Proximate Cause and Traumatic Neurosis

NATHAN T. SIDLEY, MD

Causation is frequently an issue to forensic psychiatrists in negligence cases involving psychic trauma. Typically a plaintiff asserts that a given mentally upsetting event occurred and that event "caused" subsequent neurotic symptoms (that is, a "traumatic neurosis"). For example, a plaintiff may allege that the defendant did something that frightened the plaintiff and that the fright caused harmful anxiety symptoms.

To prove causation in relation to a specific case, one logically has to show that the case is an example of a general causal law. A general law is, in effect, a statement that whenever a member of a given class of individuals is exposed to a given class of conditions, a given outcome occurs. For example, "Whenever water is heated to 100° under standard conditions, it boils," is a causal law.

To establish such a causal law, however, it also must be shown that when a member of the given class is not exposed to the given condition, the given outcome does not occur.*

Demonstrating specific causation in a specific case presupposes that a general law of causation has been established. Indeed, the absence of a general law is a great handicap in trying to establish that in a specific situation a given event was the cause of a given subsequent event. (The absence of a causal law may vary from the almost complete absence of such a law to the presence of a causal relationship that is only partial or that is statistical. From the standpoint of science, a causal relationship valid in 50 percent of cases is surely better than one valid in only 20 percent of cases, but neither is as helpful as desired where a determination must be made in one legal case.)

Under negligence law, causation is an issue in relation to the concept of "proximate cause," which is the same as saying, "legally responsible cause." In a negligence suit, the plaintiff alleges the defendant committed an act defined by the law as negligence and that the negligent act caused harm to the plaintiff. The law recognizes that events generally have multiple causes, some that go far back in history. (In a broad sense the landing of the Pilgrims is a cause of the Korean War.) The critical question in negligence, however, is whether the act of negligence set into play a train of events such that the ultimate harm was foreseeable. If in a negligence case some independent cause intervenes between the negligence and the ultimate harm, the negligence ordinarily is not regarded as the proximate cause of the harm.

For example, if Jones, walking on a sidewalk in front of Smith's house,

---

Dr. Sidley is Director, Court Clinic, Fourth District Court of Eastern Middlesex, Woburn, MA 01801.

*Of course, there may be multiple possible causes for a given outcome. The critical test to show that *a* causes *b* must be a situation in which some comparable individuals are exposed to *a* while others are not. One can only prove that *a* causes *b* if the situations without exposure to *a* do not result in *b*.

trips on a roller skate and twists his ankle, Smith is likely to be considered negligent and to be found liable for the damage. If after twisting his ankle, Jones sits to rest on Smith's steps and a tree-branch above him is hit by lightning, falls, and breaks Jones's back, Smith is not likely to be liable for Jones's broken back, because the lightning represents an independent, unforeseeable cause. However, if Jones sits on Smith's steps and a neighbor's dog running after a ball bowls him over and breaks his arm, Smith is likely to be held liable for the broken arm as well as for the twisted ankle, since, at least to a certain extent, that is a foreseeable consequence of a person's getting injured from tripping on a roller skate left on a sidewalk. Of course, the fact finder in a case of negligence must determine whether any given act of negligence is a proximate cause of the ultimate harm.

Legally the concept of proximate cause is simplified as the "but for" test. That is, an act of negligence is considered as a proximate cause if the ultimate harm would not have occurred but for the act of negligence. That test oversimplifies causation, however, and does not completely comprehend proximate cause. (In the example just presented, neither of Jones's possible injuries would have occurred but for the roller skate negligently left on the sidewalk. However, if lightning hits the tree, the roller skate negligence is not the proximate cause of his broken back; if the dog breaks Jones's arm, the roller skate negligence probably would be regarded as the proximate cause of the arm injury.) An act of negligence cannot be a proximate cause unless it passes the "but for" test, but merely passing the test does not establish the act as a proximate cause if a subsequent unforeseeable independent event causes the damage. Indeed, the degree of foreseeability and the relative contributions to harm produced by the negligent act and by subsequent events may become important issues in litigation.

Note that if there are multiple causes to a harm, if the alleged negligence is a principal or a substantial cause giving rise to the foreseeable harm, that is enough for it to be a proximate cause. There can be more than one proximate cause to a harm if there are several negligent parties or acts; some proximate causes may occur at the same time; others may come in sequence. Again using the example, if Jones's ankle were negligently treated in a hospital, one proximate cause of harm is Smith's negligence, while another proximate cause is the negligence of the treater. The wisdom of the fact finder must ultimately be involved in order to resolve the causal and the liability issues,† and complex cases can be a real challenge.

The state of knowledge of general causal relationships is at base the state of the art of science. In psychiatry and the field of causation of neurosis, the state of science is primitive; in the field of traumatic neurosis, it probably will always be primitive because of some basic difficulties.

†"An act or omission is the proximate cause of a loss where there is no intervening, independent, culpable and controlling cause, but the application of this rule, and similar rules, is not without difficulty." (25 C.J.S.651)

The first difficulty in forming general causal laws in traumatic neurosis is that of adequately defining individuals before they are exposed to traumas. People are so different on so many dimensions of characterization, that there is no single individual who is identical to any other individual. That problem alone compromises any kind of general conclusions that may be sought with respect to causation of neurosis. (Whatever difficulties there are in classifying persons, doing so retrospectively after the fact of trauma multiplies the problems.)

Second, differences among traumas on characteristics that might be important in causing neurosis also compromise any general laws that might relate to traumatic neurosis. (That includes the lack of identity of trauma impact. Even if individuals were exposed to the same trauma, such as a natural disaster, the significance or impact of the trauma may differ among the individuals exposed.)

Third, outcome situations are so varied as not to be adequately comparable. A neurosis in one person may be very unlike a neurosis in another individual.

Because experiments cannot be conducted with human beings to try to establish causal laws relating neurosis to trauma, such causal relationships as can be ascertained must be inferred from post facto observations of people exposed to traumatic situations. First, the "before" and the "after" states in exposed persons are examined; second, mental changes over time in people exposed to traumas are contrasted with time changes in people not exposed to traumas.

The best causal conclusion that can be arrived at seems to be a statistical probability that people exposed to traumas subsequently develop neurotic symptoms to a greater extent than people not exposed to traumas. Causation is not at all strict. Another way of saying the same thing is that on the average, exposure to trauma increases the risk of neurotic symptoms.

That broad approach seems to have been taken by the DSM-III. There the notion is that a "traumatic event," to be regarded as such, must be "generally outside the range of human experience," and that "the stressor... would evoke significant symptoms of distress in most people... Some stressors frequently produce the disorder... others produce it only occasionally." The vagueness and subjectivity of the DSM-III definition reflect the limitations of the science of stress reactions.

Does that mean that traumatic neurosis should never be counted in assessing damages in negligence litigation? Surely an argument can be made that it should not. Indeed for many years that was judicial policy, seemingly because of the difficulty in distinguishing bona fide mental damages (resulting from an act of negligence) from alleged mental damages that actually sprang from some other source. It may be better policy to exclude from recovery meritorious claims of mental damage than to allow the possibility that many undocumentable and undeserving claims would be rewarded.

An alternative approach is to allow for legal consideration the possibility

of mental damages resulting from an act of negligence. Fact finders, on a case-by-case basis, would try to make decisions about causation on subjective grounds. These subjective grounds would require a judgment of the severity of the trauma in relation to the vulnerability of the individual,‡ a procedure that involves commonsense inference from two concepts: The first is that people often seem to be adversely affected by stress. The second notion is that the reaction to stress seems generally to be proportional to the severity of the stress and to the vulnerability of the individual. (Except for some scientific measures of stress, however, each fact finder is left to create his own ad hoc stress-gradation scale.)

When the law gives such latitude to fact finders, the skills of plaintiff's and of defendant's advocates may be of excessive importance in bringing about a decision in a negligence case.

Another conceivable policy is to have arbitrary limits placed on the definition of stress and/or the allowable damages for mental pain and suffering. The situation would be analogous to the definitions in an insurance policy. For example, mental damages might be allowed only if an individual were physically injured in a negligence situation to a degree requiring two days' hospitalization; similarly a policy might be that if mental damages were claimed because of the death of someone else, that person would have to be a first degree relative or a spouse. Alternatively, limits of monetary damages could be prescribed in relation to different types of cases, so that damages for mental suffering might not be allowed to exceed, say, $10,000.

Such a system would be very complex and also very arbitrary, though perhaps on the average it might be more equitable in distributing damage payments than would the previously mentioned system. Of course, there is no good way, absent a scientific backing, of establishing whether any one method is any better or any more "just"§ than other methods. Policy must be made in the absence of good information as to its effects on society. Of course, policy issues must be taken up by legislatures and courts.

At any event, in the presently predominant system fact finders make their awards subjectively. If proximate cause is an issue in a case of traumatic neurosis, the fact finders weigh the evidence, and, other considerations aside, if it appears to them more likely than not that the negligence was a proximate cause of the posttraumatic neurotic symptoms, they rule for the plaintiff.

What can the psychiatrist do when he or she is called to make an evaluation of the cause of a mental disorder in relation to a traumatic event, when well-founded principles of decision making do not exist? The assumption is that the context is a jurisdiction in which decisions are made on a

‡Vulnerability is important in causation in the sense that a minor trauma might be an important cause of neurosis in a highly vulnerable individual, even though the same trauma would be an unlikely cause of neurosis in a less vulnerable person. Judgments of vulnerability and severity are, of course, highly subjective.

§"Justice" being, for practical purposes, policies that contribute to social harmony.

case-by-case basis.

Though one cannot provide hard rules, there are some general suggestions:

First, in the absence of a scientifically precise and reliable cause-and-effect relationship between trauma and neurosis, integrity demands that one acknowledge that one cannot express an opinion that is demonstrably valid. One can only offer informed speculation, which is, at best, perhaps a cut above reasoned lay opinion based on common sense. Of course, informed speculation can be very helpful to a lay jury as well as being influential in their consideration of an issue.

Second, in forming his or her opinion, one is advised to evaluate the case in terms of the logic of causation. That involves considering the individual, the stress, and the succeeding situation.

The initial issue is trying to determine whether the alleged harm would have, or might well have, occurred to the individual had he or she not been exposed to the trauma. If, for example, it is alleged that symptoms occurred following a trauma, but history indicates that comparable symptoms occurred at different times in the person's life in the absence of trauma, that casts doubt on a hypothesis that symptoms occurred because of exposure to the postulated trauma.

If, in addition to the alleged trauma on which the negligence case is based, there were other traumatic situations to which the individual also was exposed prior to the occurrence of symptoms, that too renders it difficult to determine with much confidence which trauma, if any, produced the symptoms. However, one can formulate reasonable hypotheses based on the degree of trauma, and so on, to try to resolve the legal causation issue.

In that context it is important to define accurately the trauma to which the individual was exposed and how the trauma impacted on him or her. It is not always apparent to fact finders that mere exposure to a traumatic situation does not in itself determine the effects on a person. One requires additional information to formulate some kind of quasi-quantitative appraisal of the degree of trauma. It is appropriate to investigate what specifically were the external facts to which the individual was exposed; what was his or her perception of those facts; how did he or she anticipate the effect, and what were his or her emotional responses to those perceived anticipations. The degree to which the trauma itself was foreseen by the person (and could thus have been defended against) may be another important issue.

In addition, the neurosis alleged to have resulted from the traumatic exposure must be well defined. Some types of symptoms are more common than others following trauma exposure. The degree to which the plaintiff's symptoms conform to expected symptoms is important to evaluate in trying to ascertain to what extent a trauma might have been the cause of a neurosis.

The timing of onset of symptoms is a last issue that might be investigated. While common sense suggests that symptoms should start immediately after a trauma (at least immediately after the effects are

psychologically perceived) and that the individual should get better with time as mental recovery processes occur, there are some conceptual problems with that view: (1) It may take significant time for a person to become aware of the full psychological impact of a trauma. (2) There may be overriding mental inhibitory factors that prevent the expression of symptoms. (3) There may be a threshold below which a trauma reaction might not be manifested. The individual might, for example, be left in a state of increased vulnerability so that subsequent exposure to a minor stress, of either an external or an internal physiological nature, which ordinarily would give no trouble, would precipitate severe symptoms or would result in the manifest expression of a previously latent neurosis.

The first issue, that it might take time for the fully psychological meaning to sink in, still implies an onset of neurosis fairly soon after the trauma, when the effects of the trauma are fresh and the individual is still highly motivated to deal with its effects. Once the trauma is out of the immediate focus of the mind, it appears reasonable to assume that the initial gross impression has been coped with and that partial effects of the trauma, subsequently experienced by the person retrospectively, are less stressful than the initial exposure; absent some other factor that compromises adaptation, they also can be coped with by the individual. After exposure, once a person has resumed a more or less usual pretrauma comfort level, he or she has coped with the initial trauma impact, and barring further exposure to new stress, additional perceptions of the trauma are also likely to be adapted to.

If the second possibility occurs and the individual avoids neurosis by great inhibitory effort, that effort would be expected to begin immediately with the trauma, and the person would be observed one way or another to be an individual under tension.

If the third potentiality ensues, and subsequent to trauma the individual is rendered foreseeably vulnerable to the ordinary stresses of life, a defendant might still be held liable for a late sequel of a negligent act. (Ordinarily, though, a defendant is not held liable merely for a plaintiff's increased vulnerability resulting from the defendent's negligence.) Because other events supervene between the negligent act and the symptoms, though, the logical problems involved in sorting out the relative contributions of possible vulnerability and foreseeability factors are at least formidable, if not impossible, to resolve adequately for purposes of the individual case.

In the view of the writer, a time delay between trauma and symptoms with a period of well-being between this implies one or both of the following: (1) There is a fresh stress, which somehow either activates latent problems caused by the former stress or independently causes the symptoms. (2) There is some internal change in threshold resulting in the uncontrolled expression of issues previously held in check.

From the standpoint of proximate cause, the implication in either case is that there is an intervening or an independent cause operative in producing the ultimate symptoms. The initial trauma can possibly be regarded as a cause of an increased susceptibility to neurosis. But the fact finder would

have to determine to what extent the defendant should be liable for the neurosis, and that would in part be dependent on the fact finder's perception of the relative importance of the original trauma and the subsequent stresses. Generally, the longer the delay between the trauma and the onset of symptoms, and the greater the apparent precipitating stress in relation to the symptoms, the less likely will the defendant be to be found liable for the neurosis.

Example

The case example of alleged traumatic neurosis presented illustrates forensic psychiatric contributions that can be made in a lawsuit in attending to the logic of proximate cause when considering the "facts," that is, case observations, made in the course of evaluation. The plaintiff was a 62-year-old truck driver. While driving his tractor-trailor truck on a snowy, slippery day, he became involved in an accident. A pickup truck operated by the defendant, driving behind a small car, collided with the car, pushing it into the plaintiff's lane of traffic. The plaintiff's rig then collided with the car and injured the driver, who died a few days later. The plaintiff sued the defendant for negligence, claiming that psychiatric difficulties the plaintiff subsequently underwent, as well as loss of income, pain-and-suffering, and so forth, were caused by the defendant's negligence, which allegedly had brought about the accident.

Three years after the accident, the plaintiff was hospitalized for what were diagnosed as depression and a conversion contracture of his left arm. He was obsessed with thoughts of the accident and cried whenever it was mentioned. He was hospitalized on the anniversary of the accident.

In the hospital he improved rapidly on treatment with psychotherapy and antidepressants. Subsequently, following his discharge from the hospital, he discontinued medication. He relapsed six months later, a time which happened to be a month after he returned to work. He did not reenter the hospital, and he never worked again. He did undertake a few efforts at treatment, but they were unsuccessful. (He lived in an area with few psychiatric facilities.)

The plaintiff's psychiatrist, who had managed the case in the hospital, deemed the neurosis to be an "anniversary reaction" to the trauma of the accident, in effect, a late sequel of the reaction to the event, particularly to the death of the driver of the car.

The view of the defense psychiatrist was somewhat different, emphasizing other features of the prior history: The plaintiff had not shown any real psychic reaction at the time of the accident. He had not been in any danger, and he had not been personally afraid at the time of the collision. After the accident he had continued to work in his usual manner, and he had shown no apparent personality changes.

He was seemingly an accident-prone person, having undergone several accidents in the course of his employment with his company. (In fact, he had been criticized several times for not having reported on time accidents that

were allegedly workmen's compensable.) His illness absenteeism record from work was greater than average.

His medical records also revealed significant discrepancies in his accounts when histories of the same medical incidents had been taken by different individuals.

He missed considerable work because of various physical illnesses that had occurred in the year prior to his hospitalization. Five months before his hospitalization he had injured his left elbow and had required some painful treatments.

A month after his elbow injury he was reassigned at work. His reassignment meant a drastic change in his work conditions, including operating out of a location distant from his home. The change necessitated separation from his wife during the working days of the week. Additional expenses and new assignments in unfamiliar territory also were required. In addition there were many unfamiliar coworkers and unfamiliar outsiders to contend with, at times on an intimate basis.

Three days after beginning the new assignment he reported to his physician, complaining of nervousness and inability to work and appearing anxious and depressed. He received a month's medical leave.

Three days after returning to work following the medical leave, he again reported to his physician, again complaining of nervousness and inability to work. He exhibited a flexion contracture of his left arm, a symptom he also regarded as preventing him from working. He was sent to an orthopedist, who referred him to a psychiatric consultant, whence he was admitted to the hospital on the anniversary of the accident.

It should be noted that the facts used by plaintiff's psychiatrist were fundamentally those presented by the plaintiff during his admission to the hospital. The psychiatrist evaluated the case for treatment purposes and did not undertake the kind of evaluation he might have for legal purposes. There was a broader range of information available to the defendant's psychiatrist because it was sought in the context of the legal case and because defense counsel had undertaken an exhaustive investigation of the case. (Note that by request of the attorneys the psychiatrists did not confer about the case.)

Discussion

One's approach to determining proximate cause of traumatic neurosis is generally most effective when performed in a systematic manner. That calls for evaluation of the previous life of the individual, the nature of the alleged trauma, the possible presence of alternative symptom-producing factors, and the appropriateness of symptoms alleged to have resulted from the trauma, both as to the nature of the symptoms and as to their timing.

In the case mentioned the alleged trauma, the accident with the death of the car driver, was an event which perhaps fulfilled DSM III criteria in being outside the ordinary range of human experience. Because the plaintiff realistically felt himself to be in no danger, however, it is very questionable that the accident was a stressor that would evoke significant symptoms of

distress in most people.

The history indicated no prior symptoms of depression nor of conversion problems of the extremities, but there was a suggestion of much job absenteeism secondary to illness, and for some medical problems more time was taken off work than is ordinarily the case. There was also an implication of a greater orientation toward compensation for illness than is usual.

There did not seem to be other potential symptom-causing stresses in operation at the time of the initial accident. The timing of symptoms is important, however, and the occurrence of symptoms after a three-year interval without difficulties implies that if the incident were involved as a cause of the subsequent symptoms, it could only be a partial cause and that there must have been other causes involved as well.¶

The subsequent trauma in the man's life, occurring in relation to his work and having a major impact on his life, did, of course, take place in an immediate time relationship to the development of his symptoms.

The symptoms themselves would have been surprising in relation to the alleged causative incident. Although depression is appropriate to loss (and the death of the other driver might be regarded as a loss), yet depression to the extent of hospitalization is uncommon even after the death of an important figure like a spouse, let alone the death of a complete stranger, as in the case at hand. There is also nothing in the nature of a motor vehicle accident, such as in this case, that would be likely to lead to a conversion flexion contracture of an arm, especially beginning almost three years later. Such a symptom would be far more expected in relation to the stress of a major undesired work change and might be helpful in enabling the person to avoid the work situation.

Thus in this case a conclusion was drawn by the defense psychiatrist that the accident ought not reasonably to be regarded as the proximate cause of the neurosis that followed. Insofar as the disorder should be regarded as caused by events occurring in the patient's experience, it was considered to be far more reasonable to regard the work change as the stress that proximately caused the symptoms and the motor vehicle accident as a factor that only influenced the form of the symptoms in providing a focus for the expression of the depressive symptomatology.

While many cases of alleged traumatic neurosis may not as readily lend themselves to analysis as did the one presented, approaching such a case

---

¶In this case the problem was similar to the general problem in delayed-onset traumatic neurosis, namely that the content of the symptoms seems to be related to the traumatic situation. In this case it was preoccupation with the accident; in other cases, the content of symptoms might be related to military experience, for example.

In these cases it is apparent that the traumatic prior experience is a partial cause in the sense that without the experience the individual at the time following the traumatic experience would not have symptoms related to the trauma. But the trauma does not explain the more critical issue of the very existence of the symptoms at the time they occur. In legal terms, the form of the symptoms fulfills the "but for" test with respect to proximate cause. However, the form of symptoms is by itself not a sufficient condition to establish proximate cause. It is more likely, by the argument above, that independent causes subsequent to the trauma led to the very fact of occurrence of the symptoms.

systematically in terms of the different points involved in causation should enable the forensic psychiatrist to make an important contribution to that critical legal issue in the case in which he or she is involved. To do so, of course, means investigating each of the possible points of causation with whatever information can be obtained regarding the point.

Incidentally, the example case was closed by a settlement that occurred during the presentation of the plaintiff's case in the trial.

Summary

Because there is not definite scientific knowledge about posttraumatic neurosis, a psychiatrist involved in evaluating allegations of psychiatric impairment in negligence law cases can make only a circumscribed contribution. Nevertheless, the contribution can be an important one.

References

Andreasen NC: Posttraumatic stress disorder. Section 21.1d in Kaplan H., Freeman AM, Sadock B, eds. Comprehensive Textbook of Psychiatry. New York, Williams & Wilkins, 1980

Clark and Lindsell on Torts. 14th Ed. London, Sweet and Maxwell, 1975

DSM III. Washington, American Psychiatric Association, 1980

Kuster: Toxic substance contamination: The risk benefit approach to causation analysis. 14 U Mich J L Ref 53 (1980)

Prosser W: Handbook of the Law of Torts, 4th Ed. St. Paul, West, 1971

Voght: Tests of causation and the Florida jury instructions—The current conflict and the need for a change. 32 U Fla L Rev 308 (1980)

Wigmore JH: Evidence in Trials at Common Law, v.2. (Rev by Chadbourne JH) Boston, Little Brown, 1979 □