are guaranteed to lose money. To verify this fact, observe that the draw has exactly three possible outcomes:

> *Possibility 1*
> Some card other than red-red is drawn, and it lands with a white side up, that is, not-RR and W-up.
>
> *Possibility 2*
> Some card other than red-red is drawn, and it lands with a red side up, that is, not-RR and R-up.
>
> *Possibility 3*
> The red-red card is drawn, and it lands (of course) with a red side up—that is, RR and R-up.

It is impossible to obtain the combination RR and not-R-up.

Now if possibility 1 arises, then you lose bet (a), win bet (b), and neither win nor lose bet (c′); overall, you lose $.10. If possibility 2 arises, then you lose bet (a), lose bet (b), and win bet (c′); overall, you lose $.10. If possibility 3 arises, then you win bet (a), lose bet (b), and lose bet (c′); overall you lose $1.80. Thus, you lose no matter what happens.

To recapitulate, your judgment about the three-card problem led you to assign certain probabilities to statements (a), (b), and (c′) in table 2.1. These probability attributions led you to accept as fair bets (a), (b), (c′) of (5). It turns out, however, that the joint outcome of these bets is bound to be unfavorable. In the terminology of probability theory, a *Dutch book* has been made against you.

Most people (including the present author) are lured into Dutch books on first encountering the three-card problem. Let us now see how to avoid them in principle, and then analyze the psychological factors that lead us into situations of this kind.

2.3 How to Avoid Dutch Books

Not every probability function is open to a Dutch book. Those immune to such traps can be characterized with the aid of probability theory.[6] A few preliminary concepts are necessary.

2.3.1 Logical Truth, Exclusion, and Equivalence

Roughly, a *logical truth* is a statement that is true in every conceivable circumstance—for example:[7]

6. Systematic presentations are available in Resnik (1987), Ross (1988), Skyrms (1986).
7. A fuller discussion of logic is available in Allen and Hand (1992). See also chapter 9 in the present volume.

Either all frogs croak or some frog does not croak.

Two statements are called *logically exclusive* just in case there is no conceivable circumstance in which both are true—for example:

The heaviest poodle weighs more than 80 pounds.
The heaviest poodle weighs less than 60 pounds.

Two statements are *logically equivalent* just in case they have the same truth value in every conceivable circumstance—for example:

Not all philosophers like Mozart.
Some philosopher does not like Mozart.

It is assumed in what follows that if S_1, S_2 are statements in the domain of a given probability function, then so too are certain logical combinations of them, namely:

(6) a. not-S_1

b. S_1 or S_2

c. S_1 and S_2

Statement (6c) is called the *conjunction* of S_1 and S_2.

2.3.2 Coherent Probability Functions

Consider a probability function $\mathbf{P}$ such that for all statements S_1, S_2 the following conditions hold.

(7) a. $\mathbf{P}(S_1) \geq 0$.

b. If S_1 is logically true, then $\mathbf{P}(S_1) = 1$.

c. If S_1 and S_2 are logically exclusive, then $\mathbf{P}(S_1 \text{ or } S_2) = \mathbf{P}(S_1) + \mathbf{P}(S_2)$.

d. If $\mathbf{P}(S_2) \neq 0$, then $\mathbf{P}(S_1 \text{ and } S_2) = \mathbf{P}(S_1 \text{ assuming that } S_2) \times \mathbf{P}(S_2)$.

Conditions (7a–d) may be paraphrased:

a. No probability is negative.
b. The probability of a logical truth is 1.
c. The probability that one of two logically exclusive statements is true equals the sum of their respective probabilities.
d. The probability of the conjunction of two statements equals the probability of the first, assuming the second times the probability of the second.

A probability function that satisfies conditions (7) automatically satisfies a variety of other conditions, in particular these (with respect to any statements S_1, S_2):

(8) a. $\mathbf{P}(\text{not-}S_1) = 1 - \mathbf{P}(S_1)$.

b. If S_1 and S_2 are logically equivalent, then $\mathbf{P}(S_1) = \mathbf{P}(S_2)$.

c. If S_1 logically implies S_2, then $\mathbf{P}(S_1) \leq \mathbf{P}(S_2)$.

d. *Bayes's theorem*: If $\mathbf{P}(S_2) \neq 0$ then $\mathbf{P}(S_1 \text{ assuming that } S_2) =$

$$\frac{\mathbf{P}(S_2 \text{ assuming that } S_1) \times \mathbf{P}(S_1)}{\mathbf{P}(S_2)}$$

e. $\mathbf{P}(S_1 \text{ and } S_2) \leq \mathbf{P}(S_1)$.

The deduction of (8) from (7) is problem 2.2 at the end of this chapter. The conditions in (7) are known as *Kolmogorov's axioms* and they suffice to develop the elementary portion of probability theory. Probability functions that satisfy (7) are called *coherent*, the others *incoherent*.

2.3.3 The Dutch Book Theorem

Coherent probability functions stand in the following relation to Dutch books:

(9) *Dutch Book theorem*
Suppose that individual I is willing to accept any bet that is fair for I (in the sense of section 2.2.2). Then a Dutch book can be made against I if and only if $\mathbf{P}_I$ is not coherent.

For proof of the Dutch Book theorem, see Lehman 1955, Kemeny 1955, and the discussion in Gustason 1994, Resnik 1987, and Skyrms 1986. The theorem shows that coherent judgment offers protection against falling for a Dutch book. Conversely, some Dutch book can be contrived against anyone manifesting an incoherent probability function—and willing to accept apparently fair bets. The latter proviso is nontrivial because not everyone likes to wager. The Dutch Book theorem nonetheless provides striking evidence for the reasonableness of the axioms in (7).

Notice that coherent judgment is no guarantee against foolish bets in general. A person may accept an even-money wager that Luciano Pavarotti will finish first in the next Boston Marathon without violating (7a–d). This bet is not a Dutch book. Coherence is protection only against combinations of bets whose logic guarantees a loss.

Now let us return to the three-card problem. From table 2.1 we see that your probability function yields these judgments:

(10) a. $\mathbf{P}_{you}(RR) = \frac{1}{3}$.

b′. $\mathbf{P}_{you}(\text{R-up}) = \frac{1}{2}$.

d. $\mathbf{P}_{you}(\text{R-up assuming that RR}) = 1$.

Plugging these values into Bayes's theorem (8d), we see that $\mathbf{P}_{you}$ is coherent only if

(11) $\mathbf{P}_{you}(\text{RR, assuming that R-up}) =$

$$\frac{\mathbf{P}_{you}(\text{R-up assuming that RR}) \times \mathbf{P}_{you}(\text{RR})}{\mathbf{P}_{you}(\text{R-up})} = \frac{1 \times (1/3)}{1/2} = \frac{2}{3}.$$

However, (c) in table 2.1 reveals that $\mathbf{P}_{you}(\text{RR assuming that R-up}) = \frac{1}{2}$, not $\frac{2}{3}$ as required by (11). We conclude that $\mathbf{P}_{you}$ is not coherent. Theorem (9) thus implies that you are open to a Dutch book. And this is what we saw in section 2.2.3.

2.4 Incoherence or Momentary Illusion?

Incoherence has such far-reaching consequences that we should be cautious about imputing it to ourselves or others. In particular, answers to the three-card problem seem too slender a basis for global evaluation of a person's probability function. Let us therefore consider the possibility that the responses recorded in table 2.1 result from a fleeting illusion that is not typical of our considered judgment.

2.4.1 Illusions in Other Domains

A perceptual analogy may be helpful. Consider the Müller-Lyer illusion in figure 2.1. Although the two horizontal lines are equal in length, the presence of the arrowheads disturbs our perceptual judgment and favors the impression that the top line is longer than the bottom one. Does this

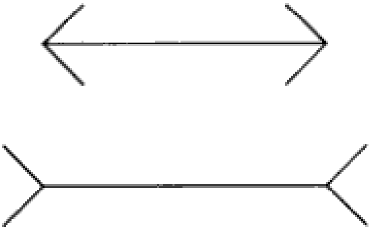

Figure 2.1
The Müller-Lyer illusion.

illusion imply that the human visual system lacks a veridical mechanism for comparing line lengths? The answer must be no, because in the absence of arrowheads we judge the comparative length of parallel lines with great precision.

In the same way, we may qualify as illusory the impression of nongrammaticality engendered by sentences like:

The horse raced past the barn fell.

The sentence is well-formed English, and is perceived as such upon reflection. Its apparent nongrammaticality arises from minor imperfections in human parsing mechanisms.[8] As before, these imperfections ought not to blind us to the existence of mental systems that yield, in usual circumstances, a correct verdict about the grammaticality of sentences.

There are also estimation illusions, such as the following "anchoring" effect discovered by Tversky and Kahneman (1974). These investigators asked people to estimate various quantities stated in percentages. For example, one estimate stated the percentage of African countries in the United Nations. Prior to each estimate, a number between 0 and 100 was determined by spinning a wheel of fortune. Participants were then asked to indicate whether the randomly obtained number was higher or lower than the value of the quantity to be estimated, and finally to estimate the quantity by moving upward or downward from the number given. Despite the evident, arbitrary character of the "anchors," they influenced estimates considerably. For example, the median estimates for the percentage of African countries in the United Nations were 24 and 45 percent for groups that received 10 and 65, respectively, as starting points. The effect of anchoring was not dissipated by offering monetary rewards for accuracy.

A similar phenomenon is reported in Slovic et al. 1980. They describe a study in which people were asked to judge the lethality of various potential causes of death using different, arithmetically equivalent formats. For example, one group of people judged the lethality of heart attacks by responding to question (12a), whereas another group responded to the equivalent question (12b).

(12) a. For each 100,000 people afflicted, how many died?

b. For each person who died, how many were afflicted but survived?

To facilitate comparison of the judgments, answers to (12) were converted into estimates of deaths per 100,000 people afflicted. The average esti-

8. For more on the topic of grammatical parsing, see chapter 8 in volume 1.