illusion — just another case where the human mind, not surprisingly, fails to be aware of the seething activity constantly going on below the surface of its familiar linguistic behavior.

Technomorphism — an Analogue to Anthropomorphism

Sometimes it’s not the choice of a computer-oriented word that betrays the tendency of technical ideas to slip into everyday thinking, but simply a computer-oriented habit that unconsciously pops up and tries to insinuate itself into an unfamiliar new situation. After all, when one is constantly dealing with technology and using the Web on a day-in day-out basis, one can’t help starting to see this familiar old world with fresh new eyes. This can happen, for instance, when one is trying to find a favorite passage in a book. While randomly flipping through its pages, seeking the desired passage much like someone seeking a needle in a haystack, one can easily grow very frustrated, since one is painfully aware that if one only had an on-line version of the book, it would be a piece of cake to find the desired passage.

The following anecdotes illustrate the very human and very natural — indeed, irresistible — tendency for computer-based concepts to pop to mind in everyday situations as if they applied, when in fact they don’t apply at all (or at least not yet!).
A woman was driving her ten-year-old son, a video-game addict, to another city. After several hours, the boy started complaining of being bored to death. His mother replied, “Would you please stop your whining? I’ve got to concentrate on the road.” Her son shot back, “But it’s not fair — you get to play the whole time!” For the boy, who had virtually driven virtual cars for several years, the act of real driving was exactly like playing a video game, and thus to him, it seemed that his mother had the good fortune to be having lots of fun while he was unable to do so.

A little girl walked into a room where two digital photo frames were sitting on a buffet, each of them periodically flashing one image, then another, then another, and so on. Between them was an old-style frame containing a standard still snapshot. She commented, “Look, Daddy, the frame in the middle is broken!” For her, the old-style frame was automatically seen as a member of the category digital photo frame, and as such, it was clearly a defective member of the category.

An eight-year-old girl was in her family’s car when a heavy rain started to come down as they were driving through the countryside. She remarked, “Watch out, Dad, you can’t see anything on the screen any more!”

A computer addict confessed that when he was sitting at his computer and a fly landed on the screen, his knee-jerk reflex was to try to get rid of it by clicking on it and then dragging it off the screen.

When you’re at home and can’t find an object, you often wish you could simply search for it just as one can search for things on one’s hard disk or on the Web, or for that matter on one’s cell phone. If only you could type in a couple of key words, you feel, you could instantly retrieve any lost article! Take Alice, for example. She was very groggy when she woke up, and badly needed her morning coffee. While drinking it, she got a phone call from her mother, so she set the cup down. When she hung up, she’d forgotten where she’d placed it, so she blithely picked up her cell phone and started typing the word “coffee” in the “Search” slot of her list of contacts. Then there’s Bob (another example we found on the Web), who for five straight days had been working all day long at his computer, trying in desperation to finish a grant proposal before the deadline. At some point, he misplaced his glasses and realized that he couldn’t squelch the desire to type a couple of words into a search engine to find them. Examples of this sort are a dime a dozen on the Web.

Unfortunately, no tool at that level of sophistication exists to help us locate lost items in the physical world, and for that reason the on-line world’s concept of search, which by now is second nature to nearly everyone, has become the source of analogies in the physical world, rather than the other way around. It is very tempting to think that objects that we’ve misplaced over and over again, such as our keys, our wallet, our checkbook, or our glasses (or their case), should be just as easily coaxed to chime out their presence to us as is our misplaced cell phone. In short, we’d like to be able to summon anything that we can’t find and have it instantly chirp back at us, telling us just where to find it.
By now, the Web and cell phones have given us the sense that pretty much anything should be within easy reach at any time. Wherever we are located, just about anything that matters to us is within clicking range, emailing range, cell-phoning range, or text-messaging range. This gives us the illusory feeling that anything of any sort should be available to us instantaneously.

The great ease of obtaining things from afar these days has had the effect of reinforcing the sense of presence of people who are in fact absent. Some people, for example, have started to feel ill at ease about making the slightest critical comment about anyone, any time their cell phone is nearby — as if the person being criticized might somehow overhear the comments even without any call having been initiated. This scenario is a bit reminiscent of living in an apartment with very thin walls, where one never knows for sure whether one might be overheard by others in neighboring apartments. Along much the same lines, certain people, when in a chat room on the Web, often start whispering to physical people who are physically near them, as if the virtual occupants of the virtual room could overhear every word they say aloud. In short, some computer addicts develop the jittery feeling that the physical world all around them is populated by virtual people who are able to overhear conversations about themselves.

One very useful feature of computers is that they offer us all sorts of chances to undo mistakes that we have made. Thus when we are using a word processor or a photo editor, it seems only reasonable that we should instantly be able to reverse any action that we’ve done — including massive deletions carried out by accident — and we get very used to such luxuries, taking them for granted. The habit then spills over into other domains of life and becomes an expectation, and when we realize there is no “undo” button in nearly any of them, we become frustrated. Here are two anecdotes illustrating this tendency:

A student who was cooking a cake had put too much flour in her dough, and she would have liked to go back a few moments in time and undo her mistake. All of a sudden, she caught herself wishing that she could just push the “undo” key that she was so accustomed to.

A teen-ager who was plucking hairs from her eyebrows suddenly realized that she’d done more than she intended, and she thought to herself, “Oh, no problem, I’ll just revert to the older version.”

Not only the act of “undoing” but many other types of frequent computer actions can become such strongly ingrained habits that they wind up shaping the way people see and behave in the material world, as the following set of anecdotes demonstrates:

A teen-ager was looking at a photo in a magazine in which she recognized some people but not others, and she said, “My first reflex was to wonder, ‘How come they’re not tagged?’ Then I remembered I was looking a magazine, not a Facebook page.”
A man reported that one time when driving he found himself using two fingers on his rear-view mirror in an attempt to blow up the image. A woman replied that “the same thing” had happened to her but on her flat-screen television, and yet another said “exactly the same thing” had happened to her when she was looking at her reflection in her bathroom mirror.

An avid moviegoer said that she wanted to go to the restroom during a film and for a split second she found herself trying to grab her mouse to stop the film momentarily.

Another film buff confessed, “When I was in the movie theater, I tried to jiggle my mouse a little bit to see how many minutes were left in the movie.”

A schoolboy said that during a quiz, he wasn’t sure whether he had spelled a certain word right, so he waited for a few seconds to see whether what he’d just written in pencil on paper was going to get underlined in red.

A college student described how she was cramming for an exam and had finally, with great effort, figured out how a certain complex biochemical reaction worked. At that point, she was sitting in front of her computer, and she mindlessly hit the pair of keys that she always would hit when saving a file.

These days, amusing mental contaminations due to these types of crossed wires involving computer concepts and pre-technological concepts are a dime a dozen, and there are Web sites aplenty “where” (if we may analogically borrow this concept from the physical world) people gleefully report and laugh at their own gaffes of this sort.

**Some Equations Are More Equal than Others**

Now that we have taken a careful look at the ways that naïve analogies originating in recent technology have insidiously invaded our lives, we can turn back to the field of education, focusing specifically on the role that naïve analogies play in how children pick up basic mathematical concepts in school.

Is the equation “3 + 2 = 5” completely clear? Is there just one way to understand it? Do all educated adults understand the equals sign in the same way? Theoretically, an equation symbolizes a perfect equivalence or interchangeability; that is, an equals sign tells us that the two expressions flanking it stand for one and the same thing. The notion of equality, when described this way, seems so simple and straightforward that it would seem hard to imagine any other way of interpreting it. And yet there is another side to the notion of equality, and it comes out of a naïve analogy that we will call “the operation–result analogy”.

In this alternate interpretation, the left side of an equation represents an *operation*, while the right side is the *result* of the operation. This is a naive analogy in which equations are tacitly likened to processes that take time, and it crops up in situations that have nothing to do with school or mathematics, and which influence everyone, including young children. For instance:
point at + cry = obtain a desired object
vase + knock over = shards of glass on the floor
mud + hands = mess
DVD + DVD player + remote control = watch a movie
chocolate + flour + eggs + mix + bake = cake
cheese + lettuce + tomato + bread = sandwich
3 + 2 = 5

Here, the equals sign is a symbol that links some sort of action in the world to its outcome, and it can be read as “gives” or “yields” or “results in”. When seen that way, “3 + 2 = 5” is not the statement of an equivalence at all; rather, it expresses the idea that the process of adding 3 and 2 results in 5.

The ideas of interchangeability and operation–result are different. The second point of view clearly embodies an asymmetrical conception of equations, in which the two sides play different roles, one side always standing for a process and the other always representing its outcome. To write “5 = 5” would be incompatible with this viewpoint, since no process is indicated. Likewise, writing “7 − 2 = 8 − 3” is also troublesome, since now there is no result. And lastly, writing “5 = 3 + 2” would be disorienting, because the operation and its result occupy the wrong sites. Indeed, many first- and second-graders understand equality in just this fashion, insisting that “5 = 3 + 2” is “backwards”, and that “7 − 2 = 8 − 3” makes no sense because “after a problem there has to be an answer, not just another problem”. Some even balk at “5 = 5”, replacing it with something such as “7 − 2 = 5”.

The operation–result naïve analogy guides children before they encounter the concept of equivalence, because the notions of a process and its result are familiar even to toddlers. These notions are close cousins to the notions of cause and effect, as well as to the idea that certain means have to be used to reach certain ends.

Although today’s children may acquire a fairly decent understanding of equality in elementary school, coming up with the symbol “=” took a long time for humanity as a whole. A symbol for equality in mathematics first appeared only in the year 1557, in a book by the Welsh mathematician Robert Recorde. He wrote:

I will sette as I doe often in woorke use, a pair of paralleles, or Gemowe lines of one lengthe, thus: ==, bicause noe 2. thynges, can be moare equalle.

The word “gemowe” means “twins”, and the “twinnedness” of the upper and lower horizontal lines was intended to symbolize the general idea of equality. The fact that a symbol for equality took such a long time to occur to anyone, even though mathematics had existed for at least two millennia, reveals that it is far from a self-evident notion.

Although for many adults today the idea that “equality equals equivalence” may seem obvious in a mathematical context, it doesn’t follow that the operation–result view of equality has disappeared from their minds. In fact, people often write down, and read aloud, equations in a way that reflects their unconscious understanding. For instance,
in reading “4 + 3 = 7”, many people will say “four plus three makes seven”, whereas for “7 = 4 + 3” they might say “seven is the sum of four and three”. If education always resulted in equations being seen as statements of interchangeability, then by the end of high school, the operation–result view of the equals sign would surely have disappeared for once and for all. The order of the two sides in an equation would be completely irrelevant, and both ways of writing an equation down would elicit exactly the same commentary. However, this turns out not to be the case. Let’s take a look at some specific cases, starting out with some that are very remote from mathematics.

**Naïve Equations in Advertisement**

It’s standard practice for advertisements to appeal to the child inside each of us rather than to the budding mathematician. Here are a few examples of “equations” culled from real ads:

- buy two items = 50% off on the second one
- buy a pair of glasses = get a pair of sunglasses for free
- buy a loyalty card = free home delivery for a year
- buy a TV set = a DVD player for just $1
- buy any pizza = get another one free

Just to convince ourselves that interchangeability is not the idea behind these equations, let’s flip them around. As you will see, the resulting “equations” sound utterly silly, even nonsensical.

- 50% off on the second one = buy two items
- get a pair of sunglasses for free = buy a pair of glasses
- free home delivery for a year = buy a loyalty card
- a DVD player for just $1 = buy a TV set
- get another one free = buy any pizza

As these examples reveal, people’s first glimmerings of understanding of the equals sign come from the naïve analogy of an operation followed by a result, and even if the concept of interchangeability gains some ground in the course of twelve years of education, the operation–result viewpoint is never fully eradicated. It can always be coaxed out of dormancy when the right cues are presented. We thus see that education does not eliminate the first naïve ideas about a mathematical notion – even one that we tend to think of as completely trivial because it is taught in elementary school, when children are only six or seven years old. In childhood and even when one is fully grown, the naïve view coexists with a different view, which is instilled at school but which is also dependent on a pre-existent and familiar notion: that of *same thing* (that is, *identity*). The transition from the earliest viewpoint (operation–result) to the more sophisticated one (identity) does not obliterate the earlier viewpoint, which remains
triggerable, and on which one frequently relies on a day-by-day basis, sometimes even in scientific contexts, such as mathematics or physics, as we shall now see.

Of Equations and Physicists

For physicists, the most fundamental formula of classical mechanics is doubtless Newton’s second law, which describes how a force affects the motion of an object. The basic idea of this celebrated law is compatible with the naïve analogy that says that one side of an equation should represent a process, with its other side representing the result of that process. In this case, the process (ideally occupying the left side of the equation) would be the action of a force of size \( F \) on a mass of size \( m \), and the result (ideally on the right) would be an acceleration of size \( a \) imparted to the mass. Rendered symbolically, this yields the equation \( \frac{F}{m} = a \). Unfortunately, though, Newton’s law is virtually never written this way. Instead, it is almost always cast as follows: \( F = ma \). This famous formula is quite confusing to many students, since neither side of it cleanly symbolizes either the process or the result. The alternative notation \( \frac{F}{m} = a \) encodes the naïve analogy much more clearly, and would therefore be easier for students to relate to, but it is seldom if ever found in textbooks. From a purely logical standpoint, these two versions of Newton’s law are completely equivalent and interchangeable, but from a psychological and pedagogical standpoint, they certainly are not.

Luckily, physicists are often sensitive to such psychological pressures, and most of the time they try hard to cast their equations in the form of clean and clear cause-and-effect relationships, with one side giving rise to the other side. Take, for instance, the first of Maxwell’s four equations for electromagnetism:

\[
\text{div } E = 4\pi \rho
\]

where \( E \) represents an electric field, \( \rho \) represents electric charge density (basically, a description of how much electric charge there is in each point of space), “div” stands for a certain operation in differential calculus called the “divergence”, and \( \pi \) is the familiar circular ratio 3.14159…

This formula is universally seen by physicists as saying, “A certain distribution of electric charges in space (the cause) always gives rise to a certain pattern of electric fields in space (the effect).” For historical reasons, however, the cause (the charge distribution) is conventionally placed on the right side of this equation and the effect (the electric field) on its left side, thus reversing the usual operation–result order. Why do physicists always write it in this flipped fashion? That’s hard to say, but basically it’s just a harmless “professional deformation”. In any case, Maxwell’s first equation intuitively embodies a physical cause-and-effect relationship, with the cause on the right side and its effect on the left side. (In fact, all four of Maxwell’s equations embody similar cause-and-effect relationships, and they all have this same kind of right-to-left causal flow.)

There is, however, another way of looking at Maxwell’s equations. For concreteness’ sake, let’s once again consider the first one, as shown above. It says that if
you calculate the divergence of the electric field, you will obtain the charge density. Now such a calculation can also be seen as a kind of cause-and-effect or process-and-result relationship, wherein certain quantities are fed into a calculating machine that churns for a while and eventually outputs new quantities. Seen this way, the “cause”, or initial event (namely, the feeding of input values into the computing device), is always on the left side, while the “effect”, or subsequent event (namely, the number that the device spits out), is always on the right. So now we have a left-to-right causal flow!

But one must keep in mind that this is only a mathematical kind of causality, meaning you can calculate the charge density if you’re given the electric field everywhere in space. However, as we pointed out above, the equation can also be read as a physical kind of causality, asserting that if you arrange electrical charges in a certain way in space, you will always find that a specific pattern of electric fields surrounds them: in short, the charges produce the fields. When the equation is read in this latter way, the causality flows from right to left (i.e., from charges to fields). And that’s how physicists view this equation, whether they do so consciously or unconsciously. Indeed, it would strike a physicist as absurd if someone were to say that Maxwell’s first equation means that an electric field spread out all over space gives rise to a tiny electric charge sitting somewhere. That would sound as backwards as saying that a strong stench wafting all through a neighborhood will give rise to a frightened skunk crouching under a bush (note the use of a caricature analogy here).

In summary, the equals signs in Maxwell’s equations can be understood either as expressing physical causality (a physical cause giving rise to an effect), when they are read from right to left, or as expressing mathematical causality (a calculation giving rise to a result), when they are read from left to right. And Maxwell’s equations are in no way exceptional. Physicists always try to manipulate their equations so that they will have this quality — namely, with causes on one side and effects on the other. Doing so is certainly not logically necessary, but it contributes greatly to clarity. For example, here are two alternative ways of writing Maxwell’s first equation that are both perfectly correct yet would make physicists scratch their heads in puzzlement and ask, “What’s the point of writing it that way?”

\[
\text{div } E / 2 - 2\pi \rho = 0 \quad \text{div } E / 4\rho = \pi
\]

Indeed, these equations both cloud up the crux of the law, which is the fact that one phenomenon gives rise to another.

In short, physicists, no less than other people, have a weakness for, and also derive benefit from, the naïve analogy likening equations to cause-and-effect relations.

**Does Multiplication Always Imply Getting Bigger?**

For some concepts that one learns in school, there is an early naïve analogy that is very helpful, but there is no other familiar category that helps one develop the concept more deeply. In such cases, the naïve analogy will very likely be one’s only means for
grasping the concept, and it will retain this primary role even after many years of education. In such cases, refining the concept so that it becomes more general and flexible will be far harder. It so happens that multiplication and division, two of the most basic notions in mathematics, are cases of this sort.

Addition, subtraction, multiplication, and division are taught in elementary school and are presumed to have been fully understood by middle school. Since so many other mathematical notions are built on them, they are often called the four basic arithmetical operations. These classic notions feel as if they are part of the cultural heritage of every member of our society, and any adult claiming to have no idea what multiplication or division is would be looked at askance. Taught all the way through childhood, these notions should be clear as a bell to high-school and college students. And yet the belief that these operations have been mastered by most adults is an illusion. The next few sections illustrate how this can be so.

Let’s consider multiplication. We have found, in surveys of many quite advanced university students (we don’t mean advanced math majors), that if they are asked, “What is the most precise possible definition of multiplication?”, they are generally very satisfied with either of the following two definitions that we suggest:

- Multiplying is repeatedly adding a value a certain number of times.
- Multiplication is taking \( a \) times \( b \), which means adding \( b \) up \( a \) times.

It’s hard to find anyone who disagrees in any way with these definitions, and virtually no one sees any way to improve upon them. We have also asked groups of advanced university students to supply definitions themselves, and exactly the same themes reappear. It always comes down to the idea that multiplication means, by definition, adding a given number over and over to itself, counting how many times it is done, even if the formulation is not always as concise or clear as the two definitions offered above. For example:

- A multiplication is the iterated addition of a given number a specified number of times.
- Multiplying means adding a given figure to itself as many times as one is told to do so.
- To multiply is to add a particular number to itself as many times as the other number tells you to do so.
- Multiplication is a calculation in which one is told how many times one should add a given quantity to itself.

On the Web, definitions of this sort abound. One site proposes: “Multiplication is thus nothing but an addition in which the numbers being added up are all equal to each other. This is why we say that it amounts to repeating the multiplicand as many times as there are units in the multiplier.”

For a bit of historical perspective on the question, one can take a look at definitions along these lines proposed by professional experts. Thus in 1821, the renowned French
Naïve Analogies

mathematician Étienne Bezout wrote, in his Treatise on Arithmetic Intended for Sailors and Footsoldiers: “Multiplying one number by another is summing up the first of these as many times as there are units in the other.”

Well, now… is this view of multiplication as repeated addition really as indisputable as it would seem from all the above? Hardly! As a matter of fact, this view is a naïve analogy that falls far short of the target, and in the long run, it is almost sure to lead anyone who relies on it into confusion.

First of all, this view of multiplication requires that one of the two values be a positive whole number, since otherwise “as many times” has no meaning. What would it mean to speak of adding a number to itself $2^{1/2}$ times or $1/3$ of a time, let alone $\sqrt{2}$ times or $\pi$ times? And yet, requiring one of the factors in a multiplication to be an integer should raise suspicions, since everyone knows that multiplying two non-integers is not forbidden; indeed, in school we all learn how to do it, and pocket calculators don’t balk at all at multiplying any two numbers they are given. What on earth would the expression “$\pi \times \pi$” mean if at least one of the two factors had to be an integer?

The next stumbling block lurking in this definition is the common belief that when one adds $b$ to itself and over again, the result will always be greater than $b$. We would not merely expect the result to be somewhat greater than $b$, but, by definition, a times greater than $b$. Dictionaries confirm this naïve idea, as does everyday speech. Indeed, the words “multiply” and “multiplication” suggest a clear image of growth, never an image of shrinking (even though, as we pointed out in Chapter 4, things that shrink can be said to be “growing smaller”). Thus rabbits are said to multiply, quickly resulting in overcrowding; in good times, one’s assets multiply, making one wealthier; in bad times, risks multiply, making one less secure; and so on. The prefix “multi” also tends to make one think of growth, as in words like “multinational”, “multicolored”, “multilingual”, “multimillionaire”, and so forth. However, this preconception runs into a brick wall when one is instructed to multiply by a value less than 1, as doing so yields a result smaller than the multiplicand. This is incompatible with repeated addition.

There is still more trouble. The best-known property of multiplication is that it is commutative — that is, for any numbers $a$ and $b$, it is always the case that $a \times b = b \times a$. Why this should hold for every pair of numbers $a$ and $b$ is not at all obvious if multiplication is conceived of as repeated addition. In fact, the naïve analogy would suggest that multiplication is intrinsically asymmetric, since it treats the multiplicand and the multiplier differently: the former is repeatedly added to itself, while the latter counts how many times the operation is carried out. This certainly does not fit the image of commutativity, in which the two numbers play totally interchangeable roles. Since the naïve analogy gives no insight into this key property, a child (or an adult!) may be baffled by the fact that $a$ added to itself $b$ times always gives the same result as $b$ added to itself $a$ times. To be sure, one can enrich one’s notion of multiplication by arbitrarily tossing in the fact of commutativity like icing on the cake, but the naïve analogy of multiplication as repeated addition makes this fact seem mysterious rather than natural.

These minor stumbling blocks turn into serious obstacles when pupils who depend on the naïve analogy are given word problems to solve. For instance, when middle-
school students in England were given the problem “If one gallon of petrol costs 2.47 pounds, what is the price of 0.26 of a gallon?”, only 44% of them recognized that this is in fact a multiplication problem. The remaining 56% took it to be a division problem (namely, 2.47 divided by 0.26)! And thus a multiplication problem that should be very easy even for elementary-schoolers stumped roughly half of the middle-schoolers.

What happens if one changes the numbers in this problem? If one merely replaces “0.26” by “5” and asks the question again (“If one gallon of petrol costs 2.47 pounds, what is the price of 5 gallons?”), then 100% of the middle-schoolers solve it correctly. This discrepancy is due to the fact that the first problem doesn’t meet the naive analogy’s image of adding a number repeatedly, since the idea of adding a number to itself 0.26 times makes no sense. On the other hand, using the naive analogy of repeated addition works just fine in the modified problem (2.47 + 2.47 + 2.47 + 2.47 + 2.47). Discrepancies between participants’ performances on the two problems reflect the fact that the naive analogy is of no help in the first one, yet is appropriate in the second one.

Adding Thrice and Fifty Times are Different Kettles of Fish

It’s enlightening to compare the preceding findings with some experiments carried out in Brazil. The participants were teen-aged boys who had dropped out of school and were making a living as street vendors. The following simple problem was given to a group of them:

A boy wants to buy some chocolates. Each chocolate costs 50 cruzeiros. He decides to buy 3 of them. How much money will he need?

The same problem was also given to a different group, except that the two numbers were interchanged, as follows:

A boy wants to buy some chocolates. Each chocolate costs 3 cruzeiros. He decides to buy 50 of them. How much money will he need?

To all readers of this book it will surely be trivially obvious that each of the two boys will have to fork over 150 cruzeiros, even if one of them winds up with far fewer chocolates (at least in number) than the other one. When we read the two problems, they appear equally easy; the scenario is the same, they involve the same numbers (50 and 3), and they both involve the same arithmetical operation: multiplication. But were they equally easy for the two groups of street vendors? Not in the least.

The first problem was handled pretty well by most: 75% got it right. The second problem, by contrast, was not solved by any of the street vendors. The reason behind this discrepancy is relatively simple; it comes down to reliance on the naive analogy of repeated addition. To solve the first problem, all one needs to do is add 50 + 50 + 50, to get 150 cruzeiros. This is just two additions, and it involves only very simple facts: first, that 50 + 50 = 100, and second, that 100 + 50 = 150. The variant problem,
however, is another ball of wax entirely. To compute the answer, one has to carry out a very long process of iterated addition — namely, $3 + 3 + 3 + \ldots + 3 + 3 + 3$, involving fifty 3’s. As one can easily imagine, this is not a challenge that an elementary-school dropout would be very likely to be able to handle.

This might seem the moment to shower praise on our educational system, thanks to which we educated adults are all instantly able to solve a problem that to school dropouts seems impossibly hard. We all know in a flash that \(50 \times 3\) equals \(3 \times 50\), end of story. Given this contrast with the 0% success rate of the school dropouts, one might be tempted to conclude that schooling very effectively gets across the true nature of multiplication. However, things are not that simple.

There is no disputing the fact that schooling teaches us that the two numbers in a multiplication can be interchanged — we all know that multiplication is commutative, that \(a \times b = b \times a\) — and we all carry out such switches without a moment’s thought. However, carrying out multiplications using one’s knowledge of commutativity doesn’t mean that one’s understanding of multiplication, as an adult, goes far beyond that of elementary-school children. Indeed, a quick informal survey reveals that almost no one, aside from serious math enthusiasts, knows why it is the case that, for instance, \(5 \times 3\) equals \(3 \times 5\). Middle-school students, high-school students, even university students are generally unable to say why the two numbers in a product are switchable. How, then, do they justify to themselves the idea that five threes equals three fives, or symbolically, the fact that \(3 + 3 + 3 + 3 + 3 = 5 + 5 + 5\)?

Most people, if asked this question, will answer readily that one can check this out in any specific case (“Just go get a calculator and try it out for whatever pair of numbers that you wish!”). Some people will state it more as an axiom: “In multiplication, you have the right to switch the two factors”; others will baldly assert, almost as if some kind of magic were involved, “That’s just how it is” or “Well, it’s known to be a fact.” In short, for most well-educated adults, since multiplication is conceived of as repeated addition, its commutativity appears simply as a kind of miraculous coincidence, lacking any clear explanation or reason.

The above-cited treatise on arithmetic by Étienne Bezout provides a somewhat wordy and obscure justification for the commutativity of multiplication. If one’s vision of multiplication is rooted in the naïve analogy of repeated addition, then asking why the order of the factors makes no difference in multiplication amounts to asking why two different repeated additions give the same answer, and there is no obvious symmetry between the two operations involved. Bezout tries to resolve this dilemma, but his words are not terribly clear:

As long as one considers numbers as abstract entities — that is, as long as one ignores the units attached to them — it makes little difference which of two numbers to be multiplied is taken as the multiplier and which as the multiplicand. For example, 3 times 4 is nothing but the triple of 1 taken four times, while 4 times 3 is the triple of 4 taken one time. Now it’s self-evident that 1 times 4 is the same thing as 4 times 1; and one can apply the same reasoning to any other number.
Most adults consider the fact that $a \times b$ always equals $b \times a$ to be a very useful but unexplained coincidence, which simply is empirically true. They “understand” the commutativity of multiplication in much the same way as they “understand” why bicycles don’t topple over and why airplanes can stay airborne: simply because they’ve seen such things for most of their lives and have long since forgotten that these phenomena are mysteries that crave explanation. And so, although education certainly drills into students the rote fact that multiplication is commutative, it fails to instill a deep understanding of multiplication’s nature; instead, it leaves them dependent on their initial naïve analogy.

What about Division?

For division, too, there is a widespread naïve analogy that dominates people’s thought, profoundly affecting how people conceive of the operation. Although one would tend to think that division is perfectly understood by anyone who has gone through school, this is an illusion.

Here is a very simple “home experiment” that reveals the hidden presence of this naïve analogy, demonstrating how concrete and standard it is. There is no trap here, just a couple of straightforward challenges. The first challenge is a warm-up exercise: Invent a word problem involving division — that is, a problem whose solution requires just one operation of division. Hopefully, this will pose no problem for any reader. Here, for example, are a few division word-problems that were invented by university students:

- 4 friends agree to share 12 candies. How many candies will each one get?
- 90 acres of land is going to be divided into 6 equal parcels. What will the area of each parcel be?
- A mother buys 20 apples for her 5 children. How many apples will each child get?
- A theater has 120 seats arranged in 10 rows. How many seats are in each row?
- A teacher buying food for a class picnic filled 4 grocery carts with a total of 20 watermelons. How many watermelons were in each cart?
- It takes 12 yards of cloth to make 4 dresses. How many yards does it take to make one dress?

Each of these problems involves a starting figure to be divided by something, and the result of the division is always smaller than that starting figure. Thus the first problem reduces 12 candies down to 3, the second problem reduces 90 acres down to 15, the third problem reduces 20 apples down to 4, and so forth.

This observation is already significant, since it shows that when people are asked to invent division problems, they come up with situations where the key idea is “making smaller”. Just as the standard image of multiplication as repeated addition locks in the belief that multiplication necessarily involves growth, here it seems that there is a naïve
analogy at work that locks in the image of division as involving shrinking. And this supposed fact about division jibes perfectly with the way the word “division” is used in everyday speech. When one speaks of dividing X up, one sees X as being broken into pieces, with each piece obviously being smaller than X itself. The 1988 edition of Webster’s New World Dictionary confirms this vision, defining division as follows: “a sharing or apportioning; distribution”. Moreover, division is often associated with the notion of weakening. For instance, the slogans “united we stand; divided we fall” and “divide and conquer” imply that if an entity is divided into pieces, it will be weaker than the original entity.

Well, that was the first of our two small challenges. The second one is simply to invent one more division problem, subject to one extra constraint: the answer must be larger than the starting figure. Readers, to your marks!

As you probably have noticed, this slight modification of the assignment changes everything. Observe, for instance, that none of the problems in the preceding list meets this constraint. We assume that most of our readers experienced a jump in difficulty between the first and second challenges. Whereas inventing a word problem involving division is a piece of cake for nearly everyone, inventing a division problem where the answer is bigger than the starting number is generally not easy at all. It requires a bit of mind-stretching, and for many people it simply is beyond their reach. After all, how can dividing something possibly result in something that’s larger?

The reactions of typical university students to the second challenge are quite diverse. Some students are categorically negative: the challenge is simply impossible. For them, division is by definition incompatible with the idea of making larger. Therefore, instead of inventing a word problem meeting the requirement, they explain why the task makes no sense:

- “Can’t be done. Division always makes things get smaller.”
- “When you have a certain value at the outset and you divide it up, you necessarily have less at the end, so it’s not possible.”
- “Division means sharing, and with equal-sized shares. So each person gets less than what there was at the start. Therefore, it’s impossible to invent a division problem where someone winds up with more than there was at the beginning.”
- “Impossible, because dividing means cutting something up into pieces. To get more, you have to multiply, not divide!”
- “No way, because whenever you divide something, you always reduce it!”

Some other students acknowledge that division problems can indeed have the requested property, because from school they recall the fact that dividing by a number between 0 and 1 has this effect. However, they are convinced that this kind of formal mathematical operation doesn’t correspond to any situation in the real world, and so they assert that there can be no word problem that meets the requirement. At least they can’t think of any. Here are some comments along these lines:
• “I could say ‘10/0.5’, which gives 20, but that’s just a calculation. You can’t make up a corresponding word problem, because in the real world you always divide by 2, 3, 4, and so on. That is, you always divide by numbers bigger than 1.”
• “Yes, it’s possible — for instance, ‘5/0.2’ — but I can’t think of any actual situation that this formula would describe.”
• “Any time you divide by a quantity less than 1 you get a larger answer, but I can’t think of any real situation where it works like that.”
• “When you divide something by one-half, you get more, sure — but the thing is, it’s not possible to divide anything by one-half!”

Then there are some students who invent various problems that seem to them to work, but they cheat in one way or another, because the problems they give don’t match the assignment. For example:

• “Rachel has 20 bottles of wine. She sells half of them at 8 dollars apiece. How much money does she get?”
• “Eric had 8 marbles. In a game, he won half again as many. How many marbles did he wind up with?”

Despite all these protests, it is perfectly possible to devise a division word-problem whose answer is larger than the starting number. Some people find good examples:

• “How many half-pound hamburgers can I make with 4 pounds of meat?”
• “If I have 3 days to prepare for an exam, and it takes me 1/5 a day to read a book, how many books can I read before my exam?”
• “I have 10 dollars, and a chocolate mint costs a quarter (of a dollar). How many chocolate mints can I buy?”
• “How many scarves can I make out of a 3-yard roll of cloth if each scarf requires 3/8 of a yard?”

It turns out, however, that to come up with a problem such as these last four is quite hard. Among 100 undergraduate students, roughly 25 came up with a problem of this type, while the other 75 couldn’t do so, and were split into roughly equal-sized groups associated with the three types of failures quoted above. And so we see that an arithmetical operation that in theory should have been completely mastered in elementary school still gives a great deal of trouble to adults, even university students. Could it be that division problems lie so far back in their past that they’ve forgotten what they once knew about division? Well, no, because the same challenge was set to 250 seventh-graders, all of whom had been studying division for the previous three years, and so for them this kind of challenge was very fresh in their minds (indeed, they had studied problems involving divisors smaller than 1 for at least one full year), and yet
it turned out that over three-fourths of them said that it’s impossible to invent a situation where division gives a larger answer, and of the 250, only one single student invented a word problem that correctly met the challenge.

**Why is it So Hard to Dream up Such Problems?**

It’s a common belief that when situations are concrete, people think more clearly, but this challenge shows that concreteness is no guarantee of clear thinking. The kinds of problems invented by university students in both parts of our little test featured essentially the same kinds of everyday items (cakes, candies, glasses of water, books, scarves, and so forth), and they were set in the same kinds of environments (kitchens, schools, trips, shopping, and so forth). What, then, is the nature of the conceptual gulf between solving the first challenge, which virtually everyone was able to do, and solving the second challenge, which so few people could do?

The explanation is that the two challenges belong to two quite different categories of problems. They do not rest on the same naïve analogy. To be specific, the problems dreamt up in response to the first challenge, which didn’t ask for a larger answer than the initial value, were all problems involving sharing. The examples we quoted above were selected in order to give readers some variety, but in truth, two-thirds of the problems invented were extremely routine, always involving sharing the same kinds of things — relatively uniform everyday objects — among the same kinds of recipients — children, siblings, or friends. From the examples cited, it’s obvious that the division word-problems that people spontaneously come up with nearly always involve the concept of sharing, and more specifically, the splitting-up of a certain quantity into a number of precisely equal shares. The most typical case involves countable items (candies, apples, marbles) shared among people, and the word “sharing” often shows up explicitly in the problem’s statement. Nonetheless, there are more abstract kinds of sharing that show up in a few of the word problems suggested.

In such cases, one has to imagine a more abstract manner of sharing than merely distributing a given set of objects to a given set of people. It might still involve the distribution of entities, but not to human recipients — say, the sorting of cookies into bags, or the arrangement of chairs into rows. It can also involve non-countable substances, such as flour, water, sugar, or land, which get split up into several equal-sized portions. Here there is no sharing in the marked or narrow sense of the term — that is, a counting-out of items, similar to dealing cards out to players in a card game — but there is still sharing in a more general or unmarked sense of the term, in which a whole is divided, through some process of measurement, into smaller chunks. But in any case, none of the responses given by students to the first challenge, whether they involved the marked or the unmarked sense of the concept of sharing, was a division whose result was larger than the initial quantity. And this is no surprise, because the nature of sharing is that it makes something smaller. Sharing involves breaking an entity into smaller parts, with each recipient necessarily receiving less than the whole that was there to start with. A part cannot be larger than the whole from which it came.
By contrast, in word problems that successfully meet the second challenge, a different naïve analogy operates behind the scenes — that of measuring something. (In mathematics education, such problems are said to involve “quotative division”.) Division problems of this type can always be cast in the form, “How many times does $b$ fit into $a$?” This is a measuring situation, in the sense that $b$ is being treated as a measuring-rod with which $a$’s size is being measured. If the size of $b$ is between 0 and 1, then there will be more $b$’s in $a$ than the size of $a$, which means that the result is bigger than the initial size. For example, the calculation $5/0.25$ can be phrased: “How many times does $1/4$ go into $5$?” The answer, 20, is of course larger than 5. What all this shows is that if a division problem is of the sharing sort, then its answer can’t be larger than the starting value, but if it is of the measuring sort, then its answer can be larger.

It turns out that from a historical and scholarly point of view, measuring is a more fundamental way of looking at division than sharing. The definition of division given by Bezout in his 1821 treatise is quite explicit: “To divide one number by another means, in general, to find out how many times the first number contains the second.” Indeed, the etymology of the terms involved in division reflects the view of division as a measuring process. As readers will recall from elementary school, the result of a division is called its quotient. (As Bezout explained it: “The number to be divided is the dividend; the number by which one is dividing is the divisor; and the number that tells how many times the dividend contains the divisor is the quotient.”) The English word “quotient” stems from the Latin word “quotiens”, which is a variant of “quoties”, meaning “how many”, and which derives from “quot”, a word that refers to the counting of objects. In sum, today’s terminology echoes the conception of division as measurement, since “quotient” means “how many times”.

Bezout is aware that seeing division as measurement is not the only possible point of view, but he wants his readers to act as if it were: “One’s goal in doing a division is not always to find out how many times one number is contained in another number; however, one should always carry out the operation as if this were indeed one’s goal.” This shows that the view of division as being primarily a kind of sharing did not come from mathematicians, for they tend to favor the view of division as measurement or counting. To the contrary, the origins of the naïve analogy of division as sharing lie outside of mathematics. As we mentioned earlier, dictionaries tend to define “to divide” in its everyday sense along the following lines: “to separate into parts; split up; sever; to separate into groups; classify; (Math) to separate into equal parts by a divisor” (this taken again from the 1988 edition of Webster’s New World Dictionary).

Is Division Mentally Inseparable from Sharing?

The experimental results we’ve just described show that for most people, division is understood through the naïve analogy of sharing; after all, most people find the first challenge very simple and invent word problems that involve sharing, while the second challenge, which is easily handled if one simply uses the analogy of measuring, is much harder for most people. Although children spend years learning about division in
school and are thus presumed to have mastered this basic operation by the end of middle school (and adults are assumed to know division yet better), it turns out that people of all ages have trouble thinking of division other than through the naïve analogy that equates it with sharing.

Most people use the term “division” not to describe a concept that they learned in school, but to describe a category of situations that was part of their lives before they started school — sharing. When sharing comes up in a mathematical context, they have learned from school to use the term “division” instead. In other words, most people think that “division” is just a technical term to denote the concept of sharing, especially when a calculation is called for, and that’s all there is to it. When one is in math class, sharing has a fancier name, just as in certain arenas of life people use various special terms to designate familiar concepts, even though such terms don’t lend any particular insight. Thus one learns that when one is at the opera, it’s better to say “aria” than “song”, and likewise, when one has truck with wine connoisseurs, one soon gets used to hearing about the “bouquet” rather than the “smell” of the wine; one also gets used to the fact that one’s doctor will tend to speak of “apnea” rather than of “having trouble breathing”, or of “hypertension” instead of “high blood pressure”.

To summarize, although it is tempting to think that schooling teaches people the full-blown concept of division, thus allowing them to throw away the naïve analogy of sharing like a no-longer-needed crutch, the truth is that the crutch remains the central way of understanding division — it merely disguises itself by donning the more impressive-sounding mathematical label of “division”.

**Mental Simulation in the Driver’s Seat**

To solve either of the following two word problems is a challenge as easy as they come:

Paul had 27 marbles. Then during recess, he won some, and now he has 31. How many marbles did he win?

Paul lost 27 of his 31 marbles during recess. How many does he have left now?

Both of these problems are solved by carrying out exactly the same operation — namely, subtracting 27 from 31. At first glance, they thus seem identical in terms of what is going on mentally when we solve them, but let’s set aside the formal operation by which we solved them; instead, let’s try to visualize these situations in our mind’s eye — that is, we’ll try to mentally simulate each of them. What happens?

In the first case, it’s easy to imitate what happened by counting on one’s fingers or in one’s head. Paul’s marble count moved up from 27 to 28 (“1”), then to 29 (“2”), then to 30 (“3”), and finally it reached 31 (“4”). The solution takes four simple steps.

The second case, however, is very different. This time, starting from 31, one has to move downwards 27 steps: first to 30 (“1”), then to 29 (“2”), then to 28 (“3”), then to 27 (“4”), then to 26 (“5”), … , and after a long time one will finally hit 4 (“27”).
We thus see that these two word problems, although they’re both solved by the same formal operation \(31 - 27\), are not imagined or mentally simulated in the same fashion at all. One process involves just four easy counting steps, while the other takes 27 steps, which, to make matters worse, involve counting backwards.

This contrast should recall a similar one from earlier in the chapter — namely, that of the teen-aged street vendors in Brazil. As we saw then, the product of 50 and 3 can be mentally simulated either as \(50 + 50 + 50\) or as \(3 + 3 + 3 + \ldots + 3 + 3 + 3\), depending on how the problem was stated; here, likewise, the subtraction \(31 - 27\) can correspond to two very different mental simulations, one very short and one very long.

Let’s now compare the following four word problems:

1. If we break a stack of 200 photos into piles of height 50, how many piles do we get?
2. If we break a stack of 200 photos into 50 piles, how many photos are in each pile?
3. If we break a stack of 200 photos into 4 piles, how many photos are in each pile?
4. If we break a stack of 200 photos into piles of height 4, how many piles do we get?

The first two are solved by the division “200/50”, and the last two by “200/4”. That is quite obvious. But are some of these problems easier or harder than others? Perhaps it seems as if we are asking if the division “200/50” is easier (or harder) than the division “200/4”. If so, the first two problems would be easier (or harder) than the last two. But things are trickier than that, as a bit of mental simulation will show.

Let’s try to envision the first situation: *If we break a stack of 200 photos into piles of height 50, how many piles do we get?* We first imagine a tall stack of 200 photos; now we want to break it into smaller stacks of height 50. In order to find the answer by simulation (as opposed to doing it by formal division), we take the number 50 and add it to itself until we get 200. 50 plus 50 makes 100, and then another 50 makes 150, one last 50 to make 200. Four 50’s altogether — that’s our answer.

Now let’s try to imagine the second situation: *If we break a stack of 200 photos into 50 piles, how many photos are in each pile?* Formally speaking, this problem also involves the division “200/50”. We had a stack of 200 photos and we divided it up into 50 smaller ones. We need to add up some number 50 times, but we don’t know which number. In fact, not only do we have to do 50 additions to find the right answer, but we may have to do it a bunch of different times, guessing about what to add to itself! 2 + 2+ 2…? After much toil, we get 100, and see that “2” was wrong. 3 + 3 + 3…? Again a lot of toil to wind up with the wrong answer. 4 + 4 + 4…? Well, this time, if we add right, we’ll get 200. But solving this problem is not a piece of cake.

So although the first two problems are both solvable by the same division (200/50), from the point of view of seeing what’s going on in one’s mind’s eye, the first is much easier than the second.

Now let’s look at the third one: *If we break a stack of 200 photos into 4 piles, how many photos are in each pile?* Here we break our tall stack into four shorter stacks. How many photos in each short stack? This is very similar to the second situation, where we want
to do a repeated addition in order to reach 200, but this time we only need to add the mystery number to itself four times, instead of 50 times. On the other hand, we have to guess at the mystery number’s identity. But if we’re clever, we may hit on “50” without too much trouble. For instance, if we recall from everyday life that 50 + 50 = 100, then we can quickly figure out that 50 + 50 + 50 + 50 = 200, and we’re done.

Finally, the fourth problem: If we break a stack of 200 photos into piles of height 4, how many piles do we get? We know we’re dealing with repeated addition of the number 4, but the question is: how many additions? We know that 4 + 4 + 4 + … + 4 + 4 + 4 = 200, but the mystery is how many copies of “4” there are in this sum. This is tough, because we’ll need to keep track of two things in our head at once — firstly, how many copies of “4” we have added up so far, and secondly, what the running tab is. We thus see that mentally simulating the third situation is far easier than mentally simulating this one.

To recapitulate, it turns out problems #1 and #3 are fairly easy to simulate in one’s mind’s eye, while problems #2 and #4 are challenging. If we go back to our earlier distinction between division as sharing and division as measuring, we see that two of these problems are easy to simulate mentally: there is an easy sharing problem (#3, where 200 photos are shared among 4 piles) and there is an easy measuring problem (#1, in which a 200-photo stack is measured using big stacks of height 50). There are also two problems that are very hard to simulate mentally: there is a difficult sharing problem (#2, where 200 photos are shared among 50 piles), and there is a difficult measuring problem (#4, where a 200-photo stack is measured using small stacks of height 4).

Let’s rephrase this in another way. One problem involving the division “200/4” is easy to simulate mentally (#3: 200 photos shared among 4 piles), while another problem that involves exactly the same division is hard to simulate mentally (#4: “Measure a 200-photo stack using piles of size 4”). Likewise, one problem involving the division “200/50” is easy to simulate mentally (#1: “Measure a 200-photo stack using piles of size 50”), while another problem involving exactly the same division is hard to simulate mentally (#2: 200 photos shared among 50 piles).

Here is a short summary of what can be said about these four word problems, all of which are very similar in form (each involves a division whereby a pile of photos is broken into smaller piles), yet differ greatly in how one conceives of them (some involve sharing, some involve measuring) and also in terms of their difficulty for a solver who is mentally simulating them (some are easy, some are hard):

1. If we break a stack of 200 photos into piles of height 50, how many piles do we get?
   [200/50; easy measuring problem]
2. If we break a stack of 200 photos into 50 piles, how many photos are in each pile?
   [200/50; difficult sharing problem]
3. If we break a stack of 200 photos into 4 piles, how many photos are in each pile?
   [200/4; easy sharing problem]
4. If we break a stack of 200 photos into piles of height 4, how many piles do we get?
   [200/4; difficult measuring problem]
Let’s presume that when subjects solve such a problem, they do so by making a mental simulation rather than by instantly carrying out an arithmetical calculation. If that’s the case, then the first and third problems should both be easy, while the second and fourth should be difficult; moreover, knowing which arithmetical calculation is involved (200/4 or 200/50) does not tell us how hard the problem is.

This way of looking at word problems stands in marked contrast to the traditional view, which takes the formal arithmetical operation needed to solve a problem as a gauge of the problem’s difficulty. In place of this, the new perspective highlights the spontaneous way in which the situation is framed — that is, the analogy that allows one to solve the problem in a very direct way, namely by counting in one’s head. This point of view shows why different word problems, even if they are equally down-to-earth, and even if they all involve exactly the same formal arithmetical operation, can nonetheless have extremely different levels of difficulty. The surprising findings that we described earlier concerning Brazilian teen-aged street vendors tackling multiplication problems now seem to apply more generally, both to other kinds of word problems and to other groups of people. The key variable is seen to be the simplicity of the mental simulation that will yield the correct solution.

Rémi Brissiaud, a developmental psychologist, has done pioneering research into these ideas in the context of learning to do arithmetic, and he has written very innovative and efficient new mathematics textbooks inspired by his discoveries. In collaboration with him, we have studied how seven-year-olds who are just beginning to learn the basic arithmetical operations tackle different kinds of word problems. Our results show a clear distinction between problems that tend to be solved by mental simulation and ones that tend to make children resort to formal arithmetical operations. In cases where the mental simulation is not unwieldy, it is always preferred. This principle is illustrated by the following examples:

Paul has 10 boxes containing 4 cookies apiece. How many cookies does he have in all?

Paul has 4 boxes containing 10 cookies apiece. How many cookies does he have in all?

The first problem, if solved by simulation and not multiplication, requires adding 4 to itself 10 times. In such a simulation, children might imagine Paul at the grocery store, taking boxes one by one off the shelf and placing them in his shopping cart. Thus it would go as follows: “Box #1 (4 cookies); #2 (8); #3 (12), … , #9 (36), #10 (40)”.

This process requires one to count and to keep a running tab at the same time, and for children just learning how to add, repeatedly adding up all these 4’s is far from easy. Mental simulation is hard here. By contrast, the second problem is solved by mental simulation quite handily. It takes just four additions, and what’s more, they’re all easy; in fact, each sum along the way echoes the counting number just preceding it, as follows: “Box #1 (10 cookies); #2 (20); #3 (30); #4 (40)”.

Our experiments have shown that children are much better at solving word problems in which mental simulation comes easily than they are at solving problems in which it does not. Not only does this hold before the relevant formal arithmetical
operation has been taught to them (not too big a surprise!), but it also holds after it has been taught (this, by contrast, is quite surprising). We observed that even two years after the relevant arithmetical operation has been taught, if mental simulation provides a short solution path, the problem is solved much more easily than via formal calculation.

Some of our findings fly in the face of received ideas about the relative difficulty of arithmetical operations. For instance, subtraction is usually considered to be an easier operation than division, and is thus taught in schools already roughly at age 6, whereas all mention of division is put off for another two years or so. And yet our experiments have shown that children who have supposedly mastered subtraction but have heard nary a word about division manage to solve certain division problems (those that can be done via mental simulation) better than they can solve subtraction problems in which mental simulation is inefficient. Here is an example of what we are talking about:

Jill passes out 40 cookies to her 4 children. How many cookies does each child get?

This division problem is much more easily solved by children at the above-described stage than the following subtraction problem:

Paul has 31 marbles. He gives 27 to his friend Peter. How many does he have left?

This is not simply due to the fact that 40 breaks easily into 10 + 10 + 10 + 10, because the problem “Jill has 40 cookies and wants to make little packets of 4 each. How many packets will she make?” turns out to be far harder than the one given above, involving 40 cookies given out to 4 children, although both have the same answer (namely, 10).

Our experimental findings show that whenever it’s possible, children opt for using analogies to real-world situations rather than making formal arithmetical calculations. If a word problem can be conceived of in such a way that formal calculations can be bypassed, then simulation is the pathway that children tend to follow. The formal technique will be wheeled out only when there is no alternative — that is, in situations where mental simulation would be inefficient, either because it would require too many steps (adding up ten 4’s) or because it would require the use of arithmetical facts about which the child is still a bit shaky (e.g., “4 + 16 = 20”).

The Influence of Language on Naïve Analogies

Does sharing’s dominance over measurement as a naïve analogy for division mean that the former is a simpler concept than the latter? Is sharing such a natural and familiar idea that it automatically and irrepressibly jumps to mind as a ready-made analogue for division? And is measurement such a rare and unfamiliar idea that it is unlikely to be used as an analogue for division? Is this why sharing enjoys the lion’s share of mental imagery for division?
The answer is no. The predominance of sharing as the naïve analogy for division is not due to intrinsic simplicity, but merely to an accident of language. Division is unconsciously associated with sharing in the minds of most speakers of English because the English word “division” has both a mathematical meaning and an everyday meaning, and connotations of the everyday meaning inevitably spill over into the technical meaning; as a consequence, the naïve analogy of division as sharing overwhelms that of division as measuring.

Suppose there were an arcane mathematical notion called “surgery” (indeed, it exists). If you were told that surgery sometimes involves smoothly tying things together and other times involves tearing things asunder, it seems likely that, thanks to your prior familiarity with medical surgery, the naïve analogy that you would unconsciously exploit in trying to make sense of the notion would tilt more towards tying together than tearing apart. A similar story can be told about division and sharing. Suppose that hundreds of years ago, the English word assigned to the mathematical concept of division had not been “division” but “measurement”. Had that been the case, then children today, on first hearing about the arithmetical notion called “measurement”, would tend to create a very different primary naïve analogy for it. And the idea of sometimes getting a larger answer than what they started with (that is, \( a/b \) being sometimes greater than \( a \)) wouldn’t strike them as strange or confusing in the least.

In sum, the predominance of the naïve analogy “division is sharing” doesn’t imply that envisioning a measurement (“How many B’s will fit inside A?”) is cognitively more demanding than envisioning an act of sharing (“If I cut A into B parts, how big will each part be?”). Indeed, quite to the contrary, we’ve seen that a measurement problem such as “How many 50’s are there in 200?” is much easier than a sharing problem that involves exactly the same numerical values: “Dole out 200 candies to 50 kids!”

What we learn from this example and similar ones is that the prevalence of sharing as the naïve analogy for division is not because sharing is easier to imagine than measuring; it is because there is an unconscious bleed-through of the everyday meaning of “to divide” into the technical term “to divide”, and this semantic contamination gives a big head start to sharing as the source of the naïve analogy, even if there are many cases where measurement would be a more apt analogy. In other words, it’s easy to solve “How many 50’s are there in 150?” in one’s head (“50 and 50 and then 50 again — that makes 150, so the answer is 3”), but to realize that one has just solved a division problem is not easy at all, because the usual feeling of carrying out a division is pervaded by the everyday sense of that word, which has no connection with the idea of measuring anything.

**What Schooling Leaves Untouched in Our Minds**

Does what we learn in school profoundly affect how we see situations? Does school teach us to think “formally” about situations? By this, we mean acquiring the ability to zoom straight to the abstract core of a situation, not deflected by its concrete details. We all do this in some aspects of daily life — that is, we routinely ignore many aspects
Naïve Analogies

of certain situations, though we are fully aware of them intellectually. Thus, we know
but we forget that our closet door was once part of a tree, that Adolf Hitler was once a
baby, that the meat on our table was not long ago inside an animal grazing in a field,
and so forth. Even if we accept the truth of these facts, we systematically ignore them, so
that it’s fair to say that we simply don’t see an ex-animal in the steak, nor an ex-tree in
the door, nor an ex-baby in photos of the Führer. This is not stupidity but intelligence.

In the same way, we pay no attention to all sorts of properties of the objects that
surround us. Who would think of using a painting hanging on the wall as a tray on
which to carry the dirty dishes into the kitchen, or as a bulletin board onto which we
could post a bunch of family photos, or as a throw rug that might decorate our floor?
And yet in theory, and in case of extreme need, any of these uses might come to mind,
perhaps even seeming eminently reasonable. Only when one is extremely angry or
frightened does it ever occur to one that a candle, a plate, a vase, a glass, a statuette, a
chair, and a mirror are all potential weapons, and this “forgetfulness” is as it should be.

Categorization involves taking a certain point of view, and once one has chosen a
category for something in one’s environment, that act tends to suppress the perception
of all sorts of properties that are irrelevant to the chosen category. Who ever wonders if
the hamburger in their bun came from a male or a female cow? And who would ever
care if it’s a left or a right shoe when one is so starved that one is desperate to eat it? In
short, we are constantly abstracting and thus constantly ignoring thousands of
potentially observable facets of things and situations. Once again, this is not stupidity
but intelligence. Does school teach us to use this kind of “intelligent forgetting” or
“intelligent ignoring” more systematically, especially in mathematics?

When most people are given a mathematical word problem — even a very simple
one — they have great trouble ignoring some of its irrelevant aspects. Instead of
treating such a problem in a formal manner, they tend to be influenced by some of its
salient concrete features. Even if someone eventually discovers the abstract
mathematical structure in a word problem, that recognition never fully overrides the
person’s more spontaneous initial view of the situation; various concrete aspects get
blurred in with more abstract ones. Our ability to perceive mathematical situations
formally — that is, in such a way that our thinking is not contaminated by some of their
irrelevant, surface-level aspects — is very limited.

We will illustrate this using one of the easiest possible cases for an adult —
multiplication, which, as we stated earlier, is one of the first operations taught in school,
and which is a concept that few people feel they have anything more to learn about.
To make it even simpler, we’ll look only at cases of multiplication involving positive
integers. We compare three situations:

I go into a store where every item costs 4 dollars, and I buy 3 pens. How much money
do I spend?

For each of my 4 children, I buy 3 pens. How many pens do I walk out with?

For each of my 4 children, I buy one blue pen, one green pen, and one red pen. How
many pens do I walk out with?
Seeing things purely formally would involve instantaneously filtering out the store, the narrator, the pens, the colors, the children, and the money, and jumping directly to the bare-bones idea that “These problems all involve multiplication of 4 and 3”, and then simply carrying out that operation. Casting the problems this way should involve no asymmetry between the two factors, because if these problems are seen on a purely abstract level, they are all just multiplication problems. And indeed, given the above problems, few us would wonder, “Is this a ‘4 x 3’ situation or is it a ‘3 x 4’ situation?” We would instead simply tend to think, “Multiply 3 and 4 together.” Or at least this is one’s first introspective impression when, as an adult, one tackles these problems. But do we really perceive, in our mind’s eye, nothing but a “Multiply 3 and 4” situation in all three cases? Do we really see and do exactly the same thing when we solve these three problems? Let’s take a closer look at them.

In the first problem, where every item costs 4 dollars, I buy one pen for $4, then another, and then another. Thus in the end, I spend $4 + $4 + $4 = $3 x $4 = $12 dollars.

In the second problem, where I’m buying pens for my children, I buy 3 pens for each child. Thus I buy 3 pens for the first child, 3 for the second, 3 for the third, and 3 for the fourth child. Altogether, then, I buy $3 + 3 + 3 + 3 = 4 x 3 = 12 pens.

In the third situation, there are two possible perspectives. From the first perspective, we focus on what each child will get — that is, we see things much as in the second problem. Each child gets 3 pens, and that makes 4 cases of “getting 3”, which means $3 + 3 + 3 + 3 = 4 x 3 = 12 pens.

From the second perspective, I focus on colors rather than on children. I’m buying 4 blue pens (one for each child), 4 green pens, and 4 red pens. Therefore, since there are 3 colors concerned, I am buying $4 + 4 + 4 = 3 x 4 = 12 pens.

If a person solving one of these word problems immediately perceived the abstract idea that it involved multiplication, then the way in which the problem happened to be concretely embedded in the world should have no effect on the order of the factors. By contrast, if these word problems are subliminally perceived through the filter of the naïve analogy repeated addition, as opposed to being perceived formally as multiplications, then the naïve analogy should guide the way the problems are solved, and the solutions that people find should be influenced by the preceding considerations.

To test this hypothesis, we asked older elementary-school students and also university students to solve these problems without using multiplication. It turned out that nearly everyone, children and adults alike, used repeated addition, which is no surprise, and moreover that the additions that were chosen depended crucially on how the problems were stated, which may be more of a surprise.

Thus roughly 90% of the subjects used the addition “$4 + 4 + 4” for the first problem, and roughly the same percentage solved the second problem using the addition “$3 + 3 + 3 + 3”. For the third problem, involving colored pens, roughly 50% of the subjects went for “$3 + 3 + 3 + 3” (thus seeing it in terms of children), roughly 40% of them went for “$4 + 4 + 4” (thus seeing it in terms of colors), and the remaining 10% didn’t group anything together explicitly; that is, the repeated addition they saw was “$1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1”. Of course those subjects still
made unconscious groupings — either “(1+1+1) + (1+1+1) + (1+1+1)” (each group representing a child) or else “(1+1+1+1) + (1+1+1+1)” (each group representing a color).

This finding clearly supports our prediction, which is that given a word problem, people will try to solve it not just by perceiving its formal structure, but also by doing their best to find, based on the way the problem is worded, an analogy to repeated addition. In summary, to solve these word problems, people tend to mentally simulate the situations described in them.

If, on the other hand, the situation had been spontaneously perceived “formally”, meaning that the abstract concept multiplication had been instantly evoked in subjects’ minds, then (since using multiplication had been explicitly banned) the repeated addition “4 + 4 + 4” would have been suggested by all the subjects for all the problems, because that sum involves the least computation, and there would have been no good reason to opt for the repeated addition “3 + 3 + 3 + 3”, since it uses more summands.

To summarize, we have shown that even in a bare-bones mathematical situation, people are very seldom able to ignore all of its superficial, concrete aspects and to home in on just its abstract formal structure. For better or for worse, people are influenced by how situations are concretely described, by their familiarity with similar situations, and by the naïve analogies that these situations evoke naturally.

Sometimes Situations Do Our Thinking for Us

Here is a problem to solve, which contains no hidden trick:

Lawrence buys an art kit for $7, and also a binder. He pays $15 altogether. John buys a binder and also a T-square. He pays $3 less than Lawrence did. How much does the T-square cost?

Of course you’ve already gotten the correct answer, but that’s not the main point here. We would instead ask you to think carefully about this problem and try to find the most streamlined, efficient way to solve it, showing exactly how the shortest, simplest solution would work, step by step.

Usually, people suggest a solution that involves three calculations, as follows:

- Price of the binder: $15 – $7 = $8.
- Price paid by John: $15 – $3 = $12.
- Price of the T-square: $12 – $8 = $4.

This a perfectly correct way to solve it. But now, how about tackling the following problem, again looking for the shortest route to the solution:

Laurel took ballet lessons for 7 years and stopped at age 15. Joan started at the same age as Laurel but stopped 3 years earlier. How long did Joan take ballet?
Once again we are interested in the most economical way of solving this problem, and in seeing what exactly are the steps that must be taken. We thus ask you to indulge us once again in trying to find and spell out the minimal pathway to the solution.

One idea is to carry out essentially the same steps as in the preceding problem:

- Age at which Laurel (and hence also Joan) started ballet: $15 - 7 = 8$ years.
- Age at which Joan quit: $15 - 3 = 12$ years.
- Total time that Joan took ballet: $12 - 8 = 4$ years.

Although this pathway to the solution is totally correct, another pathway might have come to your mind. If Laurel and Joan started taking lessons at the same age and Joan stopped 3 years earlier than Laurel did, then Joan took ballet 3 years less than Laurel did, so she took lessons for $7 - 3 = 4$ years (we were explicitly told that Laurel took lessons for 7 years). This pathway involves just one arithmetical operation!

The existence of this alternate route to the solution of the second problem suggests to most people that the two word problems differ fundamentally from each other, since the ballet-lesson one can be solved in just one step, whereas no similar shortcut exists for the school-supply problem. Is this really true, though?

What would it mean to use the subtraction “7 – 3” in the context of the first problem? To be sure, it gives the right answer — 4 dollars — but does it mean anything? Most people who are given this problem tend to think it doesn’t, or that if it is a meaningful thing to do, then it would take a long time to figure out why, and it’s not worth it. And yet, the situation can be described as follows: “Both Lawrence and John bought a binder plus some other item”; this view leads one to a solution using just one single operation. One of them paid 3 dollars less than the other one paid. Therefore, the difference between what the two boys shelled out is due totally to the other item. Hence the price of John’s other item (the T-square) must be 3 dollars less than the price of Lawrence’s other item (the art kit): $7 - 3 = 4$.

What’s remarkable, when one compares the solutions of these two problems, is that fewer than 5% of elementary-school children and fewer than 5% of adults (in fact, of highly educated adults — namely, university students and schoolteachers) find the direct one-subtraction solution to the school-supply problem, whereas roughly 50% of the children and also 50% of the adults spontaneously find the one-subtraction solution to the ballet-lesson problem. And so, although all that’s required here is to carry out very trivial subtraction operations in extremely concrete situations, nonetheless the angle of attack that yields a one-step solution is almost never found in the first context while it is very often found in the second context.

What Does It All Mean?

It’s no accident that the very same one-operation method can be used to solve both problems, as both involve situations that could be handled by using the following formal rule, which has the feel of a theorem one might find in a set-theory textbook:
If two sets overlap, then the difference between their sizes equals the difference in the sizes of their non-overlapping parts.

If we apply this rule to the school-supplies problem, it tells us that the difference between what Lawrence paid and what John paid must be equal to the difference between the non-overlapping parts of their purchases — that is, the difference between Lawrence’s art kit and John’s T-square. If we apply the rule to the ballet lessons, it tells us that (since Laurel and Joan had equally long periods with no lessons), the difference between their ballet-quitting ages is equal to the difference between their ballet-lesson periods. If this formal rule were learned and fully absorbed by everyone, then we would expect that both problems would be always solved by the single-subtraction method. As we have seen, though, nothing of the sort happens. How come?

Simply because the formal rule is not part of most people’s mental repertoire. Even people who discover the one-operation method for the second problem are unlikely to be aware of any such rule. Rather, they just allow the problem itself to direct their thoughts. If they come up with a one-step solution, it’s because that is what they are naturally led to. Each situation is defined, in a person’s mind, by the categories it effortlessly evokes, and that perception, rather than the application of any formal rule, is what guides the person’s thinking.

The ballet-lesson problem is thus perceived as follows. If two events start at the same moment, and one of them lasts $N$ time-units less than the other, then it will end $N$ time-units before the other one does. This is so patently obvious to us all that the sentence tends to sound like a mere tautology, a vacuous triviality. Let us nonetheless restate it slightly differently: if two events start simultaneously, then the difference between their durations equals the interval between their cutoff moments. People’s perception of time is so deeply anchored and they so intuitively understand this basic principle — taking less time means ending earlier — that they often recall the statement of the ballet-lesson problem in a distorted fashion.

To be specific, if students who have read the statement “Joan started ballet lessons at the same time as Laurel but took them for three years less” are asked to write it down by memory, they often do not reproduce it correctly, writing instead: “Joan started ballet lessons at the same time as Laurel but quit three years earlier.” Their deduction is so deeply fused with their perception of the situation that they don’t see it as such; they are unaware of having transformed the sentence in committing it to memory.

Indeed, the transformation of the problem from the initial phrasing (“Joan took three years less”) to the final phrasing (“Joan finished three years earlier”) converts a difference between two lengths of time into a difference between two temporal stopping points. In the isomorphic school-supply problem, this would be comparable to someone converting the three-dollar price difference between Lawrence’s total outlay and John’s total outlay into the difference between just the art kit and the T-square. However, in our experiments we have never run across any subject who read the phrase “John’s total outlay was three dollars less than Lawrence’s” and subsequently wrote it down by memory as “the T-square costs three dollars less than the art kit”.

Naïve Analogies

431
What we see here is that prior knowledge about relations involving time (“three years less” can be converted into “three years earlier” or *vice versa*) allows people to solve the temporal word problem in one single step. In this sense, one could quite reasonably claim that it’s the *situation* that’s “doing the thinking” for the subjects. Finding the one-step solution is not a consequence of having mastered some general set-theoretic rule such as we quoted at this section’s outset.

As is probably quite clear, students who solve the ballet-lessons problem do not rely on its abstract, formal structure (as expressed in the above-stated “theorem”); they have no need at all to evoke the abstractions of set theory in order to solve it. When they carry out their reasoning, they don’t perceive the time intervals involved as *sets* (if they did, each set would contain infinitely many infinitesimal moments!), nor do they perceive the common age at which Laurel and Joan started their lessons as the *overlap* of two sets (their “intersection”, in set-theoretic language), nor do they see the lengths of time that the two girls took lessons as the *non-overlapping* parts of sets, nor do they see the girls’ ages when they stopped their lessons as the *sizes* of two sets. In fact, it would take careful intellectual work to recast this problem in set-theoretical language, because set theory is not the framework in which humans naturally perceive it, and that’s why efficient solving of the problem by a person should not be taken as showing that the formal rule (the “theorem”) was correctly applied. Rather, the act of perceiving the ballet-lesson situation in terms of familiar time-categories does the bulk of the work for the student, and there is no need whatsoever to code the situation into an arcane, abstract technical formalism such as set theory.

The diagrams on the facing page show three ways of conceiving these isomorphic word problems. The diagram at the top shows how the school-supplies problem is imagined by nearly everyone to whom it is given. The typical assumption is that in order to figure out the price of the T-square, you have to subtract the price of the binder from John’s total outlay (which itself is found through a subtraction), and the binder’s price is found by subtracting the price of the art kit from Lawrence’s total outlay — thus three subtractions *in toto* seem necessary. In this diagram, one doesn’t see that the difference between John’s and Lawrence’s total outlays is identical to the difference between the prices of the T-square and the art kit. That key idea, a prerequisite to solving the problem in a single step, is missing, and so three arithmetical operations seem to be needed.

The middle diagram shows the way that many people very naturally envision the problem of the ballet lessons. They begin with the idea that Joan and Laurel took up ballet at the same age, and so the difference between the *lengths of time* they took ballet has to equal the difference between their *ages* when they quit. This idea is screamingly obvious in the diagram, and that explains why the one-step solution is often found for the ballet-lessons problem. (*Let’s not fail to note the frame blend on which this diagram is tacitly based, and which Gilles Fauconnier and Mark Turner would delightedly point out — namely, the fact that we are imagining aligning the two girls’ lives on a single horizontal time axis. Aligning their lives means placing their births at the same spot on the time axis, and as a result of this maneuver, their first lessons will also coincide.*)
Each of the two upper diagrams (the top one, involving three operations, and the middle one, involving just one) was tailored to fit one specific problem, and it is not apparent that they have much in common with each other. Looking at the upper two diagrams alone, one might think that the problems they represent are of very different sorts; in fact, this is what most people claim who try to solve them both. But the bottom diagram not only shows how the two problems can be seen in a single unified manner; it also reveals how a one-step method works just as naturally for the school-supply problem as for the ballet-lesson problem. While the “theorem” enunciated above was abstruse and difficult to grasp, this visual encoding of the two problems in one single picture reveals their isomorphism (i.e., the fact of having the same underlying structure), as well as how they both can be solved with just one arithmetical operation.

These three diagrams show how the way a person visualizes a word problem can either bring out or hide a pathway to the solution. The bottom diagram could be said to be more abstract than the upper two in that it unifies them in a single diagram, but on the other hand, the image of two boxes standing on the same pedestal is very concrete, and as soon as these problems are cast in a form that involves a shared “pedestal” (the binder, in the first problem, and the starting age, in the second one), the
abstruse idea expressed in the “theorem” suddenly becomes crystal-clear, as it has been fleshed out in a concrete manner, using simple, everyday images, such as boxes resting on a shared pedestal. Teaching a young student to see the “pedestal” in these two word problems imparts an elegant insight that is unavailable to most untrained adults.

A key challenge for educators is thus to take into account the way people manage to adroitly sidestep the formal encoding of situations by exploiting the way their familiar categories, built up over years of interactions with the world, work. Although most teachers are quite aware that the way in which a problem is “dressed” can profoundly affect its difficulty, educators have not yet figured out how to make the art of dressing mathematics problems into a powerful teaching tool. To achieve this would be a great advance, but of course doing so constitutes a great challenge as well.

A Naïve Analogy that has Ill Served Psychology

Do you find it hard to see naïve categorizations such as multiplication is repeated addition and division is sharing as constituting analogies? Despite the multiple arguments we’ve mustered and the many situations we’ve dissected with a fine-tooth comb, perhaps a little voice inside you keeps insisting, “I’m sorry, but categorization and analogy-making are just not the same thing. Taking two notions that initially seem very far apart and then building a mental bridge between them because one sees that they have certain abstract qualities in common is a profoundly distinct type of act from seeing something and merely recognizing that it belongs to a familiar category, such as table.”

Why do so many people, perhaps even you, have an inner voice that so strongly resists the thesis that the building of analogies between things is just the same activity as the assigning of things to categories? How come your inner voice hasn’t gradually calmed down and grown silent over the course of reading this book? How come it hasn’t listened to all the reasons that we had hoped would convince you of our thesis?

The answer — a rather ironic one — is that our thesis itself explains why so many people have so much trouble accepting it: namely, the belief that analogy-making and categorization are separate processes springs from none other than a certain naïve analogy about the nature of categories. This naïve analogy, which has the dubious honor of having seriously held back progress in the field of analogy and categorization for a long time, has already been cited in this book. Here it is, once again:

Categories are boxes, and to categorize is to put items into boxes.

This is the everyday, down-home view of categorization. Let’s think about it a little bit.

If categories really were boxes and if there really were a reliable, precise mechanism for assigning things to their boxes, then it would make eminent sense to distinguish between two types of mental process. First would be categorization — a rigorous, exact algorithm reliably placing mental items in their proper boxes; second would be analogy-making — a subjective and fallible technique for dreaming up fanciful, unreliable bridges between mental items that do not enjoy a contents-to-container relationship.
However, as research in psychology has shown, and as we have stated throughout this book, the vision *categorization = placing things in their natural boxes* is highly misleading, for categorization is every bit as subjective, blurry, and uncertain as is analogy-making. A categorization can be outright wrong, can be partially correct, can be profoundly influenced by the knowledge, prior experiences, prejudices, or goals, conscious or unconscious, of the person who makes it, and can depend on the local context or the global culture in which it is made. In addition, categorizations can be just as abstract as analogies, can be nonverbal as well as explicitly verbal, can be shaded, and so forth.

Recent work, in fact, makes this point so obvious that the old view of categories as boxes is now usually called “the classical approach to categories”, because in cognitive science there is scarcely anyone around any longer who still puts stock in it; today the classical approach tends to be looked upon as simply a quaint historical stage in the development of a far more sophisticated theory of categories and categorization.

And yet, the “categories are boxes” naïve analogy is still seductive, and leads us all to fall victim to the nearly irrepressible belief that all objects and situations we encounter have a privileged category to which they belong, and which constitutes their intrinsic identity. Let’s recall the object Mr. Martin purchased; though it subsequently bounced merrily back and forth among such motley categories as *fragile object, dust-gatherer, spider carrier, and home for tadpoles*, it always seemed to remain in truth just one thing — namely, a *glass* — and this label constituted its genuine, intrinsic identity. The naïve unspoken analogy “categories are boxes” implies that each item in the world, just like Mr. Martin’s glass, has a proper box to which it belongs, and that this connection between a thing and its natural box is universally shared in all people’s heads, and finally that this is simply the nature of the world, having nothing to do with thinking or psychology. In this view, not only would category identity exist, but it would be precise and objective.

Over the years, the insidious motto “categories are boxes” has underwritten much scientific research reinforcing the belief that there is a clear-cut distinction between analogy-making and categorization. Indeed, the view of categorization that reigned in cognitive psychology for decades, though expressed in far more sophisticated terms, was essentially indistinguishable from this motto. That view was a restatement of the motto in technical terms borrowed from mathematical logic. It portrayed each mental category as possessing a set of “necessary and sufficient conditions” for membership. An entity belonged to the category *if and only if* it had all those properties. Thus each category-box was thought of as being precisely defined and having rigid, impermeable walls. In this theory of categories, there was no room for degrees of membership, nor for contextual effects on membership in categories. Only in the mid-1970s, when psychologist Eleanor Rosch published her seminal series of articles on categorization, was this erroneous but nearly universally accepted theory at last discredited.

And so, after finally being liberated from the motto, how do today’s researchers view analogy-making and categorization? Has a clear consensus emerged on what, if anything, makes the two different? Well, let’s listen to some of the top authorities. On the one hand, Thomas Spalding and Gregory Murphy write, “Categories let people treat new things as if they were familiar”; on the other, Mary Gick and Keith Holyoak...
state, “Analogy is what allows us to see the novel as familiar”. If there is a distinction between these two characterizations, it eludes us! Or, to cite Catherine Clement and Dedre Gentner, “In an analogy, a familiar domain is used in order to understand a new domain, especially to predict new aspects of this new domain”, whereas for John Anderson, “If one establishes that a given object belongs to a certain category, then one can predict a great deal about the object.” Once more, these descriptions of supposedly different processes make them sound as alike as Tweedledum and Tweedledee.

The great overlap of these experts’ definitions confirms (if confirmation was needed) the tight relationship of analogy-making and categorization, and provides grist for the mill we are defending — namely, that the idea that analogy-making and categorization are separate processes is illusory. Still, someone who wished to play devil’s advocate might argue that by overhauling the definitions of analogy-making and categorization, one might be able to show that they are indeed two processes and not just one. In fact, after Chapter 8 there is a dialogue that carefully explores this possibility, and we hope that that dialogue, in addition to closing our book, will also close the book on this issue.

However, if you, even after having read this far in our book, still feel reluctant to accept our thesis, rest assured that you are in excellent company, for even experts in the fields concerned fall victim all too often to this same naïve vision. Indeed, we — the authors of this book, the most fervent proponents of our thesis — find ourselves from time to time falling into the very trap we’ve tried so hard to warn our readers about! Yes, we too fall occasionally for the tempting illusion that categories are boxes. The following way of putting it would probably have warmed the cockles of Aristotle’s heart:

People make naïve analogies.
Analogy experts are people.

Therefore, analogy experts make naïve analogies.

Specialists of any ilk, be they experts in psychology, mathematics, physics, or any other field, do not belong to a different species from ordinary people; they make naïve analogies not only in their daily lives, but also in their professional lives, even ones that involve the concepts with which they are the most proficient.

In short, we sympathize with those readers who still have some doubts about the identity of categorization and analogy-making. It is, after all, a counterintuitive view, and a little voice inside, prompted by a beguiling naïve analogy, continually whispers, “It’s wrong! It’s wrong!” Nonetheless, we harbor fond hopes that in the remaining pages of our book, we might still get some doubters to swing around to our view.

In any case, speaking of the remaining pages, it is high time that we moved from the often troubling world of naïve analogies to the ever-admirable role of analogies in scientific discovery. And yet in so doing, will we really leave naïve analogies far behind?