Addiction is widely considered to be a chronic, relapsing disease. For instance, in an editorial in Science, we read that addiction is a "chronic relapsing disease of the nervous system" (Bloom 1997). In a companion article to that editorial, the director of the National Institute on Drug Abuse likens addiction to Alzheimer’s disease and schizophrenia, disorders that have no cure (Leshner 1997). And in an article in Time magazine, addiction is coupled with diabetes and hypertension, two diseases that likewise are chronic (Nash 1997).

However, research shows that many addicts recover. Among a group of inner-city heroin addicts in St. Louis, all claimed to have kicked their addiction by the time they were in their thirties (Robins and Murphy 1967). Although the sample in the St. Louis study was small, the result may be representative of most of those who become addicted. Large-scale epidemiological surveys reveal that there are millions of recovered smokers, alcoholics, and drug addicts (e.g. Robins and Regier 1991; Schelling 1992). In-depth studies of small populations of cocaine addicts (Waldorf, Reinarman, and Murphy 1991) and heroin addicts (Biernacki 1986) tell much the same story: several years of heavy drug use followed by an apparently enduring period of abstinence or controlled drug use. Possibly these studies are misleading, reflecting biased methods rather than the nature of addiction. On the other hand, perhaps the claim that addiction is a chronic, relapsing disease is misleading. The issue is an empirical one, and we now know enough about addiction to settle it.
The question of whether addiction is a chronic disorder is central to policy, treatment, and research. For instance, some male heroin addicts commit crimes at rates that approach one a day during periods of heavy drug use (e.g., Ball, Shaffer, and Nurco 1983). If addiction is typically chronic, then these men can be expected to commit scores if not hundreds of crimes every time they are released from prison. However, if addiction wanes with age or with the responsibilities that usually accompany age, it would be wrong to set sentences on the expectation of a lifelong pattern of drug-related crime. Clinicians who treat addicts face a different set of problems. Are they misleading their clients when they endorse the idea that recovery is the norm or when they endorse the idea that relapse is the norm? Presumably the ability to remain abstinent can be influenced by information about relapse rates, and it would be irresponsible to tell an addict that he or she had a chronic disease if in fact this was not true.

Neuroscientists who study drug-induced changes in the brain often identify these changes as the substrates of a chronic, relapsing disease, without mention of the reports indicating that recovery is the rule. That is, even in the scientific community there is not a general awareness of the conflicting findings regarding addiction relapse rates. This is unfortunate. Researchers may be assuming irreversible damage when in fact the brain changes are temporary and/or readily reversible by means of environmentally induced experiences. More generally, the brain may be a good deal more dynamic and plastic than assumed in current biological accounts of addiction. In short, whether we are judges, clinicians, or scientists, we need to know if addicts typically recover, or if "once an addict, always an addict."

Much of the research reviewed in this chapter was made possible by advances in the ability to reliably diagnose psychiatric cases. In the late 1970s the American Psychiatric Association revised its criteria for identifying psychiatric disorders. The goal was to ensure higher interrater reliability. By this standard, the new nosology (APA 1980) was a success (e.g., Spitzer and Forman 1979; Spitzer, Forman, and Nee 1979). On average, reliability increased by about 50 percent, and for substance-use disorders reliability scores were usually above 80 percent. An immediate consequence of diagnostic progress was scientific progress, especially in the area of psychiatric epidemiology. For instance, recent estimates of national frequencies of psychiatric disor-

ders often agree within a few percentage points, whereas earlier epidemiological research produced notoriously inconsistent results (e.g., Sandifer et al. 1968). Consequently, a good starting point for the investigation of relapse rates is the American Psychiatric Association's (APA) criteria for identifying addiction.

The most recent version of the APA diagnostic manual (1994) defines addiction ("substance dependence") as a "cluster of cognitive, behavioral, and physiological symptoms indicating that the individual continues to use the substance despite significant substance-related problems." Features of the disorder include tolerance, withdrawal, and loss of control over drug use. Loss of control, which is also referred to as "compulsive drug use," means such things as taking more of a drug than was initially intended, persistently trying and failing to curtail or quit taking drugs, and spending less time in conventional activities in order to pursue drug use. Tolerance and withdrawal are neither necessary nor sufficient for the diagnosis, and although the same can be said for loss of control, it is this feature of addiction that has been most emphasized by clinicians and scientists. "Compulsive" use leads to the adverse consequences that typify addiction and also to tolerance and withdrawal, and "compulsive" use is what is so hard to explain. For example, purchasing illicit drugs in the amounts required to maintain addiction requires planning and guile. And yet this pursuit, according to many clinicians, is "compulsive" and "out of control." But how can behavior that is planned also be "out of control"? (At the end of this chapter, this contradiction will be resolved. I will argue that "compulsive drug use" is better described as "ambivalent" drug use. The difference is important. Ambivalent users can be persuaded to stop, compulsive users cannot.)

Relapse Rates

The View from the Clinic

One of the major sources of information on addiction is research on treatment. The typical finding is that within a year or so of leaving the clinic the patient has resumed drug use (e.g., Stephens and Cottrell 1972; Vaillant 1966; Wasserman et al. 1998). Figure 3.1 shows some often-cited results. On the x-axis is time since the completion of treatment. On the y-axis is the percentage of patients who resumed drug use. Despite treatment, within 12 months most addicts had resumed
The question of whether addiction is a chronic disorder is central to policy, treatment, and research. For instance, some male heroin addicts commit crimes at rates that approach one a day during periods of heavy drug use (e.g., Ball, Shaffer, and Nurco 1983). If addiction is typically chronic, then these men can be expected to commit scores if not hundreds of crimes every time they are released from prison. However, if addiction wanes with age or with the responsibilities that usually accompany age, it would be wrong to set sentences on the expectation of a lifelong pattern of drug-related crime. Clinicians who treat addicts face a different set of problems. Are they misleading their clients when they endorse the idea that recovery is the norm or when they endorse the idea that relapse is the norm? Presumably, the ability to remain abstinent can be influenced by information about relapse rates, and it would be irresponsible to tell an addict that he or she had a chronic disease if in fact this was not true.

Neuroscientists who study drug-induced changes in the brain often identify these changes as the substrates of a chronic, relapsing disease, without mention of the reports indicating that recovery is the rule. That is, even in the scientific community there is not a general awareness of the conflicting findings regarding addiction relapse rates. This is unfortunate. Researchers may be assuming irreversible damage when in fact the brain changes are temporary and/or readily reversible by means of environmentally induced experiences. More generally, the brain may be a good deal more dynamic and plastic than assumed in current biological accounts of addiction. In short, whether we are judges, clinicians, or scientists, we need to know if addicts typically recover, or if “once an addict, always an addict.”

Much of the research reviewed in this chapter was made possible by advances in the ability to reliably diagnose psychiatric cases. In the late 1970s the American Psychiatric Association revised its criteria for identifying psychiatric disorders. The goal was to ensure higher interrater reliability. By this standard, the new nosology (APA 1980) was a success (e.g., Spitzer and Forman 1979; Spitzer, Forman, and Nee 1979). On average, reliability increased by about 50 percent, and, for substance-use disorders reliability scores were usually above 80 percent. An immediate consequence of diagnostic progress was scientific progress, especially in the area of psychiatric epidemiology. For instance, recent estimates of national frequencies of psychiatric disor-
not matter. However, if general health also mattered, then clinic populations could greatly overestimate relapse rates, especially if most addicts did not seek treatment.

Recovery and Relapse in Non-clinic Heroin Addicts
One way to avoid Berkson’s bias is to select subjects independently of whether they end up in treatment. The next two studies take this approach. In both, the subjects were heroin addicts.

Robins and Murphy (1967) studied the behavioral and familial antecedents of heroin addiction in African-American men who had grown up in St. Louis right after World War II. The men were identified on the basis of their elementary school registration forms, not drug use. While they were in their late teens and early twenties, some 13 percent of the sample experimented with heroin, and of this subgroup, about 75 percent became addicted. However, as they approached their late twenties and early thirties, they stopped using heroin. For the year prior to the interview, 84 percent claimed no heroin use, and the other 16 percent said they used occasionally but were not addicted. That is, according to self-report, the recovery rate was 100 percent. Official records support these results. Two-thirds of the men were known to the Federal Bureau of Narcotics, and of these, 74 percent did not have a record of heroin use in the five years prior to the study. Robins and Murphy add that according to health and judicial records, the men typically told the truth about their drug-use history (and there was no obvious advantage in misleading the researchers).

The second non-clinic study involves American servicemen who began using opiates (usually heroin) while in Vietnam. More than 40 percent used heroin at least once. Of those who tried the drug at least five times, about 90 percent went on to become regular users. In 1971 several thousand soldiers were returning from Vietnam each month. Given the clinic relapse rates (see Figure 3.1), it was widely, and sensibly, believed that a domestic heroin epidemic was imminent. President Nixon requested a study of the problem, and Lee Robins, who had directed the St. Louis study, was asked to head the project. She and her colleagues collected data on drug use in a sample of 898 men who
not matter. However, if general health also mattered, then clinic populations could greatly overestimate relapse rates, especially if most addicts did not seek treatment.

Recovery and Relapse in Non-clinic Heroin Addicts

One way to avoid Berkson's bias is to select subjects independently of whether they end up in treatment. The next two studies take this approach. In both, the subjects were heroin addicts.

Robins and Murphy (1967) studied the behavioral and familial antecedents of heroin addiction in African-American men who had grown up in St. Louis right after World War II. The men were identified on the basis of their elementary school registration forms, not drug use. While they were in their late teens and early twenties, some 13 percent of the sample experimented with heroin, and of this subgroup, about 75 percent became addicted. However, as they approached their late twenties and early thirties, they stopped using heroin. For the year prior to the interview, 84 percent claimed no heroin use, and the other 16 percent said they used occasionally but were not addicted. That is, according to self-report, the recovery rate was 100 percent. Official records support these results. Two-thirds of the men were known to the Federal Bureau of Narcotics, and of these, 74 percent did not have a record of heroin use in the five years prior to the study. Robins and Murphy add that according to health and judicial records, the men typically told the truth about their drug-use history (and there was no obvious advantage in misleading the researchers).

The second non-clinic study involves American servicemen who began using opiates (usually heroin) while in Vietnam. More than 40 percent used heroin at least once. Of those who tried the drug at least five times, about 90 percent went on to become regular users. In 1971 several thousand soldiers were returning from Vietnam each month. Given the clinic relapse rates (see Figure 3.1), it was widely, and sensibly, believed that a domestic heroin epidemic was imminent. President Nixon requested a study of the problem, and Lee Robins, who had directed the St. Louis study, was asked to head the project. She and her colleagues collected data on drug use in a sample of 898 men who
were scheduled to be discharged in September 1971 after serving in Vietnam (Robins, Helzer, and Davis 1975).

Of those who became regular heroin users, about 70 percent met the study’s criteria for addiction (withdrawal symptoms and difficulty quitting). But one year after returning to the United States, 95 percent were no longer regular users (ibid.), and three years later the remission rate was still close to 90 percent (Robins et al. 1980). This dramatic decrease in heroin use was not simply a matter of heroin’s becoming less available. About 50 percent of those addicted in Vietnam had tried heroin after returning home, yet they did not resume regular use of the drug.

Unfortunately, the authors provide little information about the recovery process. We can surmise that it was aided by a wide array of informal methods, as only 6 percent of the Vietnam opiate users went to drug treatment centers. (They were eligible for care at VA hospitals.) Figure 3.2 shows the Vietnam results and those of a typical clinic study. They are virtually mirror images of one another. Addicts in conventional treatment facilities typically returned to drug use; addicts who did not seek treatment typically recovered.

The clinic-based and non-clinic-based studies could not have produced more discrepant results. Nevertheless, the differences may be illusory. For example, the St. Louis men and the Vietnam enlistees may have been opiate users but not real opiate addicts. This distinction has a precedent. Some heroin users are able to regulate their intake so that their drug use does not interfere with other aspects of their life ("chippers"). For instance, a common pattern for the controlled heroin user is to restrict use to Saturdays, thereby ensuring that periods of intoxication do not interfere with work and allowing a day for recovery (Zinberg et al. 1977). Thus we should evaluate the possibility that the subjects in the two non-clinic studies were heroin users but not really heroin addicts.

The St. Louis men typically injected heroin for several years or more, 50 percent were sentenced to federal hospitals for addiction, and almost all were known to public officials as addicts. They were recognized by their peers and the authorities as street addicts. If the Vietnam sample is restricted to those who injected heroin more than once a week, the recovery rates are still more than 80 percent. The same holds for men who kept using heroin even though they knew that an opiate-positive urine test might delay their departure for home (Robins, Helzer, and Davis 1975). Thus the high recovery rates in these two non-clinic populations do not appear to be due to a too liberal definition of addiction.

A more important methodological issue is whether the recovery rate results are representative of addicts in general. Although 10 percent of the St. Louis sample became heroin addicts, this amounts to only 22 men. In the Vietnam research there were more subjects (386 opiate users), but their experience may not be relevant to conditions elsewhere. In Vietnam, heroin was cheap, use typically went unpunished, and the men were caught in a bitter, highly controversial war. Some might argue that this situation is too unusual to provide lessons about the nature of addiction. (On the other hand, these conditions
were scheduled to be discharged in September 1971 after serving in Vietnam (Robins, Helzer, and Davis 1975).

Of those who became regular heroin users, about 70 percent met the study's criteria for addiction (withdrawal symptoms and difficulty quitting). But one year after returning to the United States, 95 percent were no longer regular users (ibid.), and three years later the remission rate was still close to 90 percent (Robins et al. 1980). This dramatic decrease in heroin use was not simply a matter of heroin's becoming less available. About 50 percent of those addicted in Vietnam had tried heroin after returning home, yet they did not resume regular use of the drug.

Unfortunately, the authors provide little information about the recovery process. We can surmise that it was aided by a wide array of informal methods, as only 6 percent of the Vietnam opiate users went to drug treatment centers. (They were eligible for care at VA hospitals.) Figure 3.2 shows the Vietnam results and those of a typical clinic study. They are virtually mirror images of one another. Addicts in conventional treatment facilities typically returned to drug use; addicts who did not seek treatment typically recovered.

The clinic-based and non-clinic-based studies could not have produced more discrepant results. Nevertheless, the differences may be illusory. For example, the St. Louis men and the Vietnam enlistees may have been opiate users but not real opiate addicts. This distinction has a precedent. Some heroin users are able to regulate their intake so that their drug use does not interfere with other aspects of their life ("chippers"). For instance, a common pattern for the controlled heroin user is to restrict use to Saturdays, thereby ensuring that periods of intoxication do not interfere with work and allowing a day for recovery (Zinberg et al. 1977). Thus we should evaluate the possibility that the subjects in the two non-clinic studies were heroin users but not really heroin addicts.

The St. Louis men typically injected heroin for several years or more, 50 percent were sentenced to federal hospitals for addiction, and almost all were known to public officials as addicts. They were recognized by their peers and the authorities as street addicts. If the Vietnam sample is restricted to those who injected heroin more than once a week, the recovery rates are still more than 80 percent. The same holds for men who kept using heroin even though they knew that an opiate-positive urine test might delay their departure for home (Robins, Helzer, and Davis 1975). Thus the high recovery rates in these two non-clinic populations do not appear to be due to a too liberal definition of addiction.

A more important methodological issue is whether the recovery rate results are representative of addicts in general. Although 10 percent of the St. Louis sample became heroin addicts, this amounts to only 22 men. In the Vietnam research there were more subjects (386 opiate users), but their experience may not be relevant to conditions elsewhere. In Vietnam, heroin was cheap, use typically went unpunished, and the men were caught in a bitter, highly controversial war. Some might argue that this situation is too unusual to provide lessons about the nature of addiction. (On the other hand, these conditions
may not be too different from those experienced by addicts living in neighborhoods blighted by abandoned buildings, shooting galleries, and grudging tolerance of drug sales and use.)

This chapter began with a discrepancy: addiction is often referred to as a chronic, relapsing disorder, yet many addicts recover. The data reviewed so far provide a neat and comprehensive resolution. Research on addiction has been largely restricted to those addicts who end up in treatment; addicts in treatment typically relapse within a year or so (e.g. Brecher 1972; Thombs 1994; Wasserman et al. 1998). In contrast, those addicts who do not end up in treatment typically recover (Robins and Murphy 1967; Robins 1993). However, so far there are only two non-clinic-based studies. The next section tests the generality of the treatment-vs.-nontreatment hypothesis.

**Large National Surveys in Community Samples**

Researchers recognized that clinic samples might provide a biased picture of addiction, especially if many addicts did not seek treatment. Clinicians saw this problem somewhat differently. They were concerned that those who suffered from addiction and other psychiatric disorders were not getting the treatment they needed. Both issues pointed to the need for a survey of mental health problems in a large, representative sample. In the late 1970s circumstances fell into place to make this sort of survey possible.

Shortly following her husband’s inauguration, Rosalynn Carter, wife of President Jimmy Carter, convened a meeting of mental health experts at the White House. The experts recommended a nationwide survey of psychiatric disorders, including addiction (Regier et al. 1984). The National Institutes of Health sponsored the research, now known as the Epidemiological Catchment Area Study (ECA), and the results were summarized in a book published in 1991 (Robins and Regier 1991).

The ECA selected subjects from five major metropolitan areas, independently of their treatment history. Because of the size of the effort (nearly 20,000 subjects), the interviewers (about 200) were not professional clinicians but a specially trained lay staff. Their primary instrument was a questionnaire designed so that the answers could be classified in terms of the recently revised and field-tested American Psychiatric Association diagnostic categories (APA 1980).

For most diagnoses, the reliability between lay interviewers and psychiatrists was as good as that between different psychiatrists (Helzer et al. 1983), and the average number of symptoms per case was virtually identical for lay and professionally trained interviewers (Helzer, Spitzer, and McEvoy 1987). In addition, some ten years after the ECA survey there was a second large, nationwide evaluation of psychiatric health in community samples. This survey, known as the National Comorbidity Study (NCS), provides a convenient check on the reliability of the ECA results. Other methodological issues, such as whether the sample population was representative of addicts in general or whether the interviewees accurately reported their drug use, will be addressed later.

Figure 3.3 shows the ECA and NCS estimates of remission rates for addiction and other psychiatric disorders. In the ECA study, remission was defined as no symptoms for the year just prior to the interview. In the NCS study, which included about 8,000 respondents, the criteria for remission was anything less than the minimum set of symptoms for establishing a diagnosis. That is, the ECA criteria for remission were more conservative. In both studies substance-use disorders had the highest remission rates. According to the NCS results, 76 percent of all of those with a lifetime diagnosis were not addicted for a year or more prior to the interview. According to the ECA respondents, the remission rate was 59 percent. For other psychiatric diagnoses the remission rates were lower, and in much closer agreement. Indeed, if substance-use disorder is not included, then the average difference in relapse rates for the NCS and ECA surveys is only 2 percent. One interpretation of this pattern of findings is that the ECA and NCS used similar criteria for identifying active and recovered cases for every disorder but addiction. In support of this interpretation, when the NCS researchers use the ECA criteria for addiction, the difference in remission rates shrinks to less than 5 percent (Warner et al. 1995). Thus the “discrepancy” appears to reflect the faithful application of the diagnostic criteria, rather than unreliable instruments.

Although both surveys found that substance-use disorders had the highest remission rates, this result requires further analysis. The NCS
may not be too different from those experienced by addicts living in neighborhoods blighted by abandoned buildings, shooting galleries, and grudging tolerance of drug sales and use.)

This chapter began with a discrepancy: addiction is often referred to as a chronic, relapsing disorder, yet many addicts recover. The data reviewed so far provide a neat and comprehensive resolution. Research on addiction has been largely restricted to those addicts who end up in treatment; addicts in treatment typically relapse within a year or so (e.g. Brecher 1972; Thombs 1994; Wasserman et al. 1998). In contrast, those addicts who do not end up in treatment typically recover (Robins and Murphy 1967; Robins 1993). However, so far there are only two non-clinic-based studies. The next section tests the generality of the treatment-vs.-nontreatment hypothesis.

Large National Surveys in Community Samples

Researchers recognized that clinic samples might provide a biased picture of addiction, especially if many addicts did not seek treatment. Clinicians saw this problem somewhat differently. They were concerned that those who suffered from addiction and other psychiatric disorders were not getting the treatment they needed. Both issues pointed to the need for a survey of mental health problems in a large, representative sample. In the late 1970s circumstances fell into place to make this sort of survey possible.

Shortly following her husband’s inauguration, Rosalynn Carter, wife of President Jimmy Carter, convened a meeting of mental health experts at the White House. The experts recommended a nationwide survey of psychiatric disorders, including addiction (Regier et al. 1984). The National Institutes of Health sponsored the research, now known as the Epidemiological Catchment Area Study (ECA), and the results were summarized in a book published in 1991 (Robins and Regier 1991).

The ECA selected subjects from five major metropolitan areas, independently of their treatment history. Because of the size of the effort (nearly 20,000 subjects), the interviewers (about 200) were not professional clinicians but a specially trained lay staff. Their primary instrument was a questionnaire designed so that the answers could be classified in terms of the recently revised and field-tested American Psychiatric Association diagnostic categories (APA 1980).

For most diagnoses, the reliability between lay interviewers and psychiatrists was as good as that between different psychiatrists (Helzer et al. 1985), and the average number of symptoms per case was virtually identical for lay and professionally trained interviewers (Helzer, Spitzagel, and McEvoy 1987). In addition, some ten years after the ECA survey there was a second large, nationwide evaluation of psychiatric health in community samples. This survey, known as the National Comorbidity Study (NCS), provides a convenient check on the reliability of the ECA results. Other methodological issues, such as whether the sample population was representative of addicts in general or whether the interviewees accurately reported their drug use, will be addressed later.

Figure 3.3 shows the ECA and NCS estimates of remission rates for addiction and other psychiatric disorders. In the ECA study, remission was defined as no symptoms for the year just prior to the interview. In the NCS study, which included about 8,000 respondents, the criterion for remission was anything less than the minimum set of symptoms for establishing a diagnosis. That is, the ECA criteria for remission were more conservative. In both studies substance-use disorders had the highest remission rates. According to the NCS results, 76 percent of all of those with a lifetime diagnosis were not addicted at a year or more prior to the interview. According to the ECA respondents, the remission rate was 59 percent. For other psychiatric diagnoses the remission rates were lower, and in much closer agreement. Indeed, if substance-use disorder is not included, then the average difference in relapse rates for the NCS and ECA surveys is only 2 percent. One interpretation of this pattern of findings is that the ECA and NCS used similar criteria for identifying active and recovered cases for every disorder but addiction. In support of this interpretation, when the NCS researchers use the ECA criteria for addiction, the difference in remission rates shrinks to less than 5 percent (Warner et al. 1995). Thus the “discrepancy” appears to reflect the faithful application of the diagnostic criteria, rather than unreliable instruments.

Although both surveys found that substance-use disorders had the highest remission rates, this result requires further analysis. The NCS
As expected, cocaine and opiates had the lowest recovery rates, but the more important point is that for the major illicit addictive drugs, the remission rates were quite similar and reasonably well represented by the average value. Marijuana users did not skew the results.

The national survey results lead to the same conclusion as the St. Louis and Vietnam findings. Most addicts recover, but this is only apparent if the addicts are selected independently of their treatment history. An immediate implication is that addiction is reversible. Before addressing this issue, I will review data on the duration of addiction. Low relapse rates suggest a relatively short duration. However, this is a logical point, and the data could turn out differently. The duration results will also provide a kind of check on the relapse findings. If addiction has the lowest relapse rate of any psychiatric disorder, then it should, all else being equal, have the shortest duration of any psychi-
and ECA typically did not differentiate between the various illicit drugs. Results for opiate, stimulant, and marijuana use were averaged together, as if these drugs were sufficiently similar to be considered a single category. However, they have markedly different pharmacological and behavioral effects, which may well lead to quite different relapse rates. In particular, many experts believe that the consequences of frequent marijuana use are significantly less debilitating than the consequences of frequent stimulant and opiate use. Thus the survey results on relapse may be accurate for marijuana but not for opiates and stimulants. Figure 3.4 addresses this issue. (It is based on the one table in the ECA summary that organizes drug-use statistics by drug class).

As expected, cocaine and opiates had the lowest recovery rates, but the more important point is that for the major illicit addictive drugs, the remission rates were quite similar and reasonably well represented by the average value. Marijuana users did not skew the results.

The national survey results lead to the same conclusion as the St. Louis and Vietnam findings. Most addicts recover, but this is only apparent if the addicts are selected independently of their treatment history. An immediate implication is that addiction is reversible. Before addressing this issue, I will review data on the duration of addiction. Low relapse rates suggest a relatively short duration. However, this is a logical point, and the data could turn out differently. The duration results will also provide a kind of check on the relapse findings. If addiction has the lowest relapse rate of any psychiatric disorder, then it should, all else being equal, have the shortest duration of any psychi-
Duration

The ECA report provides estimates of the duration of addiction for active and remitted cases. From these two pieces of information plus the remission rates just reviewed, it is possible to get some idea of how long addiction typically lasts.

The ECA researchers identified the onset of a disorder as the initial expression of one or more symptoms (rather than when the full case criteria were first met). For individuals who met the criteria for abuse and/or dependence at the time of the interview (current addicts), the average time since onset was 6.1 years and the median time was between 4 and 5 years (Anthony and Helzer 1991). For individuals in remission for 3 or more years (no symptoms related to drug use for at least 3 years), the average time from onset to remission was 2.7 years and the median duration was between 1 and 2 years. When the mean and median of a distribution markedly differ, the distribution is not bell-shaped but asymmetrical. When this is true, the median is the more representative population measure. Thus for recovered addicts (no symptoms for three years), addiction typically lasted less than two years.

These estimates are based on interviews, and there were no independent checks as to their validity. They could be accurate, but they also could reflect the manner in which questions were worded, the tendency, when providing a history, to reconstruct the past in terms of current circumstances, and normal difficulties in accurately remembering subtle changes in behavior, especially ones that took place gradually over an extended period.

One way to correct for the errors inherent in retrospective research is to use relative rather than absolute measures. For instance, under the assumption that errors and distortions are more or less equally likely across the different disorders, the ratio of the remembered duration of one disorder to the remembered duration of another disorder should reflect the actual ratios. Figure 3.5 is motivated by this logic. It shows the ECA estimates of duration for the more frequent psychiatric disorders.

Substance-use disorder had the shortest duration, and the differences are substantial: in remitted cases, the average psychiatric disorder lasted about four times longer than did addiction, and schizophrenia lasted about seven times longer.

Despite the methodological problems in estimating duration, there are reasons to have some confidence in the results shown in this figure. First, there is an inverse relationship between duration and remission, which is the simplest possible relationship. Second, the estimated durations agree with clinical experience. Those that were longer, such as schizophrenia, clinicians find least tractable, and those that were shorter, such as anxiety disorders, are thought to be more treatable.
Drugs Addiction and Drug Policy

atric disorder. But the claims published in Science and Time that introduced this chapter suggest a quite different outcome.

**Duration**

The ECA report provides estimates of the duration of addiction for active and remitted cases. From these two pieces of information plus the remission rates just reviewed, it is possible to get some idea of how long addiction typically lasts. The ECA researchers identified the onset of a disorder as the initial expression of one or more symptoms (rather than when the full case criteria were first met). For individuals who met the criteria for abuse and/or dependence at the time of the interview (current addicts), the average time since onset was 6.1 years and the median time was between 4 and 5 years (Anthony and Helzer 1991). For individuals in remission for 3 or more years (no symptoms related to drug use for at least 3 years), the average time from onset to remission was 2.7 years and the median duration was between 1 and 2 years. When the mean and median of a distribution markedly differ, the distribution is not bell-shaped but asymmetrical. When this is true, the median is the more representative population measure. Thus for recovered addicts (no symptoms for three years), addiction typically lasted less than two years.

These estimates are based on interviews, and there were no independent checks as to their validity. They could be accurate, but they also could reflect the manner in which questions were worded, the tendency, when providing a history, to reconstruct the past in terms of current circumstances, and normal difficulties in accurately remembering subtle changes in behavior, especially ones that took place gradually over an extended period.

One way to correct for the errors inherent in retrospective research is to use relative rather than absolute measures. For instance, under the assumption that errors and distortions are more or less equally likely across the different disorders, the ratio of the remembered duration of one disorder to the remembered duration of another disorder should reflect the actual ratios. Figure 3.5 is motivated by this logic. It shows the ECA estimates of duration for the more frequent psychiatric disorders.

![Figure 3.5. Duration of psychiatric disorders for those in remission. Source: ECA estimates, Robins and Regier 1991.](image)

Substance-use disorder had the shortest duration, and the differences are substantial: in remitted cases, the average psychiatric disorder lasted about four times longer than did addiction, and schizophrenia lasted about seven times longer.

Despite the methodological problems in estimating duration, there are reasons to have some confidence in the results shown in this figure. First, there is an inverse relationship between duration and remission, which is the simplest possible relationship. Second, the estimated durations agree with clinical experience. Those that were longer, such as schizophrenia, clinicians find least tractable, and those that were shorter, such as anxiety disorders, are thought to be more treatable.
phrenia. (Individual biographies and survey data show that addiction often comes to an end, so it is sensible to ask how long it lasts. In contrast, for many other psychiatric disorders the notion of an endpoint is really not sensible, at least for a large portion of the victims. Schizophrenia and depression are often lifelong, chronic maladies. This observation does not make the figures meaningless; rather it reinforces the point they make: among psychiatric disorders, addiction is usually the outlier.)

**Summary of Relapse, Remission, and Duration Results**

According to the idea that addiction is a chronic disorder, when addicts go on the wagon, they are soon to fall off. In support of this view are the results from clinic outcome studies showing high relapse rates (e.g. Wasserman et al. 1998). However, studies that did not use clinic populations (e.g. Robins and Murphy 1967; Robins, Helzer, and Davis 1975) showed just the opposite result: addicts recovered. Large-scale epidemiological surveys in which subjects were selected independent of treatment history showed the same pattern as the non-clinic research. Of all the psychiatric disorders, addiction had the highest remission rate and the shortest duration. Thus, once you sample addicts in a nonbiased manner, addiction no longer appears to be a chronic, relapsing disorder.

However, the survey findings do not imply that the clinic results are invalid or unimportant. Rather, they show that there are large individual differences. Although most addicts recover, many struggle for years, cycling back and forth between sobriety and heavy drug use (Brecher 1972; Vaillant 1992). To understand addiction, it is necessary to understand why the route to recovery is so varied. The most obvious starting point is “Why are addicts who seek treatment more likely to relapse than those who do not seek treatment?”

**Addicts Who Seek Treatment vs. Those Who Do Not**

The contrast between clinic relapse rates and the ECA and NCS estimates of relapse implies that most addicts do not seek treatment and suggests that the correlates of the differences between those who do and do not seek treatment will provide clues to the factors that deter recovery. The ECA data support the logic. Approximately 70 percent
phrenia. (Individual biographies and survey data show that addiction often comes to an end, so it is sensible to ask how long it lasts. In contrast, for many other psychiatric disorders the notion of an endpoint is really not sensible, at least for a large portion of the victims. Schizophrenia and depression are often lifelong, chronic maladies. This observation does not make the figures meaningless; rather it reinforces the point they make: among psychiatric disorders, addiction is usually the outlier.)

Summary of Relapse, Remission, and Duration Results
According to the idea that addiction is a chronic disorder, when addicts go on the wagon, they are soon to fall off. In support of this view are the results from clinic outcome studies showing high relapse rates (e.g. Wasserman et al. 1998). However, studies that did not use clinic populations (e.g. Robins and Murphy 1967; Robins, Helzer, and Davis 1975) showed just the opposite result: addicts recovered. Large-scale epidemiological surveys in which subjects were selected independent of treatment history showed the same pattern as the nonclinic research. Of all the psychiatric disorders, addiction had the highest remission rate and the shortest duration. Thus, once you sample addicts in a nonbiased manner, addiction no longer appears to be a chronic, relapsing disorder.

However, the survey findings do not imply that the clinic results are invalid or unimportant. Rather, they show that there are large individual differences. Although most addicts recover, many struggle for years, cycling back and forth between sobriety and heavy drug use (Brecher 1972; Vaillant 1992). To understand addiction, it is necessary to understand why the route to recovery is so varied. The most obvious starting point is “Why are addicts who seek treatment more likely to relapse than those who do not seek treatment?”

Addicts Who Seek Treatment vs. Those Who Do Not
The contrast between clinic relapse rates and the ECA and NCS estimates of relapse implies that most addicts do not seek treatment and suggests that the correlates of the differences between those who do and do not seek treatment will provide clues to the factors that deter recovery. The ECA data support the logic. Approximately 70 percent
of those with a lifetime diagnosis of substance-use disorder were not treated for drug use. However, differences between treated and untreated addicts have been little studied. Other than two reports from researchers at Yale University, there is little published research on this important topic.

The available results will be organized in terms of three factors: pharmacological history; consequences of drug use, such as arrests; and individual traits, such as psychiatric history. (It is likely that those who do not end up in treatment centers get help from friends, family, and co-workers. This, though, does not rule out the possibility that "treated" and "untreated" addicts truly differ.)

Pharmacological history. It is reasonable to suppose that those who do not seek treatment are relatively new users, who have not been as exposed to drugs as those in treatment. Some research supports this view (e.g. Chitwood and Chitwood 1981; Graeven and Graeven 1983). However, authors of more recent studies (e.g. Carroll and Rounsaville 1992) point out that earlier researchers did not use field-tested diagnostic criteria or attempt to ensure that both populations met some minimum threshold for addiction. The criticism appears to be valid: a different picture emerges in the two Yale studies, in which both treated and untreated populations were selected according to the APA diagnostic criteria (Rounsaville and Kleber 1985; Carroll and Rounsaville 1992). One was with heroin addicts, the other with cocaine addicts. The treated subjects were from the clinic, the untreated subjects were contacted by word of mouth. In these studies both treated and untreated individuals "qualified" as addicts according to the APA criteria.

Level of drug use. Cocaine addicts who did not utilize clinics exceeded the clinic cocaine addicts on all measures of drug use. In particular, they were more likely to use a wide array of addictive drugs. The same pattern held for heroin addicts. Both treated and untreated addicts had been using heroin daily for about six years, but the non-treatment group reported more alcohol and marijuana use. Thus the two most complete studies fail to provide evidence that pharmacological history distinguishes treated and untreated addicts.

Adverse consequences of drug use. Treatment-seeking cocaine addicts reported significantly higher levels of recent depression and anxiety, and significantly greater negative consequences of cocaine use in relation to family, friends, and work (Carroll and Rounsaville 1992). Among heroin addicts, those who sought clinic help had more severe drug-related problems in respect to social interactions and significantly more arrests for drug possession and sales. (The "untreated" addicts obtained about 70 percent of the income for heroin by illegal acts and reported that they had been engaged in criminal activity for profit in about 14 out of the last 30 days.) In summarizing the findings for opiate addicts, Rounsaville and Kleber (1985) write, "Overall, the findings seem to indicate that while heavy drug use per se may not be a primary motivation to seek treatment, social, legal, and psychological problems that are acutely associated with the drug use may provide the incentive to apply for help." Rounsaville and Kleber's point is that the two groups did not differ in regard to their level of drug use, but did differ in regard to the number of problems they reported, and that many of these problems were immediate consequences of drug use.

Individual differences. The Yale studies show that similar pharmacological histories did not lead to similar outcomes. This implies that individual differences mediate many critical drug effects. Individual differences in drug metabolism are likely to play an important role, and Rounsaville and Kleber (1985) stress individual differences in social support. However, the best evidence for differences between treated and untreated addicts comes from the study of co-occurring psychiatric disorders ("comorbidity").

Regier et al. (1990) write that the expected likelihood of a non-drug-related psychiatric disorder among addicts in the ECA sample, assuming no increased risk, was 22 percent. In contrast, the observed frequency was more than twice as great, 53 percent. The increased liability varied by drug class. About 50 percent of marijuana abusers met the criteria for an additional psychiatric diagnosis, whereas the proportions were 65 percent and 76 percent for opiate and cocaine disorders, respectively. However, the largest difference was between those who sought treatment and those who did not. For those who did not
of those with a lifetime diagnosis of substance-use disorder were not treated for drug use. However, differences between treated and untreated addicts have been little studied. Other than two reports from researchers at Yale University, there is little published research on this important topic.

The available results will be organized in terms of three factors: pharmacological history; consequences of drug use, such as arrests; and individual traits, such as psychiatric history. (It is likely that those who do not end up in treatment centers get help from friends, family, and co-workers. This, though, does not rule out the possibility that “treated” and “untreated” addicts truly differ.)

Pharmacological history. It is reasonable to suppose that those who do not seek treatment are relatively new users, who have not been as exposed to drugs as those in treatment. Some research supports this view (e.g. Chitwood and Chitwood 1981; Graeven and Graeven 1983). However, authors of more recent studies (e.g. Carroll and Rounsaville 1992) point out that earlier researchers did not use field-tested diagnostic criteria or attempt to ensure that both populations met some minimum threshold for addiction. The criticism appears to be valid: a different picture emerges in the two Yale studies, in which both treated and untreated populations were selected according to the APA diagnostic criteria (Rounsaville and Kleber 1985; Carroll and Rounsaville 1992). One was with heroin addicts, the other with cocaine addicts. The treated subjects were from the clinic, the untreated subjects were contacted by word of mouth. In these studies both treated and untreated individuals “qualified” as addicts according to the APA criteria.

Level of drug use. Cocaine addicts who did not utilize clinics exceeded the clinic cocaine addicts on all measures of drug use. In particular, they were more likely to use a wide array of addictive drugs. The same pattern held for heroin addicts. Both treated and untreated addicts had been using heroin daily for about six years, but the non-treatment group reported more alcohol and marijuana use. Thus the two most complete studies fail to provide evidence that pharmacological history distinguishes treated and untreated addicts.

Adverse consequences of drug use. Treatment-seeking cocaine addicts reported significantly higher levels of recent depression and anxiety, and significantly greater negative consequences of cocaine use in relation to family, friends, and work (Carroll and Rounsaville 1992). Among heroin addicts, those who sought clinic help had more severe drug-related problems in respect to social interactions and significantly more arrests for drug possession and sales. (The “untreated” addicts obtained about 70 percent of the income for heroin by illegal acts and reported that they had been engaged in criminal activity for profit in about 14 out of the last 30 days.) In summarizing the findings for opiate addicts, Rounsaville and Kleber (1985) write, “Overall, the findings seem to indicate that while heavy drug use per se may not be a primary motivation to seek treatment, social, legal, and psychological problems that are acutely associated with the drug use may provide the incentive to apply for help.” Rounsaville and Kleber’s point is that the two groups did not differ in regard to their level of drug use, but did differ in regard to the number of problems they reported, and that many of these problems were immediate consequences of drug use.

Individual differences. The Yale studies show that similar pharmacological histories did not lead to similar outcomes. This implies that individual differences mediate many critical drug effects. Individual differences in drug metabolism are likely to play an important role, and Rounsaville and Kleber (1985) stress individual differences in social support. However, the best evidence for differences between treated and untreated addicts comes from the study of co-occurring psychiatric disorders (“comorbidity”).

Regier et al. (1990) write that the expected likelihood of a non-drug-related psychiatric disorder among addicts in the ECA sample, assuming no increased risk, was 22 percent. In contrast, the observed frequency was more than twice as great, 53 percent. The increased liability varied by drug class. About 50 percent of marijuana abusers met the criteria for an additional psychiatric diagnosis, whereas the proportions were 65 percent and 76 percent for opiate and cocaine disorders, respectively. However, the largest difference was between those who sought treatment and those who did not. For those who did not
seek treatment, the prevalence of other psychiatric disorders was about 29 percent, not too different from the expected value. Among treatment seekers, the prevalence of other mental disorders was more than twice as great, 64 percent (and almost three times greater than the expected value). Thus the persistence of addiction was closely tied to the presence of a co-occurring disorder.

The ECA results are supported by smaller, clinic-based studies. In a comparison of drug users who signed up for treatment versus those who signed up to be experimental subjects (Montoya et al. 1995), the treatment seekers scored higher on eight of nine measures of psychopathology (and the non-treatment group was in the “normal” range on all measures). In the work done at Yale, treatment seekers were significantly more likely to be suffering from depression.

Why do co-occurring psychiatric disorders deter recovery from addiction? The correlational structure across disorders is complex. Addicts show elevated risks for almost all diagnoses, with the correlations for conduct disorder, bipolar disorder, and depression ranking highest (Regier et al. 1990). This pattern supports the idea that people who use addictive drugs to medicate themselves are more likely to become addicted (e.g. Khantzian 1985). Also, psychiatric disorders may undermine activities that would normally provide a compelling alternative to the addictive drug. For example, addicts who become involved in athletics, hobbies, and clubs are more likely to successfully abstain from drug use (Waldorf 1983). Unfortunately, those with psychiatric disorders are less likely to join groups and take up hobbies. Thus, with too much time on their hands, the psychiatric patients turn to addictive drugs. This point is similar to the self-medication hypothesis, but differs in that drug use fills a void rather than functions as medication for a specific disorder.

That most addicts recover has not been widely appreciated, and thus little has been written about differences between those who successfully abstain and those who do not. However, the available evidence is consistent. Pharmacological history did not make any obvious difference in recovery rates, whereas individual differences in psychiatric history did. This is not to say that drug pharmacology is unimportant. For example, it is possible that the biological effects of addictive drugs are quite different as a function of additional psychiatric disorders, and, as noted in the conclusion of this chapter, pharmacological treatments for addiction are of proven worth and should be further developed. Nevertheless, the simplest account of the available data is that recovery from addiction is significantly influenced by nonpharmacological factors.

Is Addiction a Disease?

There are several senses in which it is legitimate to call addiction a disease. According to one dictionary, a disease is a “departure from a normal condition in a negative way that can be identified by a characteristic group of signs or symptoms.” The definition applies. Most consider addiction a departure from the norm, and it can be reliably identified by clinicians. However, by the same criteria, behaviors that we are more likely to call bad habits become diseases. For instance, there are people who spend an abnormally large amount of time watching television—about eight to twelve hours a day (Goleman 1990). At the end of their television marathon, they feel bad about how much time they have wasted and state that they wish they had spent the time more productively. Nevertheless, the next day they are in front of the television again. Their behavior is excessive and, from their own perspective, deleterious. Thus, by the dictionary definition, excessive television watching is a disease. Similar analyses easily apply to long hours at the office, surfing the net, and the scores of other familiar yet excessive behaviors. But most of us believe these activities are clearly different from having cancer, heart disease, or even diabetes (which has a large behavioral component). In support of this intuition, everyday use of the term “disease” reveals a more discriminating understanding than the dictionary definition. In everyday speech “disease” has two rather specific meanings.

First, many call addiction a disease because some instances can be tied to a biological predisposition and because addictive drugs change the brain (e.g. Lewis 1991; Leshner 1997; Maltzman 1994). However, all goal-oriented behavior is mediated by the brain, and all learned behavior depends on changes in the brain. For example, heritability studies indicate a genetic predisposition for various forms of criminal behavior (Wilson and Herrnstein 1985), and it seems likely that future
seek treatment, the prevalence of other psychiatric disorders was about 29 percent, not too different from the expected value. Among treatment seekers, the prevalence of other mental disorders was more than twice as great, 64 percent (and almost three times greater than the expected value). Thus the persistence of addiction was closely tied to the presence of a co-occurring disorder.

The ECA results are supported by smaller, clinic-based studies. In a comparison of drug users who signed up for treatment versus those who signed up to be experimental subjects (Montoya et al. 1995), the treatment seekers scored higher on eight of nine measures of psychopathology (and the non-treatment group was in the “normal” range on all measures). In the work done at Yale, treatment seekers were significantly more likely to be suffering from depression.

Why do co-occurring psychiatric disorders deter recovery from addiction? The correlational structure across disorders is complex. Addicts show elevated risks for almost all diagnoses, with the correlations for conduct disorder, bipolar disorder, and depression ranking highest (Regier et al. 1990). This pattern supports the idea that people who use addictive drugs to medicate themselves are more likely to become addicted (e.g. Khantzian 1985). Also, psychiatric disorders may undermine activities that would normally provide a compelling alternative to the addictive drug. For example, addicts who become involved in athletics, hobbies, and clubs are more likely to successfully abstain from drug use (Waldorf 1983). Unfortunately, those with psychiatric disorders are less likely to join groups and take up hobbies. Thus, with too much time on their hands, the psychiatric patients turn to addictive drugs. This point is similar to the self-medication hypothesis, but differs in that drug use fills a void rather than functions as medication for a specific disorder.

That most addicts recover has not been widely appreciated, and thus little has been written about differences between those who successfully abstain and those who do not. However, the available evidence is consistent. Pharmacological history did not make any obvious difference in recovery rates, whereas individual differences in psychiatric history did. This is not to say that drug pharmacology is unimportant. For example, it is possible that the biological effects of addictive drugs are quite different as a function of additional psychiatric disorders, and, as noted in the conclusion of this chapter, pharmacological treatments for addiction are of proven worth and should be further developed. Nevertheless, the simplest account of the available data is that recovery from addiction is significantly influenced by nonpharmacological factors.

Is Addiction a Disease?

There are several senses in which it is legitimate to call addiction a disease. According to one dictionary, a disease is a “departure from a normal condition in a negative way that can be identified by a characteristic group of signs or symptoms.” The definition applies. Most consider addiction a departure from the norm, and it can be reliably identified by clinicians. However, by the same criteria, behaviors that we are more likely to call bad habits become diseases. For instance, there are people who spend an abnormally large amount of time watching television—about eight to twelve hours a day (Goleman 1990). At the end of their television marathon, they feel bad about how much time they have wasted and state that they wish they had spent the time more productively. Nevertheless, the next day they are in front of the television again. Their behavior is excessive and, from their own perspective, deleterious. Thus, by the dictionary definition, excessive television watching is a disease. Similar analyses easily apply to long hours at the office, surfing the net, and the scores of other familiar yet excessive behaviors. But most of us believe these activities are clearly different from having cancer, heart disease, or even diabetes (which has a large behavioral component). In support of this intuition, everyday use of the term “disease” reveals a more discriminating understanding than the dictionary definition. In everyday speech “disease” has two rather specific meanings.

First, many call addiction a disease because some instances can be tied to a biological predisposition and because addictive drugs change the brain (e.g. Lewis 1991; Leshner 1997; Maltzman 1994). However, all goal-oriented behavior is mediated by the brain, and all learned behavior depends on changes in the brain. For example, heritability studies indicate a genetic predisposition for various forms of criminal behavior (Wilson and Herrnstein 1985), and it seems likely that future
research will show that certain environmental events alter the brain in ways that increase the likelihood of violence (e.g. Miczek 1999). Thus, by the criterion of "biological basis," all crimes would eventually become diseases. But this would lead to foolish and morally unacceptable policies. For instance, medical hospitalization for crimes such as murder and car theft would violate widely shared ideas of justice and, if consequences count, would prove counterproductive. Thus the idea that a disorder is a disease because it has a biological basis or because it entails brain changes is too crude a standard.

The second colloquial meaning of "disease" is that it is an involuntary, as opposed to a voluntary, disorder. This turns out to be a practical and scientifically defensible distinction. The mechanisms mediating voluntary and involuntary disorders differ, especially in regard to the influence of the central nervous system. Treatments for voluntary and involuntary disorders often differ (it would be cruel to punish tics or hallucinations, but it might be quite helpful and humane to provide corrective incentives for excessive television watchers). And social institutions for the remediation of voluntary and involuntary behavioral problems differ (e.g. "wise men" and "medicine men"). In short, the question of whether addiction is a voluntary or involuntary disorder matters (whereas there should be no question that addiction has a biological basis).

The idea that addicts take drugs involuntarily has been articulately argued by clinicians and researchers. Miller and Chappel (1991), psychiatrists, explained that addicts have a disease because they have lost control over drug use. Jellinek (1960), one of the first to systematically study alcoholism, defined alcoholism as a disease on the basis of loss of control over drinking. More recently, Leshner (1997), the head of the National Institute on Drug Abuse, claimed that repeated ingestion of addictive drugs "turns off" voluntary control of drug use.

However, the idea that a behavior is "out of control" does not seem to automatically qualify it as a disease. The excessive television watchers claimed that their behavior was out of control, and yet most observers would want to distinguish their problems from those of cancer patients, diabetics, schizophrenics, and even addicts. Also, the statement that addicts are unable to control their drug use is not accurate. For example, pack-a-day cigarette smokers meet the DSM criteria for addiction, yet since the Surgeon General's report in 1964 on the health effects of smoking, some 50 million heavy smokers have quit using cigarettes (Schelling 1992). Moreover, most quit on their own, without medical help. If addiction is out-of-control drug use, how is this possible? On the other hand, it is no simple matter to quit an addictive drug. Thus the question of whether addiction is in fact involuntary drug use has remained controversial (e.g. Vuchinich 1996). Part of the problem is that there are no widely agreed upon criteria for identifying voluntary acts.

Criteria for Identifying Voluntary Acts

The distinction between voluntary and involuntary acts pervades the discussion and analysis of behavior. In philosophy and political theory, the distinction is usually made along the lines of conscious intentions (Searle 1983) and/or freedom from authoritarian regimes. In scientific studies of behavior, the criteria have to do with the factors that influence behavior. Some behaviors are elicited by stimulus conditions and are relatively immune to reward and punishment (e.g. reflexes). Other behaviors have no specific eliciting conditions, but instead are learned; their frequency is a function of deprivation and relative reward and punishment. For example, consider the contrast between two simple, topographically similar behaviors, "eye blinks" and "winks."

An eye blink is a "wired-in" behavior that is readily elicited by a directed force to the eye, such as a puff of air. The wiring admits specific eliciting stimuli, but provides few if any inputs for the influence of incentives or payoffs (contingencies). For instance, if a one-dollar incentive failed to inhibit the defensive blink, increases on the order of a hundred or even a thousand would not change this outcome. Consequently, defensive blinks are identified as reflexive as opposed to learned acts.

Winking is topographically similar to blinking, but its determinants (and hence its biology) are quite different. Winks that are reinforced by camaraderie or a shared secret tend to persist; those which are met by derision or disgust tend to fade. That is, rather weak rewards readily influence winking.
research will show that certain environmental events alter the brain in ways that increase the likelihood of violence (e.g., Miczek 1999). Thus, by the criterion of “biological basis,” all crimes would eventually become diseases. But this would lead to foolish and morally unacceptable policies. For instance, medical hospitalization for crimes such as murder and car theft would violate widely shared ideas of justice and, if consequences count, would prove counterproductive. Thus the idea that a disorder is a disease because it has a biological basis or because it entails brain changes is too crude a standard.

The second colloquial meaning of “disease” is that it is an involuntary, as opposed to a voluntary, disorder. This turns out to be a practical and scientifically defensible distinction. The mechanisms mediating voluntary and involuntary disorders differ, especially in regard to the influence of the central nervous system. Treatments for voluntary and involuntary disorders often differ (it would be cruel to punish tics or hallucinations, but it might be quite helpful and humane to provide corrective incentives for excessive television watchers). And social institutions for the remediation of voluntary and involuntary behavioral problems differ (e.g., “wise men” and “medicine men”). In short, the question of whether addiction is a voluntary or involuntary disorder matters (whereas there should be no question that addiction has a biological basis).

The idea that addicts take drugs involuntarily has been articulately argued by clinicians and researchers. Miller and Chappel (1991), psychiatrists, explained that addicts have a disease because they have lost control over drug use. Jellinek (1960), one of the first to systematically study alcoholism, defined alcoholism as a disease on the basis of loss of control over drinking. More recently, Leshner (1997), the head of the National Institute on Drug Abuse, claimed that repeated ingestion of addictive drugs “turns off” voluntary control of drug use.

However, the idea that a behavior is “out of control” does not seem to automatically qualify it as a disease. The excessive television watchers claimed that their behavior was out of control, and yet most observers would want to distinguish their problems from those of cancer patients, diabetics, schizophrenics, and even addicts. Also, the statement that addicts are unable to control their drug use is not accurate. For example, pack-a-day cigarette smokers meet the DSM crite-
ria for addiction, yet since the Surgeon General's report in 1964 on the health effects of smoking, some 50 million heavy smokers have quit using cigarettes (Schelling 1992). Moreover, most quit on their own, without medical help. If addiction is out-of-control drug use, how is this possible? On the other hand, it is no simple matter to quit an addictive drug. Thus the question of whether addiction is in fact involuntary drug use has remained controversial (e.g., Vuchinich 1996).

Part of the problem is that there are no widely agreed upon criteria for identifying voluntary acts.

Criteria for Identifying Voluntary Acts
The distinction between voluntary and involuntary acts pervades the discussion and analysis of behavior. In philosophy and political theory, the distinction is usually made along the lines of conscious intentions (Searle 1983) and/or freedom from authoritarian regimes. In scientific studies of behavior, the criteria have to do with the factors that influence behavior. Some behaviors are elicited by stimulus conditions and are relatively immune to reward and punishment (e.g., reflexes). Other behaviors have no specific eliciting conditions, but instead are learned; their frequency is a function of deprivation and relative reward and punishment. For example, consider the contrast between two simple, topographically similar behaviors, “eye blinks” and “winks.”

An eye blink is a “wired-in” behavior that is readily elicited by a directed force to the eye, such as a puff of air. The wiring admits specific eliciting stimuli, but provides few if any inputs for the influence of incentives or payoffs (contingencies). For instance, if a one-dollar incentive failed to inhibit the defensive blink, increases on the order of a hundred or even a thousand would not change this outcome. Consequently, defensive blinks are identified as reflexive as opposed to learned acts.

Winking is topographically similar to blinking, but its determinants (and hence its biology) are quite different. Winks that are reinforced by camaraderie or a shared secret tend to persist; those which are met by derision or disgust tend to fade. That is, rather weak rewards readily influence winking.
In everyday speech, we would say that the person intended to wink, could have done otherwise, or chose to wink. But in all these cases it is possible to make a simpler (and measurable) statement: “winking” is more susceptible to control by consequences than is “blinking.” The distinction is not between freedom of the will and determinism or between psychology and biology. Winks and blinks are equally determined and equally biological. However, the biology of the two behaviors differs. The anatomy of winking permits the influence of consequences, whereas the neural basis of blinks supports eliciting stimuli and admits little if any of the effects of contingent consequences.

**Voluntary as a Matter of Degree**

“Control by consequences” is not, of course, an all-or-none matter. Behaviors vary in the degree to which they are influenced by consequences, and behavioral syndromes differ in their mix of reflexive and learned components. For instance, the tics in Tourette’s syndrome or the hallucinations in schizophrenia do not seem to be readily influenced by contingencies, whereas the motor components of obsessive compulsive disorder do show some susceptibility to reward and punishment. Juvenile delinquency clearly has a biological component (given the age and gender correlations), but teenage acting out is probably more susceptible to consequences than are the motor components of obsessive compulsive disorder. In other words, one could construct a continuum in which disorders were ranked in regard to their susceptibility to the influence of various forms of persuasion (benefits, penalties, new information, and the like).

According to this analysis, the questions of whether addiction is a voluntary behavioral disorder or a disease can be rephrased as “To what extent will contingent rewards and punishments (broadly conceived) control drug consumption in addicts?” Again, the issue is not whether addiction has a biological basis or whether drugs change the brain. Rather, the issue is whether the biology of addiction results in a state such that drug consumption is no longer significantly influenced by its consequences. To answer this question, we need to test whether rewards, punishments, new information, and other forms of persuasion significantly influence drug consumption in addicts.

**“Compulsive” Drug Use as Ambivalent Drug Use**

Recall that the APA diagnostic manual defines addiction as the continued use of drugs despite their adverse consequences. According to the above discussion, this implies that addiction is involuntary. However, the authors of the manual left out a critical fact. A common feature of alcohol, nicotine, stimulants, and opiates is that they provide immediate positive consequences but delayed aversive consequences (withdrawal, penalties for intoxication, health risks, and so on). From this perspective, it is easy to see that addiction may well be a matter of contingencies: large and positive immediate consequences competing with large but delayed aversive consequences for the control of drug consumption. In these terms, addiction is ambivalence, with the decision to use drugs dependent on the temporal horizon. When the temporal frame of reference is relatively short, drug use is preferred; when the temporal frame of reference is relatively long, abstinence is preferred. Thus, what appears to be “compulsive” behavior may actually be a shift in preference, which, in turn, is a function of the temporal horizon at the moment of choice.

The difference between ambivalence and compulsion is important. By definition, a person who is ambivalent can be influenced by incentives, new information, changes in perspective, and all the sorts of acts that come under the general rubric “persuasion.” In contrast, compulsive behaviors are, by definition, not influenced by persuasion. For instance, if monetary incentives strongly influenced the frequency of hand washing in a patient with the diagnosis “obsessive compulsive disorder,” we would probably decide that the diagnosis was wrong. Note that this example suggests the practical and relative nature of these judgments. Assume for a moment that financial rewards can, in principle, have some effect on “compulsive” hand washing. The degree to which we are willing to call the act compulsive would be likely to depend on the amount of financial reward that was required for ameliorative change. If the amount was too large to have any practical application, then the available contingencies would be useless, and the hand washing would meet the criteria for compulsive behavior. Thus, to determine whether addiction entails ambivalence toward drug use or compulsive drug use, it is necessary to see if addicts can be persuaded to stop using drugs.
In everyday speech, we would say that the person intended to wink, could have done otherwise, or chose to wink. But in all these cases it is possible to make a simpler (and measurable) statement: “winking” is more susceptible to control by consequences than is “blinking.” The distinction is not between freedom of the will and determinism or between psychology and biology. Winks and blinks are equally determined and equally biological. However, the biology of the two behaviors differs. The anatomy of winking permits the influence of consequences, whereas the neural basis of blinks supports eliciting stimuli and admits little if any of the effects of contingent consequences.

Voluntary as a Matter of Degree

“Control by consequences” is not, of course, an all-or-none matter. Behaviors vary in the degree to which they are influenced by consequences, and behavioral syndromes differ in their mix of reflexive and learned components. For instance, the tics in Tourette’s syndrome or the hallucinations in schizophrenia do not seem to be readily influenced by contingencies, whereas the motor components of obsessive compulsive disorder do show some susceptibility to reward and punishment. Juvenile delinquency clearly has a biological component (given the age and gender correlations), but teenage acting out is probably more susceptible to consequences than are the motor components of obsessive compulsive disorder. In other words, one could construct a continuum in which disorders were ranked in regard to their susceptibility to the influence of various forms of persuasion (benefits, penalties, new information, and the like).

According to this analysis, the questions of whether addiction is a voluntary behavioral disorder or a disease can be rephrased as “To what extent will contingent rewards and punishments (broadly conceived) control drug consumption in addicts?” Again, the issue is not whether addiction has a biological basis or whether drugs change the brain. Rather, the issue is whether the biology of addiction results in a state such that drug consumption is no longer significantly influenced by its consequences. To answer this question, we need to test whether rewards, punishments, new information, and other forms of persuasion significantly influence drug consumption in addicts.

“Compulsive” Drug Use as Ambivalent Drug Use

Recall that the APA diagnostic manual defines addiction as the continued use of drugs despite their adverse consequences. According to the above discussion, this implies that addiction is involuntary. However, the authors of the manual left out a critical fact. A common feature of alcohol, nicotine, stimulants, and opiates is that they provide immediate positive consequences but delayed aversive consequences (withdrawal, penalties for intoxication, health risks, and so on). From this perspective, it is easy to see that addiction may well be a matter of contingencies: large and positive immediate consequences competing with large but delayed aversive consequences for the control of drug consumption. In these terms, addiction is ambivalence, with the decision to use drugs dependent on the temporal horizon. When the temporal frame of reference is relatively short, drug use is preferred; when the temporal frame of reference is relatively long, abstinence is preferred. Thus, what appears to be “compulsive” behavior may actually be a shift in preference, which, in turn, is a function of the temporal horizon at the moment of choice.

The difference between ambivalence and compulsion is important. By definition, a person who is ambivalent can be influenced by incentives, new information, changes in perspective, and all the sorts of acts that come under the general rubric “persuasion.” In contrast, compulsive behaviors are, by definition, not influenced by persuasion. For instance, if monetary incentives strongly influenced the frequency of hand washing in a patient with the diagnosis “obsessive compulsive disorder,” we would probably decide that the diagnosis was wrong. Note that this example suggests the practical and relative nature of these judgments. Assume for a moment that financial rewards can, in principle, have some effect on “compulsive” hand washing. The degree to which we are willing to call the act compulsive would be likely to depend on the amount of financial reward that was required for ameliorative change. If the amount was too large to have any practical application, then the available contingencies would be useless, and the hand washing would meet the criteria for compulsive behavior. Thus, to determine whether addiction entails ambivalence toward drug use or compulsive drug use, it is necessary to see if addicts can be persuaded to stop using drugs.
Laboratory Tests
There is a small but important series of experiments on whether contingencies can influence drug consumption in addicts. The subjects were long-term alcoholics who sought inpatient treatment for their drinking problems. The basic finding was that contingencies significantly modified drinking. For instance, in one of the more realistic studies, alcoholics were given free “priming” drinks and then offered incentives for not drinking further. Larger priming drinks were more likely to lead to a binge. However, for every priming dose there was an incentive that would promote self-control (see Cohen et al. 1971; Bigelow and Liebson 1972). It is not known whether this experience led to a lasting cure. But the question before us is whether it is possible to persuade an addict to stop using drugs. These data show that under experimental conditions, it is. (Also, according to the reports, the subjects were extremely alcoholic and would have met Jellinek’s criteria for loss of control over drinking.)

Natural Experiments: Vietnam
In the United States heroin has socially mediated aversive consequences. In Vietnam many of these aversive consequences were weakened or absent. Prohibitions against heroin were rarely enforced, and users apparently did not think of themselves as “junkies” or criminals. Thus, if addiction is a matter of choice, “simply” returning to the United States should have significantly reduced heroin addiction in the Vietnam enlists. This is exactly what happened, and the recovery rates were above 80 percent. Of course, the contingencies surrounding drug use were not the only differences between conditions in Vietnam and the United States. The men were at war in a hostile environment. However, the reports suggest that changes in the consequences of drug use played an important role. For example, the men cite heroin’s “sordid” reputation and the dangers involved in purchasing the drugs as reasons for quitting.

Historical Changes and Cohort Effects
The social consequences of illicit drug use varied markedly over the course of the twentieth century. Opiates and cocaine were legal in the United States until 1914; during the 1960s and 1970s there was wider acceptance of mind-altering drugs (including cocaine and heroin); and more recently there has been a concerted effort to dissociate addictive drugs from their earlier, positive connotations. Thus, if addiction is influenced by its consequences, there should be marked historical trends, even over rather short periods of time.

Figure 3.7, based on the NCS study, shows drug use and dependence given use (a measure of susceptibility to addiction) in men. The data are organized in terms of cohort and age. The oldest cohort was born prior to the end of World War II, between 1936 and 1945, and the youngest was born three decades later, between 1966 and 1975.

The cohort differences are large, especially for teenagers and especially relative to the oldest cohort. For example, for those born after World War II, drug use increased by more than a factor of four and dependence given use (susceptibility) also increased by a factor of four or more. Multiplying “use” by “susceptibility” should yield prevalence. By this logic, addiction increased sixteenfold for teenagers born between 1946 and 1955 and as much as fiftyfold for teenagers born between 1966 and 1975. The magnitudes are interesting. They suggest that historical factors may be among the most important determinants of addiction.

First-Person Accounts of Addiction
There is a very interesting literature on the addict’s experience of drug use and dependence. These writings include autobiographical pieces (e.g. Burroughs 1961) and ethnographic surveys that rely heavily on interviews with addicts (e.g. Courtwright, Joseph, and Des Jetteis 1989; Waldorf 1983). Many of these narratives follow a pattern that can be briefly stated, and the sequence of events, as told by those addicted, sheds light on the question of whether addicts can choose to stop using drugs.

In most cases, drug use begins in the teen years because it is “cool” and fun. There is an initial honeymoon period when aversive consequences are not apparent. This is followed by addiction (withdrawal symptoms and/or a shift in priorities with drug use becoming increasingly important) and feelings of regret in regard to how much time and money are being spent on drugs. Next comes a relatively long period during which health and welfare gradually decline. During this
Laboratory Tests
There is a small but important series of experiments on whether contingencies can influence drug consumption in addicts. The subjects were long-term alcoholics who sought inpatient treatment for their drinking problems. The basic finding was that contingencies significantly modified drinking. For instance, in one of the more realistic studies, alcoholics were given free "priming" drinks and then offered incentives for not drinking further. Larger priming drinks were more likely to lead to a binge. However, for every priming dose there was an incentive that would promote self-control (see Cohen et al. 1971; Bigelow and Liebson 1972). It is not known whether this experience led to a lasting cure. But the question before us is whether it is possible to persuade an addict to stop using drugs. These data show that under experimental conditions, it is. (Also, according to the reports, the subjects were extremely alcoholic and would have met Jellinek's criteria for loss of control over drinking.)

Natural Experiments: Vietnam
In the United States heroin has socially mediated aversive consequences. In Vietnam many of these aversive consequences were weakened or absent. Prohibitions against heroin were rarely enforced, and users apparently did not think of themselves as "junkies" or criminals. Thus, if addiction is a matter of choice, "simply" returning to the United States should have significantly reduced heroin addiction in the Vietnam enlists. This is exactly what happened, and the recovery rates were above 80 percent. Of course, the contingencies surrounding drug use were not the only differences between conditions in Vietnam and the United States. The men were at war in a hostile environment. However, the reports suggest that changes in the consequences of drug use played an important role. For example, the men cite heroin's "sordid" reputation and the dangers involved in purchasing the drugs as reasons for quitting.

Historical Changes and Cohort Effects
The social consequences of illicit drug use varied markedly over the course of the twentieth century. Opiates and cocaine were legal in the United States until 1914; during the 1960s and 1970s there was wider acceptance of mind-altering drugs (including cocaine and heroin); and more recently there has been a concerted effort to dissociate addictive drugs from their earlier, positive connotations. Thus, if addiction is influenced by its consequences, there should be marked historical trends, even over rather short periods of time.

Figure 3.7, based on the NCS study, shows drug use and dependence given use (a measure of susceptibility to addiction) in men. The data are organized in terms of cohort and age. The oldest cohort was born prior to the end of World War II, between 1936 and 1945, and the youngest was born three decades later, between 1966 and 1975.

The cohort differences are large, especially for teenagers and especially relative to the oldest cohort. For example, for those born after World War II, drug use increased by more than a factor of four and dependence given use (susceptibility) also increased by a factor of four or more. Multiplying "use" by "susceptibility" should yield prevalence. By this logic, addiction increased sixteenfold for teenagers born between 1946 and 1955 as much as fiftyfold for teenagers born between 1966 and 1975. The magnitudes are interesting. They suggest that historical factors may be among the most important determinants of addiction.

First-Person Accounts of Addiction
There is a very interesting literature on the addict's experience of drug use and dependence. These writings include autobiographical pieces (e.g. Burroughs 1961) and ethnographic surveys that rely heavily on interviews with addicts (e.g. Courtwright, Joseph, and Des Jelais 1989; Waldorf 1983). Many of these narratives follow a pattern that can be briefly stated, and the sequence of events, as told by those addicted, sheds light on the question of whether addicts can choose to stop using drugs.

In most cases, drug use begins in the teen years because it is "cool" and fun. There is an initial honeymoon period when aversive consequences are not apparent. This is followed by addiction (withdrawal symptoms and/or a shift in priorities with drug use becoming increasingly important) and feelings of regret in regard to how much time and money are being spent on drugs. Next comes a relatively long period during which health and welfare gradually decline. During this
period many addicts try to quit, but then relapse after brief periods of abstinence. However, the downward slide usually comes to an end (see recovery rates at beginning of this chapter), and often the end comes in the form of a dramatic realization that life has become intolerable. This realization is often accompanied by intense feelings of shame ("hitting bottom"), reframing of options (for example, "I do not want my kids to think of me as an addict"), and a conscious decision to stop using. Shame and reframing choices do not necessarily lead to changes in drug use, but this occurs often enough for "hitting bottom" to have become a widely recognized turning point in stories of addiction.

Recovery is not simply refraining from drug use. Success typically requires proactive measures, such as avoiding settings and friends associated with drug use and taking up new activities that can replace drug use (hobbies, new love relationships, exercise, and so on). In the initial stages of abstinence, the ex-user may suffer withdrawal symptoms and knows that "just one more hit" will significantly improve things. Moreover, abstinence increases the immediate, rewarding effects of the drug (because tolerance has been reversed). Gradually, though, new activities take over, and drug use loses its strong hold over behavior.

Recovering addicts struggle with their own desires, and success entails the ability to turn away from certain pleasure for an uncertain future. One could argue that abstinence following heavy drug use requires more effort and more self-control than refraining from initial use or never trying to stop. However, to call addiction a disease ignores these distinctions and, more generally, it fails to even acknowledge the efforts and accomplishments of those who do quit. In any case, the story of addiction is marked by choice points, inner struggle, and decisions—not by automatic, blind strivings, as implied by the disease model.

If Addiction Is Voluntary, Why Is It Hard to Quit?
Recovery from addiction is probably always a struggle, and for a significant minority it is a protracted battle (e.g. Vaillant 1992). This may appear to support the disease model. That is, if drug use is a matter of choice, why can't addicts simply quit? A detailed answer to this
period many addicts try to quit, but then relapse after brief periods of abstinence. However, the downward slide usually comes to an end (see recovery rates at beginning of this chapter), and often the end comes in the form of a dramatic realization that life has become intolerable. This realization is often accompanied by intense feelings of shame ("hitting bottom"), reframing of options (for example, "I do not want my kids to think of me as an addict"), and a conscious decision to stop using. Shame and reframing choices do not necessarily lead to changes in drug use, but this occurs often enough for "hitting bottom" to have become a widely recognized turning point in stories of addiction.

Recovery is not simply refraining from drug use. Success typically requires proactive measures, such as avoiding settings and friends associated with drug use and taking up new activities that can replace drug use (hobbies, new love relationships, exercise, and so on). In the initial stages of abstinence, the ex-user may suffer withdrawal symptoms and knows that "just one more hit" will significantly improve things. Moreover, abstinence increases the immediate, rewarding effects of the drug (because tolerance has been reversed). Gradually, though, new activities take over, and drug use loses its strong hold over behavior.

Recovering addicts struggle with their own desires, and success entails the ability to turn away from certain pleasure for an uncertain future. One could argue that abstinence following heavy drug use requires more effort and more self-control than refraining from initial use or never trying to stop. However, to call addiction a disease ignores these distinctions and, more generally, it fails to even acknowledge the efforts and accomplishments of those who do quit. In any case, the story of addiction is marked by choice points, inner struggle, and decisions—not by automatic, blind strivings, as implied by the disease model.

Is Addiction Is Voluntary, Why Is It Hard to Quit?
Recovery from addiction is probably always a struggle, and for a significant minority it is a protracted battle (e.g. Vaillant 1992). This may appear to support the disease model. That is, if drug use is a matter of choice, why can’t addicts simply quit? A detailed answer to this
question is provided in a paper that serves as the theoretical counterpart to this chapter (Heyman 1996a). The paper’s thesis is that addiction is a kind of evolutionary accident due to the poor fit between the mechanisms that guide choice and the reward structure of addictive drugs.

Addiction as Drug Preference
Experimental research shows that choice is not guided by rational bookkeeping principles, as often assumed in economic theory, but by myopic, psychological principles that reflect partial and distorted information about the competing alternatives (see Ainslie 1975; Herrnstein 1990; Heyman 1982; Heyman and Tanz 1995). Although myopic, these mechanisms are usually adequate, producing near optimal outcomes under normal conditions (e.g. Herrnstein and Prelec 1992; Heyman and Luce 1979). However, addictive drugs have unusual properties that sabotage optimal outcomes. They provide immediate positive effects and little if any satiation. Hence there are no “natural” brakes on consumption. But the delayed outcomes can be quite deleterious. This combination of properties implies a net loss for decision processes that are biased in favor of the immediate rather than the delayed value of a commodity. Also, addictive drugs have the pernicious property of undermining the reward value of competing activities (for example, withdrawal symptoms and intoxication do this; see Heyman 1996a, 1996b). This leads to a narrowing of the behavioral repertoire, with drug use crowding out other behaviors. In short, there is a mismatch, of evolutionary origin, between the rewarding properties of addictive drugs and the normal mechanisms of choice. These “bookkeeping” problems are compounded when there are also serious medical and psychiatric problems to contend with. Thus the nature of the mechanisms guiding choice, the drugs themselves, and the limited resources available to many if not most addicts conspire to make recovery from addiction a particularly difficult challenge.

This theory provides an alternative to the disease model. It says that addiction is a consequence of the normal mechanisms that guide choice operating in a context that reveals their limits—namely rewards that provide large up-front positive values in combination with hidden and delayed costs. This account also implies that to survive their own appetites, individuals require culturally transmitted practices that reinforce self-restraint (Prelec and Herrnstein 1991; Heyman 1996b). In turn, this point implies that cultural variation in the value of self-restraint and individual differences in the tendency to weight short- and long-term goals will influence addiction in important ways. Thus, like the disease model, the “evolutionary mismatch model” accounts for failures in self-command, but unlike the disease model, it views addiction as voluntary (which is to say, subject to contingencies).

The behavioral data favor the “evolutionary mismatch” model over the disease model. To be sure, the struggle for self-command is difficult, but the difficulties are ones that are amenable to contingencies. In contrast, the symptoms of diseases such as schizophrenia and Tourette’s syndrome have remained largely immune to the influences of insight, reward, and punishment. Put another way, the biological substrates of addiction are more susceptible to environmental influences than are the biological substrates of many if not all other psychiatric syndromes.

Summary and Discussion
This chapter has presented empirical findings relevant to the widely accepted claim that addiction is a “chronic, relapsing disease.” The data show that addiction typically remits, that it is the shortest-lasting psychiatric disorder, that it is the disorder most influenced by socially mediated consequences, and that addicts can curtail drug use when it is immediately beneficial to do so.

These conclusions are not inferences but simply a summary of the research findings. Thus the issue is not so much a question of how to interpret the findings, but whether to believe the findings.

Are the Findings Reliable and Valid?
Many of the graphs and statistics are based on interviews, a method subject to a host of obvious biases. The “data” depend on how questions are framed, how subjects are selected, and what the informants can remember and are willing to say about their own behavior. Moreover, the behavior in question is illegal. Thus it is reasonable to have serious doubts about the reliability and validity of the results.
question is provided in a paper that serves as the theoretical counterpart to this chapter (Heyman 1996a). The paper’s thesis is that addiction is a kind of evolutionary accident due to the poor fit between the mechanisms that guide choice and the reward structure of addictive drugs.

**Addiction as Drug Preference**

Experimental research shows that choice is not guided by rational bookkeeping principles, as often assumed in economic theory, but by myopic, psychological principles that reflect partial and distorted information about the competing alternatives (see Ainslie 1975; Herrnstein 1990; Heyman 1982; Heyman and Tanz 1995). Although myopic, these mechanisms are usually adequate, producing near optimal outcomes under normal conditions (e.g. Herrnstein and Prelec 1992; Heyman and Luce 1979). However, addictive drugs have unusual properties that sabotage optimal outcomes. They provide immediate positive effects and little if any satiation. Hence there are no “natural” brakes on consumption. But the delayed outcomes can be quite deleterious. This combination of properties implies a net loss for decision processes that are biased in favor of the immediate rather than the delayed value of a commodity. Also, addictive drugs have the pernicious property of undermining the reward value of competing activities (for example, withdrawal symptoms and intoxication do this; see Heyman 1996a, 1996b). This leads to a narrowing of the behavioral repertoire, with drug use crowding out other behaviors. In short, there is a mismatch, of evolutionary origin, between the rewarding properties of addictive drugs and the normal mechanisms of choice. These “bookkeeping” problems are compounded when there are also serious medical and psychiatric problems to contend with. Thus the nature of the mechanisms guiding choice, the drugs themselves, and the limited resources available to many if not most addicts conspire to make recovery from addiction a particularly difficult challenge.

This theory provides an alternative to the disease model. It says that addiction is a consequence of the normal mechanisms that guide choice operating in a context that reveals their limits—namely rewards that provide large up-front positive values in combination with hidden and delayed costs. This account also implies that to survive their own appetites, individuals require culturally transmitted practices that reinforce self-restraint (Prelec and Herrnstein 1991; Heyman 1996b). In turn, this point implies that cultural variation in the value of self-restraint and individual differences in the tendency to weight short- and long-term goals will influence addiction in important ways. Thus, like the disease model, the “evolutionary mismatch model” accounts for failures in self-command, but unlike the disease model, it views addiction as voluntary (which is to say, subject to contingencies).

The behavioral data favor the “evolutionary mismatch” model over the disease model. To be sure, the struggle for self-command is difficult, but the difficulties are ones that are amenable to contingencies. In contrast, the symptoms of diseases such as schizophrenia and Tourette’s syndrome have remained largely immune to the influences of insight, reward, and punishment. Put another way, the biological substrates of addiction are more susceptible to environmental influences than are the biological substrates of many if not all other psychiatric syndromes.

**Summary and Discussion**

This chapter has presented empirical findings relevant to the widely accepted claim that addiction is a “chronic, relapsing disease.” The data show that addiction typically remits, that it is the shortest-lasting psychiatric disorder, that it is the disorder most influenced by socially mediated consequences, and that addicts can curtail drug use when it is immediately beneficial to do so.

These conclusions are not inferences but simply a summary of the research findings. Thus the issue is not so much a question of how to interpret the findings, but whether to believe the findings.

**Are the Findings Reliable and Valid?**

Many of the graphs and statistics are based on interviews, a method subject to a host of obvious biases. The “data” depend on how questions are framed, how subjects are selected, and what the informants can remember and are willing to say about their own behavior. Moreover, the behavior in question is illegal. Thus it is reasonable to have serious doubts about the reliability and validity of the results.
Recall that the interviews were conducted by lay staff, not by professionally certified clinicians. This raises the possibility that as judged by trained clinicians, the diagnoses were wrongly assigned. However, in a study conducted with a clinic population, there was good agreement (Robins et al. 1982). Of the subjects who had been given a diagnosis by a psychiatrist, 75 percent were given the same diagnosis by the lay researchers. Conversely, for those subjects who were not given a diagnosis by the psychiatrist, 94 percent were also not given a diagnosis by the researchers. Some disorders were easier to agree upon than others. For substance abuse and dependence, 85 percent of cases identified by psychiatrists were also identified by the lay researchers.

But these were current or recent clinic patients, while the epidemiological research was conducted primarily with randomly selected individuals from major American metropolitan areas. The clinic subjects might be easier to diagnose and thus give a misleading account of reliability. Accordingly, the ECA researchers ran a second reliability study, this time with subjects selected independently of treatment history (Helzer et al. 1985). The subjects were first interviewed by a researcher, then by a psychiatrist. The median interval was six weeks so that there was time for symptoms to worsen or improve. The overall level of agreement on whether a subject met the criteria for a substance-use disorder was more than 90 percent. Of those identified by the psychiatrists as having a drug-use disorder, two-thirds were similarly identified by the researchers. Of those identified by the psychiatrists as not suffering from drug dependence or abuse, more than 90 percent were also not given a diagnosis by the researchers.

But reliable diagnoses do not guarantee meaningful results. The informants could systematically lie, or, more likely, misrepresent the past when it was inconsistent with their current situation or with what they thought the interviewer wanted to hear.

**Truthfulness**

The use of illicit drugs is punishable, and many of those who use illicit drugs have a history of committing illegal acts. Thus there are good reasons to suppose that informants would not be truthful about drug use, especially current drug use. This could explain the high remission rates (honesty about the past but not about the present).

In some settings, those who use illegal drugs often deny that they do so. For instance, in several studies pregnant women who tested positive for cocaine typically denied that they were using illegal drugs, including cocaine (Brownsberger 1997). Moreover, they continued to deny drug use even when confronted with positive urine and meconium samples (e.g. Ostrea et al. 1992; Nair, Rothblum, and Hebel 1994). However, the women were concerned that if they admitted to drug use they might lose their medical assistance or even have their children taken away.

The perceived risks for the interviewees in the studies reviewed in this chapter were different. The researchers went to great lengths to reassure the informants that their answers would not be turned over to the legal authorities, that their anonymity would be preserved, and that there would be no negative consequences for participating in the interviews. Objective evidence, such as urine samples, indicates that these assurances established a setting in which truthful answers were the norm. In the Vietnam studies (Robins et al. 1980), 97 percent of those who had an official record of narcotic use while in the army reported this to the interviewers. Urine tests on this population, done after the interview, did not reveal higher rates of current use than did self-report. This is significant in that the veterans did not know they were going to be tested, and the interviewer did not know the informant's history. As noted above, there was a similar level of concordance between official records and the interview data in the St. Louis study (Robins and Murphy 1967).

**Validity**

Although tests indicate that the ECA researchers reliably classified their informants, these tests do not guarantee that the estimates of relapse rates and duration are valid. For instance, there were no independent checks of when a disorder began or ended. Longitudinal studies could provide more certain information, but such data have not been collected. Thus it may be reasonable to assume that the duration and remission data are inherently flawed.

The basic question is whether the behavioral results shown in the various graphs reflect the researchers' methods or the subjects' behavior. If the findings reflect methods, then they should vary across stud-
Recall that the interviews were conducted by lay staff, not by professionally certified clinicians. This raises the possibility that as judged by trained clinicians, the diagnoses were wrongly assigned. However, in a study conducted with a clinic population, there was good agreement (Robins et al. 1982). Of the subjects who had been given a diagnosis by a psychiatrist, 75 percent were given the same diagnosis by the lay researchers. Conversely, for those subjects who were not given a diagnosis by the psychiatrist, 94 percent were also not given a diagnosis by the researchers. Some disorders were easier to agree upon than others. For substance abuse and dependence, 85 percent of cases identified by psychiatrists were also identified by the lay researchers.

But these were current or recent clinic patients, while the epidemiological research was conducted primarily with randomly selected individuals from major American metropolitan areas. The clinic subjects might be easier to diagnose and thus give a misleading account of reliability. Accordingly, the ECA researchers ran a second reliability study, this time with subjects selected independently of treatment history (Helzer et al. 1985). The subjects were first interviewed by a researcher, then by a psychiatrist. The median interval was six weeks so that there was time for symptoms to worsen or improve. The overall level of agreement on whether a subject met the criteria for a substance-use disorder was more than 90 percent. Of those identified by the psychiatrists as having a drug-use disorder, two-thirds were similarly identified by the researchers. Of those identified by the psychiatrists as not suffering from drug dependence or abuse, more than 90 percent were also not given a diagnosis by the researchers.

But reliable diagnoses do not guarantee meaningful results. The informants could systematically lie, or, more likely, misrepresent the past when it was inconsistent with their current situation or with what they thought the interviewer wanted to hear.

Truthfulness
The use of illicit drugs is punishable, and many of those who use illicit drugs have a history of committing illegal acts. Thus there are good reasons to suppose that informants would not be truthful about drug use, especially current drug use. This could explain the high remission rates (honesty about the past but not about the present).

In some settings, those who use illegal drugs often deny that they do so. For instance, in several studies pregnant women who tested positive for cocaine typically denied that they were using illegal drugs, including cocaine (Brownsberger 1997). Moreover, they continued to deny drug use even when confronted with positive urine and meconium samples (e.g. Ostrea et al. 1992; Nair, Rothblum, and Hebel 1994). However, the women were concerned that if they admitted to drug use they might lose their medical assistance or even have their children taken away.

The perceived risks for the interviewees in the studies reviewed in this chapter were different. The researchers went to great lengths to assure the informants that their answers would not be turned over to the legal authorities, that their anonymity would be preserved, and that there would be no negative consequences for participating in the interviews. Objective evidence, such as urine samples, indicates that these assurances established a setting in which truthful answers were the norm. In the Vietnam studies (Robins et al. 1980), 97 percent of those who had an official record of narcotic use while in the army reported this to the interviewers. Urine tests on this population, done after the interview, did not reveal higher rates of current use than did self-report. This is significant in that the veterans did not know they were going to be tested, and the interviewer did not know the informant's history. As noted above, there was a similar level of concordance between official records and the interview data in the St. Louis study (Robins and Murphy 1967).

Validity
Although tests indicate that the ECA researchers reliably classified their informants, these tests do not guarantee that the estimates of relapse rates and duration are valid. For instance, there were no independent checks of when a disorder began or ended. Longitudinal studies could provide more certain information, but such data have not been collected. Thus it may be reasonable to assume that the duration and remission data are inherently flawed.

The basic question is whether the behavioral results shown in the various graphs reflect the researchers' methods or the subjects' behavior. If the findings reflect methods, then they should vary across stud-
ies; if the findings reflect the nature of addiction, each type of study should tell essentially the same story.

There was, I believe, a considerable degree of consistency across studies. The research on heroin addiction in Vietnam and St. Louis and the two national surveys led to similar estimates of remission rates and duration. The laboratory findings and the first-person narratives told the same story: depending on the consequences, addicts can choose to abstain. There is even a rough quantitative agreement in the various accounts. A minority of the Vietnam veterans (6 percent) sought clinic-based treatment. Their relapse rates (67 percent) were within the range reported in the clinic-based studies. This suggests that the high recovery rates in the Vietnam sample were not unusual, but exactly what would be seen in a prospective study conducted with addicts who do not seek treatment and who became addicted in the United States.

Treatment

The data and theory presented in this chapter indicate that treatment and policy that devalue the benefits of drug use, that serve to create and strengthen activities that can effectively compete with drugs, and that teach addicts to reframe their options in terms of their full costs and benefits will prove effective. The chapters by Sally L. Satel and George E. Vaillant come to similar conclusions in regard to consequences and alternative activities, although they express these ideas somewhat differently.

Although this chapter has emphasized the role of environmental factors and individual differences in recovery, the emphasis on the biological basis of voluntary behavior implies that pharmacological treatments are potentially valuable resources for reducing drug use. The story of methadone provides a useful example.

Methadone pretreatment attenuates the effects of heroin. The rush and subsequent high are greatly diminished, and as a result preference for heroin decreases. However, whether a decrease in heroin’s hedonic effects leads to a resolved recovery appears to depend on additional factors. For instance, Dole and Nyswander’s (1967) original methadone programs were highly successful, whereas recent reports show that methadone patients frequently test positive for cocaine, alcohol, and other drugs (e.g. Silverman et al. 1996). One possible explanation for the initial high success rates is that the majority of Dole and Nyswander’s patients quickly established interests that could compete with heroin. For example, within six months of starting treatment, about 75 percent were employed. Also, Dole and Nyswander excluded subjects who were least likely to develop alternatives to drug use. Possibly other addictive drugs have practical pharmacological antidotes. This research should be pursued, as pharmacological methods are in principle a highly efficient technique for modifying preferences.

This chapter focused on research in which there was an effort to study representative drug users and not just those who were treated in clinics. When addiction is looked at in this way, it is not a chronic disease but a matter of ambivalent drug use. In the short term, the drug is the better option; in the long term, it is not. The implication for treatment and policy is that addicts can be persuaded to stop using drugs. Moreover, as there is a biology of choice, the techniques of persuasion may include pharmacological agents as well as social ones. In sum, addiction is a malleable disorder, and methods for producing change are well within our reach.

References

Preparation of this chapter was supported by a grant from the Russell Sage Foundation.


ies; if the findings reflect the nature of addiction, each type of study should tell essentially the same story.

There was, I believe, a considerable degree of consistency across studies. The research on heroin addiction in Vietnam and St. Louis and the two national surveys led to similar estimates of remission rates and duration. The laboratory findings and the first-person narratives told the same story: depending on the consequences, addicts can choose to abstain. There is even a rough quantitative agreement in the various accounts. A minority of the Vietnam veterans (6 percent) sought clinic-based treatment. Their relapse rates (67 percent) were within the range reported in the clinic-based studies. This suggests that the high recovery rates in the Vietnam sample were not unusual, but exactly what would be seen in a prospective study conducted with addicts who do not seek treatment and who became addicted in the United States.

**Treatment**

The data and theory presented in this chapter indicate that treatment and policy that devalue the benefits of drug use, that serve to create and strengthen activities that can effectively compete with drugs, and that teach addicts to reframe their options in terms of their full costs and benefits will prove effective. The chapters by Sally L. Satel and George E. Vaillant come to similar conclusions in regard to consequences and alternative activities, although they express these ideas somewhat differently.

Although this chapter has emphasized the role of environmental factors and individual differences in recovery, the emphasis on the biological basis of voluntary behavior implies that pharmacological treatments are potentially valuable resources for reducing drug use. The story of methadone provides a useful example.

Methadone pretreatment attenuates the effects of heroin. The rush and subsequent high are greatly diminished, and as a result preference for heroin decreases. However, whether a decrease in heroin’s hedonic effects leads to a resolved recovery appears to depend on additional factors. For instance, Dole and Nyswander’s (1967) original methadone programs were highly successful, whereas recent reports show that methadone patients frequently test positive for cocaine, alcohol, and other drugs (e.g. Silverman et al. 1996). One possible explanation for the initial high success rates is that the majority of Dole and Nyswander’s patients quickly established interests that could compete with heroin. For example, within six months of starting treatment, about 75 percent were employed. Also, Dole and Nyswander excluded subjects who were least likely to develop alternatives to drug use. Possibly other addictive drugs have practical pharmacological antidotes. This research should be pursued, as pharmacological methods are in principle a highly efficient technique for modifying preferences.

This chapter focused on research in which there was an effort to study representative drug users and not just those who were treated in clinics. When addiction is looked at in this way, it is not a chronic disease but a matter of ambivalent drug use. In the short term, the drug is the better option; in the long term, it is not. The implication for treatment and policy is that addicts can be persuaded to stop using drugs. Moreover, as there is a biology of choice, the techniques of persuasion may include pharmacological agents as well as social ones. In sum, addiction is a malleable disorder, and methods for producing change are well within our reach.

**References**

Preparation of this chapter was supported by a grant from the Russell Sage Foundation.


brecher, e. 1972. licit and illicit drugs. boston: little, brown.

brown, r., and r. j. herrnstein. 1975. psychology. boston: little, brown.


goelman, d. 1990. "how viewers grow addicted to television." new york times, oct. 16, c1, c8.


helzer, j. e., e. l. spitzagel, and l. t. mcEvoy. 1987. "the predictive validity of lay diagnoses in the general population." archives of general psychiatry 44: 1069-74.


heyman, g. m. 1982. "is time allocation elicited behavior?" in m. commons, r. j. herrnstein, and h. rachlin, eds., quantitative analyses of behavior, vol. 2: matching and maximizing accounts (cambridge, mass.: ballinger), 459-490.


