The Duty to Warn Revisited: Contemporary Issues within the North American Context

Fotini Zachariades PhD C. Psych. (Supervised Practice)
Eating Disorders Program,
Children's Hospital of Eastern Ontario.
Ottawa Anxiety and Trauma Clinic.
Ottawa, Ontario, Canada

Christine Cabrera PhD (candidate)
School of Psychology, University of Ottawa,
Ottawa, ON, Canada

**Abstract**

The following review discusses legal, ethical, and clinical developments that have ensued as a consequence of the case of Tarasoff v. Board of Regents of the University of California in 1976. The Tarasoff ruling is examined in terms of implications for law and the professional practice of psychology, including clinical implications and practical guidelines, and issues of malpractice, negligence and liability. The duty to warn/protect, the related duty towards third parties, expansions of the duty to warn/protect ruling, and its subsequent legal implications are discussed. Furthermore, the concepts of dangerousness and foreseeability from a legal viewpoint, and risk assessment and prediction from a clinical viewpoint are also addressed. Overall, the impact of the duty to warn/protect on the current professional practice of psychology is examined, and corresponding clinical implications are noted.

**Key words:** Duty to warn, duty to protect.

**Introduction**

The case of Tarasoff v. Board of Regents of the University of California (1976) lead to developments in law initially with regard to the duty to warn and subsequently with the duty to protect (e.g., Appelbaum & Rosenbaum, 1989; Quattrrocchi & Schopp, 2005; Truscott, Evans, & Mansell, 1995). The California Supreme Court ruled that when a therapist judges according to professional standards that a serious danger of violence to another is presented by a patient, the therapist is obligated to protect the intended victim(s) from such danger by employing reasonable care (Daniels & Ferguson, 1999). Therein, the law states that disclosure of confidential information is appropriate for the purposes of protecting clients who have expressed credible suicidal intent; or engaged in activities likely leading to imminent and substantial harm to others (Fisher, 2003). The Tarasoff rule has recently been extended to include threats disclosed by family members concerning a client, with the courts asserting that there is no difference between threats conveyed directly by a client and those communicated by a client’s immediate family member (Ewing, 2005).

The following review examines the integration of psychological considerations, ethical values, and contemporary North American legal guidelines, and explores recommendations for clinical practice, in the context of the duty to warn/protect.

**Ethical considerations**

Hadjistatropoulos and Malloy (2000) proposed that an integration of principle and virtue ethics, teleology, deontology, and existentialism, may have more application value for the various ethical dilemmas encountered by professionals. Hansen and Goldberg (1999) have suggested a matrix of considerations to be examined within decision-making processes for ethical dilemmas. Elements in this matrix include: moral principles/personal values, professional codes of ethics, agency and federal, state and local statutes, rules and regulations, and case law. These models serve to widen the scope of the decision-making process within the professional practice of psychology (Hansen & Goldberg, 1999).

The ethical dilemma regarding duty to warn can be averted with the use of appropriate informed consent procedures. Informed consent procedures prior to clinical contact should involve adequate discussion about the risks and limits of confidentiality, and of foreseeable uses of the information. The extent and limits of confidentiality can be regulated by law. As a general rule, it is recommended that the least restrictive measures to prevent violence should be employed (Felthous, 2006) in order to preserve...
the therapist-client relationship. Duty to warn/protect should be implemented only as a last resort, if previous attempts to de-escalate with the client have been ineffective. Even then, only the minimum amount of information required for protection of the victim or the public should be disclosed, such as stating the specific threat but refraining from providing any opinions or predictions (Austin, Moline, & Williams, 1990).

The presence of “clear and imminent danger” has been proposed as the signal on which to base your decision to breach confidentiality (Nixon, 1984). However, it has not been adequate in resolving the range of situations encountered professionally. Therefore, even though situations may arise in which disclosure for the prevention of “imminent” danger is usually permitted; the assessment of risk and its immediacy are not as clear. The difficulty is determining where the “public peril” begins; since danger can be construed as being of a physical, psychological, or fiscal nature and includes a myriad of threats. Lack of predictive certainty poses a disconcerting dilemma, whereby overestimation of a client's dangerousness may unnecessarily jeopardize the therapeutic relationship, yet on the other hand, underestimation could lead to potentially preventable harm (Nixon, 1984; Szasz, 2003). Current guidelines governed by concepts of “minimal amounts of information,” “imminent danger,” and “ineffective client interventions,” remain subjective and open to interpretation, hence making them difficult to standardize. As such, the therapist is required to apply a degree of personal judgement in order to govern decisions pertaining to the duty to warn/protect.

Professional ethics and ethical values

Both the American Psychological Association (APA) and Canadian Psychological Association (CPA) ethics codes relating to the principles of beneficence and non-maleficence state that although psychologists should strive to safeguard the welfare of those with whom they work, it is acknowledged that in view of the context of the duty to warn/protect, this is not always possible (Fisher, 2003). The principle of responsible caring also includes the standard to offset or correct harm whenever possible (Truscott & Crook, 2004; CPA, 2000), weighing the risks involved to each party (Morrison, 1989), and the principles of beneficence and maleficence towards the patient being balanced with social responsibility (Farnham & James, 2001).

Teleological or deontological systems of ethics such as utilitarianism, ethical objectivism, communitarianism, and feminist ethics may have application value to the duty to warn/protect context (Truscott & Crook, 2004; Fisher, 2003). For instance, ethical objectivism may have application value in instances requiring the sacrificing of confidentiality when there is a clear and present danger to a client or others. Communitarianism and feminist ethics (Fisher, 2003) are also useful approaches relevant to multicultural contexts which consider emotional factors within ethical decision-making. Considerations from the various approaches may therefore be incorporated in decisions involving dilemmas related to the duty to warn/protect.

Legal considerations

When does the duty to warn exist?

Various Tarasoff-like cases have led to a wide range of legal decisions regarding when the duty to warn should be enacted. Generally, these rulings can vary from: 1) the existence of broad threats to any victim, 2) the existence of a specific threat to an identifiable victim as suggested by the “Thompson Rule,” 3) the need to preserve confidentiality regardless of any identified threats, or 4) the conditional duty to warn only if the patient is under custodial care of the therapist (Felthous, 1989).

Insistence on first satisfying the “Thompson” or specificity rule appears to reflect an attempt on the part of the courts to limit an undefined expansion of the Tarasoff principle. Legally, it is generally accepted to enact the duty to warn/protect only when foreseeable violence towards an identifiable victim is present without evidence of an explicit threat (Felthous, 1989). Nevertheless, the question appears focused on how to protect third parties from serious harm with minimal intrusion on the client's privacy, rather than whether to warn a potential victim (Melton, 1988).

The legal test of “reasonable foreseeability” does not require “certainty,” but rather an attempt to establish whether a prudent psychotherapist would have predicted violence on the client's part (Truscott, Evans, & Mansell, 1995). A foreseeable victim is said to be in the “zone of danger” and subjected to the probable risk of a patient's violent conduct (Carmack, 2004), and furthermore this may relate to a class of persons at risk rather than one specific victim (Lamb, Clark, Drumheller, Frizzell, & Surrey, 1989).

Beyond the Tarasoff-like rulings, some North American jurisdictions have mandatory reporting laws regarding the knowledge of intent to commit a criminal act. Furthermore, some jurisdictions have expanded the duty to warn/protection laws requiring the warning of third parties (i.e., family members of the intended victim) (e.g., Fisher, 2003). Given the wide array of legal permutations on when the duty to warn/protect is appropriate; the therapist may be faced with significant challenges on how to carry out the proper due diligence when it comes to upholding these laws.

Legal considerations and the standards of clinical care

The duty to protect has since been extended to become a standard of care to not only include physical harm towards others but also to self, as well as to include other types of harm (e.g., financial, emotional, social) (Werth, Wright, & Bardash, 2003). The “should have known” standard employed in Tarasoff however, remains unclear in scope and thus potentially over-inclusive, since disclosures can be made in the interests of protection from liability even if the patient is not thought to be likely dangerous. On the other hand, reporting threats only if they are “explicit” may be under-inclusive, as circumstances requiring protective action could be inadequately addressed (Appelbaum, Zonana, Bonnie, & Roth, 1989).

The standards of the duty remain ambiguous and clarity is lacking as to whether the clinician's obligation to third parties creates a professional or non-professional standard of protection,
since the courts have articulated this obligation sometimes as a specific duty to warn, and at other times as a general duty to protect others (Quattrocchi & Schopp, 2005). Therefore, the duty apparently creates two standards for clinicians to fulfill involving the assessment of foreseeable harm, and the protective conduct triggered by foreseeable harm. In the Tarasoff ruling, the court required that clinicians determine foreseeable harm according to the standards of their profession (Quattrocchi & Schopp, 2005).

Subsequent cases and statutes in certain North American jurisdictions however, have altered this standard by requiring a specific threat to an identifiable victim, which is not a professional standard per se but a legal definition. As such, this does not necessitate any professional judgment or expertise in the determination of the foreseeability of harm (Quattrocchi & Schopp, 2005). The Tarasoff ruling appears to articulate a broad duty to protect but the protective conduct that would discharge this duty remains unclear, in particular with regard to distinguishing between professional interventions requiring expertise, versus non-professional interventions requiring no professional expertise. The limitation of the duty in circumstances involving a specific threat or identifiable victim seems to emphasize protection in terms of the non-professional action of warning. However, such protective measures do not appear to encourage the application of professional expertise such as conceptualizing risk assessment or the development of clinical interventions (Quattrocchi & Schopp, 2005).

**Assessment of risk versus assessment of dangerousness**

It has been contended that clinicians assess the risk of violence rather than dangerousness, which is a legal construct (Felthous, 2006). Therefore, a judgement of dangerousness is not the same as the assessment of risk. The former is a legal criterion, taking into account the complete body of evidence, whereas the latter informs the empirical component of the dangerousness judgement by providing evidence regarding the probability and severity of harmful conduct. Tension may thus develop between legal doctrines regarding dangerousness and contemporary approaches to risk assessment (Claussen-Schulz, Pearce, & Schopp, 2004).

The accuracy of predictions of dangerousness by mental health professionals has been extensively debated, however the duty to warn/protect requires that a professional judgement be made regardless of whether danger is “clear and present” (Keith-Spiegel & Koocher, 1985). Dangerousness is comprised of violence prediction variables (risk factors), the amount and type of violence being predicted, and the probability that harm will occur (Monahan, 1996). Although foreseeability of danger varies according to the courts, it generally pertains to whether the patient presented a serious danger of violence to others (Felthous, 1989).

What is deemed to be foreseeable has been expanded in such vague terms as “probability of danger” and “unreasonable risk of harm to others.” In some cases, substantial property damage has been included or excluded in the definition of dangerousness. Foreseeability is related to a retrospective event within the context of a lawsuit (Felthous, 1989), while clinical prediction refers to the conclusion an expert or clinician reaches when presented with a set of facts about a particular problem (Westen & Weinberger, 2004). Since the ability to concretely foresee dangerous actions by an individual is not an output of professional psychological risk assessment, it is inappropriate to base legal rulings entirely on extrapolation of probable risk assessment data. The gap between professional risk assessment and the extrapolation of risk data to generate an output of legally enforceable “foreseeable danger” can be narrowed by employing clinical as well as actuarial data in patient treatment and management (e.g., Reid, 2003).

Dangerousness can also be examined in terms of the magnitude of harm, the probability that the harm will occur, the frequency with which the harm will occur, and the imminence of the harm (Janus & Meehl, 1997). In the context of possible committal of a patient, the Mental Health Treatment Act of Canada has a requirement stating that a patient may be confined if a mental disorder will likely result in harm or danger to the person or others. However, the Provinces and Territories differ in the degree, probability, and time course of the expected harm/danger; and the type of danger/harm required for committal. The timeliness aspect refers to the imminence of the harm, which has lent itself to broad interpretation by the courts. The degree of harm has been variably defined with the use of such terms as “serious harm,” “seriously bodily harm,” “substantial harm,” and “grave and immediate danger.” The period over which symptoms can be considered relates to recent rather than historic behaviour patterns of the patient. The type of danger/harm concerns physical/bodily as well as non-physical harms to self and others- the former is included in the harm criteria of all Canadian jurisdictions. However, only seven jurisdictions allow non-physical harms to qualify to a greater or lesser degree. The harm/dangerousness criterion for involuntary hospital admission employed in the Mental Health Acts in Canada has ranged from being defined in terms of strict physical dangerousness to broader harm criteria, including mental and other non-bodily harms (Gray, Shone, & Liddle, 2000).

An advantage of less strict and more broadly defined harm criteria is that it may facilitate the early treatment necessary for a good prognosis. Judgements on future behaviour appear to be more accurate the broader the category of the behaviour being predicted (e.g., mental health professionals would be better able to assess the probable course of a certain mental illness without treatment) (Gray, Shone, & Liddle, 2000).

**Clinical considerations for predictions of dangerousness/harm**

Tarasoff is at least in part a doctrine of negligent misdiagnosis, an assessment of dangerousness, and involves behavioural prediction (Melton, 1988). Conceptualization of foreseeable harm in terms of prediction as opposed to risk appears to be a disadvantage of the duty (Quattrocchi & Schopp, 2005). The following risk factors can be considered the best predictors of imminent danger when considered in conjunction (Ferris et al., 1998): past violent behaviour, tendencies toward impulsive action, family history of violence, level of anger expressed, alcohol or other drug use, expression of specific threat, plan of action, and ownership of a weapon (Costa & Altekruse, 1994).

The importance of assessing a situation every time a patient threatens serious harm has been emphasized, even if similar threats have not been carried out in the past (Ferris et al., 1998). A patient’s intent should be re-evaluated periodically as this may change over time (Appelbaum, 1992). In distinguishing a plan...
versus intent, it must be remembered that a plan or intent to carry out a plan are symptoms that should be viewed in their entire context. Some individuals may be highly dangerous even without expressing a plan or intent to do harm, therefore other indicators (e.g., weapon ownership with impulsive temperament) of danger besides direct communication of a plan or intent must be considered (Slovenko, 1992).

**Malpractice, negligence, and liability**

Therapists are subject to liability and civil damages if they neglect their duty to warn/protect (Hansen & Goldberg, 1999; Austin, Moline, & Williams, 1990). If non-clinical protective measures such as warning the patient or notifying the police are being implemented, this may fulfill the duty. Otherwise, the omission of such measures could be regarded as negligence or malpractice possibly subject to civil liability. Liability is based on negligence with the application of elements of malpractice under tort law, involving whether a duty exists, whether there was a dereliction in carrying out this duty, and whether this dereliction proximately caused injury or damage. Negligence related to proximate cause can include failures of diagnosis, treatment/hospitalization, wrongful discharge, abandonment, or insufficiently scheduled outpatient appointments, leading to sub-optimal treatment and progress monitoring. Such errors can also be regarded as malpractice, and should not be confounded with failure to fulfill vague duties such as the duty to warn/protect (Felthous, 1989).

Liability is usually imposed for failure to follow appropriate procedures particularly in gathering and communication of information, rather than for judgement errors in light of known facts (Truscott, Evans, & Mansell, 1995). It is the therapist's responsibility to follow a “standard” method in the assessment and management of dangerous clients, which should include attempts to obtain previous medical records (Austin, Moline, & Williams, 1990).

In some cases it appears that the duty may arise without an identifiable victim, and thus liability may be extended to third parties if it is foreseeable that a patient's conduct may cause harm to another (Felthous, 1989). A third party is considered to be a family member or legal guardian of the intended victim. The related duty towards third parties characterizes the therapist's actions as reaching beyond the therapist-patient dyad. Therefore, when the therapist's behaviour is directed at protecting a non-patient, the therapist thereby assumes a duty to that person to avoid acting negligently (Appelbaum & Zoltek-Jick, 1996).

**Clinical implications and practical guidelines**

According to Monahan (1993), there is no national American legal standard for what clinicians should do when assessing risk. Felthous (1989) purported dealing with a potentially violent patient in terms of assessing dangerousness and selecting, then implementing, a course of action based on such an assessment. Monahan (1993) has recommended five guidelines for limiting therapist exposure to liability in terms of risk containment relating to: risk assessment, risk management, documentation, policy, and damage control. Some preventative measures relating to damage control include intensified treatment (e.g., frequent sessions), incapacitation (such as hospitalization), or “target-hardening” (such as warning the potential victim), and it is emphasized that these actions should be reflected in documentation (Monahan, 1993).

Reasonable efforts should be made to obtain past medical records. “Building the record” or documenting information relevant to risk should include the rationale for any actions taken (Monahan, 1993). Thorough record-keeping is important as it provides evidence that the therapist understood the situation and that reasonable steps were carried out in light of the facts (Fulero, 1988).

In terms of reviewing informed consent and the limits of confidentiality with patients, any breach of confidentiality should be discussed before it occurs, and gaining the patient's compliance should be attempted. The duty may be satisfied in terms of notifying law enforcement officers or hospitalizing the patient if it becomes necessary to inform the potential victim. The potential victim should be warned even if the patient refuses to allow any breach of confidentiality, if professional judgement deems this necessary (Felthous, 1989). Whenever possible and in keeping with the APA/CPA principles, therapists should strive to implement the least disruptive measure in order to fulfill the duty to warn/protect (Truscott, Evans, & Mansell, 1995).

It is suggested that even when danger is imminent, the patient should actively participate in helping to avoid a violent episode (Baird & Rupert, 1987). Colleagues should be consulted if there is any uncertainty with regard to the assessment of dangerousness. Consultation indicates that there is professional consensus regarding actions taken (e.g., Fulero, 1988). However, consultation with professional peers should not be viewed as an exercise in finding the “right” answer, but rather as a process through which to determine responses that would be the most defensible in terms of both legal constraints, as well as consistent with professional ethical values (Nixon, 1984).

**Conclusion**

Given a duty to warn/protect, the therapist aims at the most optimal compromise possible but still faces a professional dilemma, pulled between upholding the best interests of the client and the best interests for societal protection (Chaimowitz, Glancy, & Blackburn, 2000; Chaimowitz & Glancy, 2002). The duty to warn is a decision fraught with ethical and legal implications, and there is often not a perfect response to each unique situation. It remains that decisions related to the duty to warn should be approached with awareness of personal and professional ethics, and with consideration of relevant legalities, weighing these against clinical knowledge as applied to the context of the particular client.

**References**


