1998b, 1999, 2000b). Across these studies, approximately three quarters of trauma survivors with ASD subsequently developed PTSD.

In contrast, the predictive ability of the ASD diagnosis is less promising when one calculates the proportion of people who eventually developed PTSD and who initially displayed ASD. This approach reveals that although some reports indicated that the majority of people with PTSD initially displayed ASD, most studies found that only a minority of people with PTSD suffered ASD within the initial month after trauma exposure. That is, the capacity of the ASD diagnosis to accurately identify most people who will eventually develop PTSD appears limited. The limitations of the ASD diagnosis as a reliable and sensitive predictor of subsequent PTSD have also been underscored by recent evidence that the ASD diagnosis may not be superior to PTSD criteria (employed within the initial month after trauma exposure) as a means of identifying people who will subsequently develop PTSD (Brewin, Andrews, & Rose, 2003). Further, although Difede et al. (2002) found that 87% of burns survivors with ASD subsequently developed PTSD, they also reported that applying the PTSD criteria (except duration of symptoms) 2 weeks after burn injury identified the same individuals as developing PTSD.

One major reason for the variability in prospective studies of ASD and PTSD may be the timing of assessments of ASD. Although DSM-IV stipulates that ASD can be diagnosed after 2 days have elapsed since trauma exposure, it is likely that attempting a diagnostic decision this soon will increase the likelihood that a transient stress reaction will be incorrectly classified as a case of ASD. Indeed, Murray et al. (2002) found that the predictive value of the ASD diagnosis depended on when the patients were assessed. Among survivors of motor vehicle accidents, 77% of those who met ASD criteria at 4 weeks developed PTSD, compared with only 32% of those who met ASD criteria at 1 week after trauma exposure. The rapidly changing nature of stress reactions in the initial weeks following trauma exposure is underscored by evidence from studies of civilians involved in the Gulf War, in which many people who suffered immediate stress reactions in the initial days displayed marked adaptation in the following weeks (Solomon, Laor, & McFarlane, 1996). Attempts to distinguish between transient stress reactions and harbingers of chronic disorder on the basis of symptoms expressed within days of trauma exposure will likely be very difficult.

It appears that the major reason why the ASD diagnosis fails to identify many people who eventually develop PTSD (see the right-most column in Table 3) is that the requirement that three dissociative symptoms be present excludes many people who nonetheless develop PTSD. For example, Harvey and Bryant (1998b) reported that 60% of trauma survivors who displayed acute reexperiencing, avoidance, and hyperarousal, but no dissociation, developed PTSD. This pattern undermines the claim that acute dissociation is a necessary harbinger of subsequent pathology.

PREVENTING POSTTRAUMATIC PSYCHOPATHOLOGY

Although many people experience acute stress-related symptoms in the wake of traumatic events, only a minority develop ASD, PTSD, or both. Most people recover from traumatic events without any professional assistance. But given that a significant minority of people exposed to trauma do develop lasting psychological problems, what sort of interventions should be offered, when should they be offered, and to whom? When considering these issues, one should be mindful of important distinctions between different kinds of interventions. *Primary prevention* of PTSD and other posttraumatic problems (e.g., ASD, depression, substance abuse) entails taking steps to reduce the frequency of traumatic events (e.g., restricting adolescents' access to firearms to diminish risk of school violence). These steps usually fall within the bailiwick of law and public health rather than clinical psychology and psychiatry. *Secondary prevention* comprises crisis intervention techniques, such as psychological debriefing, that are delivered within days of the trauma and designed to mitigate distress and prevent the emergence of posttraumatic psychopathology. *Early treatment interventions* are delivered soon after posttraumatic disorders have emerged, but early in the course of the disorders.

In this review, we concentrate on whether secondary prevention, especially the widely used psychological debriefing, and early treatment interventions promote recovery from posttraumatic stress. We acknowledge that survivors and communities have many needs in the aftermath of trauma, and that the prevention of persistent symptoms of psychological distress is only one of them. It is, however, beyond our scope here to review the many different targets of crisis intervention and their effectiveness.

PSYCHOLOGICAL DEBRIEFING

Psychological debriefing has its roots in World War I (Litz, Gray, Bryant, & Adler, 2002). Following a major battle, commanders would meet with their men to debrief them. The objective was to boost morale by having combatants share stories about what had happened during the engagement. This historical group debriefing method was also used by American troops during World War II and continues to be used by the Israeli army today (Shalev, Peri, Rogel-Fuchs, Ursano, & Marlowe, 1998).

Drawing parallels between the stress of combat and the stress of emergency medical service, Mitchell (1983) reasoned that a similar approach might diminish stress reactions among firefighters, police officers, emergency medical technicians, and other people exposed to what he referred to as "critical incidents" (i.e., traumatic events). A former firefighter and paramedic, Mitchell obtained a Ph.D. in human development and developed the most widely used method of psychological de-

briefing: Critical Incident Stress Debriefing (CISD). In his seminal article, Mitchell (1983) emphasized that too many people believe that firefighters, police, and other emergency service personnel are emotionally impervious to trauma. Contrary to the John Wayne stereotype, he said, "Rescuers are vulnerable human beings who have all the normal physical and psychological responses to the horror of human suffering" (Mitchell, 1983, p. 36). That is, helping the primary victims of trauma might constitute a major stressor for the helpers themselves. Accordingly, Mitchell asserted that the mental health of emergency personnel who respond to a critical event is best served when they participate in a structured session enabling them to talk about the event and ventilate their emotions, especially in the company of peers who have experienced the same incident.

CISD is designed to mitigate the adverse psychological consequences of traumatic events by attenuating the intensity of acute symptoms of stress, thereby reducing the risk of subsequent psychiatric problems. A single debriefing session "will generally alleviate the acute stress responses which appear at the scene and immediately afterwards and will eliminate, or at least inhibit, delayed stress reactions" (Mitchell, 1983, p. 36). Originally developed for emergency service personnel, CISD is now deemed helpful for "primary victims" (i.e., the people directly exposed to trauma; Everly & Mitchell, 1999, p. 85) as well. (However, in 2002, Mitchell continued to dismiss studies showing psychological debriefing had no beneficial impact if they involved primary victims.) Debriefings now occur in businesses, schools, hospitals, and the military (Everly & Mitchell, 1999, pp. 84–85). There are different versions of psychological debriefing (Raphael & Wilson, 2000), but "Mitchell's CISD model of psychological debriefing is generally recognized as the most widely used in the world and is used across the greatest diversity of settings and operational applications" (Everly & Mitchell, 1999, p. 84).

Initially, Mitchell (1983) described a debriefing session as "either an individual or group meeting between the rescue worker and a caring individual (facilitator) who is able to help the person talk about his feelings and reactions to the critical incident" (p. 37). However, since the late 1980s, Mitchell has argued that CISD should be delivered only to groups of individuals who have been exposed to a critical incident, not to single individuals. That is, although originally deemed suitable for either individuals or groups, CISD is now recommended only for groups.

A CISD session lasts between 3 and 4 hr and is conducted between 2 and 10 days after a critical incident, except in mass disasters, in which case it occurs about 3 to 4 weeks later (Everly & Mitchell, 1999, p. 18). According to its advocates, debriefing works because it is delivered soon after the trauma, because it provides psychosocial support and an opportunity for expressing emotions and thoughts about the trauma, and because it provides tips on coping and education about stress and its management.

A debriefing session has seven phases. In the introduction phase, the debriefing facilitator explains the procedure to the participants, answering any questions they might have. He or she emphasizes that debriefing is not psychotherapy; it is a method for reducing normal stress reactions triggered by a horrific event.

The facilitator then initiates the fact phase by asking each participant, in turn, to describe what happened during the critical incident. He or she might say, "Tell me who you are, what your role in the incident was, and just what you saw and/or heard take place" (Everly & Mitchell, 1999, p. 86). The purpose is to enable each person to describe the traumatic incident from his or her perspective. "Each person takes a turn adding in the details to make the whole incident come to life again in the CISD room" (Mitchell, 1983, p. 38). However, the facilitator reassures participants that they can remain silent if they feel uncomfortable speaking in the group. The option of silence also applies to subsequent phases of the debriefing.

Next, the facilitator shifts to the thought phase by allowing each participant to describe his or her cognitive reactions to the traumatic event. The facilitator might say, "Now, I'd like you to tell us what your first thoughts were in response to the crisis" (Everly & Mitchell, 1999, p. 86). The purpose of this phase is to move closer to the expression of emotion.

The facilitator then moves to the reaction phase—the one designed to foster emotional processing of the trauma by having participants experience catharsis through expressing their feelings about the event. The facilitator might begin by asking, "What was the worst part of the incident for you personally?" (Everly & Mitchell, 1999, p. 86). The facilitator may ask how each person felt then and also how each person is feeling during the debriefing itself. As Mitchell (1983) has emphasized, "Everyone has feelings which need to be shared and accepted. The main rule is—no one criticizes another; all listen to what was, or is, going on inside each other" (p. 38).

Advancing to the symptom phase, the facilitator asks, "What physical or psychological symptoms have you noticed, if any, as a result of this incident?" (Everly & Mitchell, 1999, p. 87). The purpose of this phase is to identify stress reactions that members wish to share, and to begin the transition from the affective realm back to the cognitive one.

In the teaching phase, the facilitator tries to show that the stress reactions members have been experiencing are normal and not necessarily a medical problem, by stating, for example, "We've heard numerous symptoms that are being experienced, let me explain their nature and give you some suggestions on how to reduce their negative impact" (Everly & Mitchell, 1999, p. 87). In addition to providing stress-management tips, he endeavors to convince participants that their reactions do not signify psychopathology.

Finally, in the reentry phase, the facilitator aims to achieve closure to the traumatic event. He or she summarizes what has been covered in the session, answers any questions that have arisen, and assesses whether any members may need follow-up or referral for additional services.

Although Mitchell (1983) originally asserted that "the formal CISD should be mandatory for all personnel involved in the scene" (p. 38), he has since acknowledged that compelling people to undergo debriefing raises "intriguing" issues (Everly & Mitchell, 1999, p. 93). On the one hand, allowing participation to be voluntary "runs the risk of under utilization based upon the stigma of needing help" (Everly & Mitchell, 1999, p. 93). On the other hand, mandatory debriefing "raises issues of coercion, legal liability, and informed consent" (Everly & Mitchell, 1999, p. 93). As a possible solution to this dilemma, Mitchell and Everly have suggested providing a mandatory general information session for everyone involved in the critical incident and then following up with voluntary formal debriefings. In any event, they recommend that anyone exposed to a critical incident be offered debriefing, regardless of whether the person is experiencing any stress symptoms.

In addition to mitigating distress and preventing posttraumatic problems, Mitchell has argued, these interventions may reduce sick days taken by stressed employees (Everly & Mitchell, 1999, pp. 131–135). "Not only do [these] services make sense from a humanitarian perspective, they make sense from a business perspective, as well" (Everly & Mitchell, 1999, p. 135). According to Everly and Mitchell (1999, p. 135), a business's failure to implement some such psychological service in the immediate wake of a critical incident may constitute negligence, thereby increasing the risk of legal liability to stressed employees who may file suit. To avoid the threat of litigation for failing to meet the standard of care, some police departments in the United Kingdom have now made debriefing compulsory for people exposed to critical incidents; banks in the United Kingdom and Australia have also made debriefing compulsory for employees exposed to critical incidents in the workplace (Rose, Bisson, & Wessely, 2001). Everly and Mitchell (1999, p. 135) predicted that their approach might emerge as "the standard of care" for intervention in the wake of crisis and trauma. Indeed, between 30,000 and 50,000 individuals are trained each year by their organization, the International Critical Incident Stress Foundation, Inc. (ICISF).[3]

In the 1990s, Everly and Mitchell (1999) expanded the range of crisis intervention services offered by ICISF. The proprietary term for the entire set of techniques is Critical Incident Stress Management (CISM). The overarching purpose of CISM is "to reduce the incidence, duration, and severity of, or impairment from, traumatic stress" (Everly & Mitchell, 1999, p. 72). Hence, CISM is not a technique or method per se; it is a framework or strategy comprising a set of tactics, each designed to meet a distinct crisis intervention goal. Grouped under the CISM rubric are the following methods, in addition to CISD (Everly & Mitchell, 1999, pp. 71–92):

- *Pre-incident preparedness training* refers to educating individuals in high-risk occupations (e.g., firefighters, emergency medical technicians) about the kinds of stressors they are likely to encounter on the job, about common stress reactions, and about stress-management techniques.
- *One-on-one individual crisis support* refers to attempts to mitigate acute stress reactions, often at the scene of the trauma. The counselor attempts to provide psychological distance between the scene and the person in distress by having the person take a walk, get a cup of coffee, and so forth. As Everly and Mitchell (1999) pointed out, "In actuality, most crisis response interventions will be done individually, that is, one-on-one, rather than in groups" (p. 76).
- *Demobilization* refers to providing food, rest, and information about coping with stress reactions to large groups of disaster personnel as they rotate off duty. This method includes group informational briefing, which refers to providing facts about a critical incident (e.g., a student's suicide) to a large group of individuals indirectly affected, as well as providing information about common psychological dynamics (e.g., grief, anger) and about how to access psychological services. This tactic usually applies to schools and businesses affected by a critical incident.
- *Defusing* refers to a small-group intervention that usually takes place within 12 hr of the traumatic event. It involves having participants explore and discuss the incident and their emotional reactions to it. Group leaders teach coping skills and tell participants that stress reactions are normal and expected, and do not necessarily signify mental illness. Otherwise similar to CISD, defusing can be repeated.
- *Family support* refers to debriefing family members of the persons involved in the crisis (e.g., spouses of individuals in the military).
- *Referral mechanisms* concern procedures for referring individuals for psychiatric or psychological services, legal services, career counseling, and so forth.

Does Psychological Debriefing Work?

Most people who receive debriefing endorse it as helpful (e.g., Carlier, Voerman, & Gersons, 2000; Small, Lumley, Donohue, Potter, & Waldenstrom, 2000). But this does not mean that it prevents posttraumatic mental disorders. These reports that the method is helpful may reflect little more than polite expressions of gratitude for attention received. Given that only some trauma-exposed people will develop PTSD—and many recover on their own—the efficacy of debriefing can be gauged only by comparing the outcomes for individuals who

3. Kadet (2002) reported that 40,000 people are trained in Mitchell and Everly's methods each year. On March 2, 2002, she contacted Don Howell, the executive director of the ICISF, to ask him how many people are trained by ICISF each year. According to her notes (A. Kadet, personal communication, April 2, 2003), Howell said, "We do 30 to 50 thousand people a year; that's a conservative number. We've been at that pace for the past 4 or 5 years." When the fact checker for *Smart Money* magazine recontacted him to double-check the figures, he put the figure at 40,000 individuals trained annually. In a training workshop given by Mitchell in Columbia, Maryland, on May 30, 2002, he said his organization trains 30,000 people per year (S. Satel, personal communication, April 5, 2003).

did and did not receive this intervention. If a certain crisis intervention method fails to reduce symptoms and prevent disorder, then it should be discontinued in favor of developing something that actually does work.

Some scholars are convinced of debriefing's efficacy. Mitchell and Everly (2001) argued, "The experiences of 700 CISM teams in more than 40,000 debriefings cannot be ignored. This is especially so when the overwhelming majority of the reports of debriefing services are extremely positive" (p. 295). Mitchell and Everly (2001) claim that "numerous studies have already been published with very positive results" (p. 295), and that research on their methods "proves their clinical effectiveness far beyond reasonable doubt" (Mitchell & Everly, 2001, p. 84; see also Everly, Flannery, & Eyler, 2002).

Other scholars, who have published meta-analyses, have drawn dramatically different conclusions. Rose et al. (2001) concluded that "there is no current evidence that psychological debriefing is a useful treatment for the prevention of post traumatic stress disorder after traumatic incidents. Compulsory debriefing of victims of trauma should cease" (pp. 1–2). Another meta-analysis revealed that trauma-exposed individuals who had not received CISD experienced reductions in PTSD symptoms, whereas those who had received CISD did not (van Emmerik, Kamphuis, Hulsbosch, & Emmelkamp, 2002).

Rose et al. (2001) confined their meta-analysis to randomized controlled trials (RCTs) of psychological debriefing. The RCT is the standard method for testing the efficacy of any intervention, psychological or pharmacological. It requires that subjects be randomly assigned to receive either the intervention or no treatment. Reliable and valid measures of symptoms are administered to both groups both before and after the intervention is delivered to the treatment group, and readministered at follow-up, usually 6 months to several years later. If the intervention works as predicted, individuals who received the intervention should have better outcomes than those who did not. A failure to include a no-treatment (or assessment-only) control group makes it impossible to distinguish between improvements attributable to the intervention and those attributable to natural processes occurring with the passage of time. In the following review, we discuss all the RCTs on psychological debriefing that were included in the meta-analyses of Rose et al. and van Emmerik et al. (2002) and were published in journals; we also discuss some non-RCTs.

For several reasons, using the RCT model to test the efficacy of crisis intervention methods is more complicated than using it to test the efficacy of psychotherapy or pharmacotherapy. First, investigators may object to randomly assigning trauma-exposed individuals to a no-treatment control condition. Depriving them of a potentially helpful treatment seems to raise ethical issues. Of course, this objection presupposes that the intervention is, indeed, effective. If an intervention is not known to work, there is no ethical problem in withholding it. Moreover, some interventions may impede natural recovery, making it essential that RCTs test the efficacy of all interventions. Second, unless psychological measures are already available (e.g., predeployment measures of psychological functioning among soldiers about to embark on a military mission), it is difficult to get pretrauma assessments. Third, debriefing is often administered following man-made or natural disasters, unexpected events that produce much chaos. Unless somehow prepared in advance to conduct such studies, researchers are often unable to devise a proposal, secure ethical (human subjects) approval, and launch a study in a timely fashion.

Everly and Mitchell (1999, pp. 109–110) believe that the merits of RCTs have been overstated. More specifically, they argue that random assignment to debriefing and control groups is extremely difficult, especially in a crisis intervention context, and may undermine the probative import of the findings. Attempts to increase internal validity (experimental control) may wind up sacrificing external validity (generalizability to the real world of crisis intervention). That is, in real-world crisis situations, individuals are not randomly assigned to no-treatment control conditions preceded and followed by clinical assessments, whereas these features are integral to RCTs.

How is it possible for scholars to arrive at dramatically different conclusions about the value of psychological debriefing? Close reading of the publications of the critics and the advocates of debriefing reveals that much of the time, they are relying on different sources of evidence to arrive at their conclusions. Debriefing advocates cite allegedly positive findings that the critics regard as suffering from fatal methodological flaws (e.g., failure of randomized assignment, absence of control groups). Debriefing critics cite other studies that advocates claim fail to provide proper tests of the method (e.g., studies testing one-on-one debriefings rather than group debriefings). Therefore, to clarify this contentious issue, we first review the data debriefing advocates adduce in favor of the method. We then review the data debriefing critics adduce against the method. Finally, we review and evaluate the replies of debriefing advocates to the "negative" studies.

Studies Adduced in Support of Debriefing

A few published, peer-reviewed studies are included among the research adduced as confirming the efficacy of debriefing (Everly & Mitchell, 1999, pp. 107–129; Mitchell, 2002). Conducting methodologically sound research in this area is very challenging. Nevertheless, one's confidence in the efficacy of debriefing (or other interventions) is enhanced when researchers assign participants randomly to the treatment and no-treatment control conditions, deliver a standardized intervention with reasonable fidelity to the protocol, use reliable and valid measures of psychological distress, and conduct subsequent assessments to determine whether debriefed individuals are doing better than nondebriefed individuals at follow-up.

In an early pre-CISD study cited as relevant to CISM by Everly and Mitchell (1999), Bordow and Porritt (1979) randomly assigned 70 male patients who had been hospitalized following road traffic accidents to a one-on-one crisis intervention condition or to no intervention. The intervention comprised providing practical assistance (e.g., obtaining financial assistance from welfare agencies), exploring emotional reactions, and encouraging family members to be supportive of the patient. During assessments 3 to 4 months later, the crisis intervention group reported significantly fewer psychiatric symptoms than did the untreated group, whose members improved very little.

In another study cited in support of crisis intervention (Mitchell, 2002), Bunn and Clarke (1979) randomly assigned 30 individuals, who had accompanied a seriously injured relative to the hospital, to receive a 20-min crisis intervention counseling session or no treatment. Although this study was done before Mitchell (1983) developed CISD, the intervention did contain certain elements similar to those of CISD, such as providing information and empathic support, and encouraging participants to express feelings about the crisis. The researchers audiotaped two 5-min speech samples—one before and one after the 20-min intervention—provided by each of the individuals who accompanied their injured relatives to the hospital. (Thus, the participants were not the injured persons, but rather the individuals who accompanied them to the hospital.) The speech samples of the accompanying individuals were rated and scored for expressions of "anxiety." These scores indicated that counseled individuals had a significant decrease in anxiety, whereas participants who received no treatment did not. The brevity of the follow-up period—20 min!—and the unvalidated anxiety measures used make these findings difficult to interpret.

Wee, Mills, and Koehler (1999) published the most encouraging, albeit flawed, study on CISD. They asked emergency medical service personnel, who either had or had not received CISD after having worked during the 1992 Los Angeles riot, to complete a PTSD questionnaire. Logistical constraints prevented the 23 nondebriefed individuals from receiving the otherwise-mandatory CISD. Most participants (72.9%) reported having been attacked by the rioters, so they were primary victims of trauma. Three months after the civil disturbance, the 42 debriefed participants reported significantly fewer PTSD symptoms than did the nondebriefed participants. Unfortunately, the absence of random assignment and assessment of symptoms before the intervention diminish the probative import of the study.

Other studies adduced by Everly and Mitchell (1999, pp. 107–129; Mitchell, 2002, p. 21) in support of debriefing fail to provide much convincing evidence in favor of the method. For example, Amir, Weil, Kaplan, Tocker, and Witztum (1998) studied 15 Israeli women who had survived a terrorist attack on a bus on which they were riding. These women received a group debriefing 2 days after the attack, followed by six group psychotherapy sessions occurring during the following 2 months. Self-reported PTSD symptoms declined between 2 days posttrauma and 2 months posttrauma, but the lack of a no-treatment control group renders any symptomatic change impossible to interpret. Moreover, the outcome measures conflate the effects of the debriefing and the six group psychotherapy sessions. In any event, 27% of the women still met the criteria for a diagnosis of PTSD at 6 months, prompting the authors to state that "it can safely be concluded that the intervention had little if any effect" (Amir et al., 1998, p. 241).

Campfield and Hills (2001) randomly assigned robbery victims to receive either an immediate debriefing (within 10 hr of the crime; $n = 36$) or a delayed debriefing (more than 48 hr after the crime; $n = 41$). Mitchell's (1983) method was followed, and individuals were debriefed either individually or in groups usually ranging from 2 to 4 victims. The immediate-intervention group reported significantly fewer PTSD symptoms at 2 days, 4 days, and 2 weeks postintervention, whereas the delayed-intervention group reported no decline in symptoms during this period. The authors and Mitchell (2002) interpreted these findings as support for immediate debriefing.

Unfortunately, Campfield and Hills's (2001) findings are difficult to interpret because of the short follow-up period and absence of a no-treatment control group. That is, it is entirely possible that a group of robbery victims who received no debriefing would have shown even faster recovery than those receiving either of the debriefing interventions. Furthermore, the study violates three of Mitchell's (2002) criteria for a proper assessment of debriefing. First, in discussing published studies showing debriefing has no effect, Mitchell (2002, p. 18) criticized the researchers for committing an "egregious" error by using change in PTSD symptoms as a dependent variable to gauge the impact of debriefing. In contrast, he praised Campfield and Hills's study for showing a significant decline in PTSD symptoms (p. 24). (Moreover, a decline in symptoms during the initial 2 weeks is likely attributable to natural recovery anyway.) Second, whereas Mitchell emphasizes that standard CISD is designed for emergency service personnel, not primary victims of trauma, the participants were all primary victims of robberies. Third, although Mitchell (2002) has stated that CISD should be delivered only to groups, most participants in this study were treated either individually or as pairs (i.e., average "group" size was 2.5 people per debriefing). In summary, because Campfield and Hills's study violated several of Mitchell's criteria for an adequate test of debriefing, it is puzzling that he affirmed its probative value.

Yule (1992) conducted a study of British schoolchildren who survived the sinking of a cruise shop. Ten days after the shipwreck, clinicians had delivered a single debriefing session to two small groups of the children (total $n = 24$). Children at another school whose administration declined a debriefing ($n = 15$) served as a contrast group. There was no random assignment to groups, and there was no predebriefing assessment.

Five to 9 months after the disaster, the children completed self-report measures of PTSD, anxiety, fears, and depression. The children who had received the debriefing reported significantly fewer intrusive symptoms and fewer fears unrelated to the disaster than did the control group; there was also a nonsignificant trend suggesting that the debriefing group had fewer avoidance symptoms and fears related to the shipwreck. Unfortunately, the lack of random assignment and preintervention assessment complicate interpretation of the data. Moreover, Yule described the single session as "a problem-solving approach, based on cognitive behavioural methods" (p. 203) rather than as CISD.

Deahl et al. (2000) studied 106 British soldiers assigned to 6 months of peacekeeping duty in Bosnia. Prior to their deployment to Bosnia, all the soldiers had received an Operational Stress Training Package that included information about stress and its management. Upon their return from Bosnia, Deahl et al. asked the commanding officers to allocate the soldiers into two groups, and they "did so according to individual availability and commitment to other duties" (Deahl et al., 2000, p. 79). One group ($n = 54$) received a single session of debriefing, based on Mitchell's (1983) model and conducted by experienced debriefers, whereas the second group ($n = 52$) received no intervention. The debriefing session lasted about 2 hr, and was conducted in groups of 8 to 10 soldiers. The debriefing and no-treatment groups were both assessed prior to the first group's debriefing and again 3, 6, and 12 months later.

Although a significantly larger proportion of the control group (25%) than of the debriefed group (7.4%) retrospectively reported experiencing "intense distress" (Deahl et al., 2000, p. 80) while they were in Bosnia, the control group experienced a significant decrease in PTSD symptoms (assessed by questionnaire) by the 6-month assessment, whereas the debriefed group did not. PTSD symptoms assessed by clinical interview did not differ between the two groups at any time point. Indeed, only 2 soldiers in the control group and only 1 soldier in the debriefing group met the criteria for a diagnosis of PTSD at any time since returning from Bosnia. At the 6-month and 12-month assessments, the debriefed group reported significantly fewer symptoms than the control group on a questionnaire tapping anxiety and depressive symptoms. Both groups scored high on a self-report measure of alcohol abuse, but only subjects in the debriefed group evinced a significant reduction in these symptoms. The very low level of problems among both groups led Deahl et al. (2000) to conclude that high rates of psychiatric illness are not inevitable consequences of military conflict. Contrary to Mitchell's (2002) guidelines, the participants in this study were direct recipients of trauma (soldiers) rather than secondary recipients (i.e., emergency service personnel), and the authors themselves described their study as "not a true RCT of debriefing because selection of the sample was restricted, the method of randomization for debriefing was less than ideal and the low level of PTSD symptoms at the outset meant that there was little scope for reduction" (Deahl et al., 2000, p. 83). Mitchell (2002) interpreted this study by Deahl et al. as support for debriefing, emphasizing, for example, the significant reduction in alcohol consumption in the debriefed group.

Chemtob, Tomas, Law, and Cremniter (1997) reported that a single group debriefing session significantly reduced self-reported PTSD symptoms of people exposed to a hurricane that had struck one of the Hawaiian islands. This study did not include a nondebriefed control group. Moreover, the debriefing was delivered 6 months after the trauma. Because debriefing is specifically deemed an early intervention, it should be delivered within days after the trauma, not half a year later. Chemtob et al. acknowledged that their procedure diverged from the standard approach, noting "the length of time between the event and the intervention" (p. 417). The treatment, in effect, was a single psychotherapy session rather than a preventive crisis intervention.

In an even more striking departure from protocol, Busuttil et al. (1995) reported data on 34 individuals who received multiple debriefings within a 12-day residential treatment program. All subjects already had PTSD, 19 of them for 2 to 31 years! One year after treatment, 85% no longer had PTSD. There was no control group. Although debriefing was delivered in groups, the traumatic events suffered by members of the same debriefing group differed substantially, and the participants were primary, not secondary, victims (e.g., survivors of the Falkland Islands War and car accidents, former hostages). Finally, Busuttil et al. characterized their program as "group psychotherapy" (p. 495), not crisis intervention.

Stallard and Law (1993) reported data on 7 adolescents who had escaped from a minibus accident with only minor injuries. Six months later, the youngsters completed PTSD, depression, and anxiety measures prior to receiving two group debriefing sessions. Three months after the intervention, scores on all scales indicated improvement. However, there were too few participants for the study design to include random assignment to a no-treatment control group, making the symptom reductions difficult to interpret. Also, because the debriefing took place months after the trauma, this study is not relevant to crisis intervention. Nevertheless, Everly, Boyle, and Lating (1999) included it in their meta-analysis as support for psychological debriefing.

Nurmi (1999) conducted a single debriefing with small groups of rescue workers 3 to 7 days after the sinking of a ship near Finland. There was some evidence that scores on self-report (including PTSD) measures were lower among occupational groups whose members were debriefed (e.g., firefighters) than among occupational groups whose members were not debriefed (e.g., female nurses). But lack of randomization, lack of predebriefing measures of symptoms, and the presence of gender confounds (e.g., certain nondebriefed occupational groups, such as nurses, had only female participants) render the findings nearly uninterpretable.

Studying groups of British bank employees who had been present during robberies, Richards (2001) collected data on PTSD symptoms 3 days, 1 month, and 3 to 12 months posttrauma. The banks had been conducting single-session CISDs following robberies, but they later incorporated CISD into a comprehensive employee program (i.e., a CISM framework) involving, for example, educational programs on robberies and stress. Richards compared longitudinal symptom reductions during the CISD-only era with those occurring after CISM had been instituted, noting trends indicating outcomes at long-term follow-up were more favorable after the full-fledged CISM program was in effect. It is very difficult to make confident inferences from such a design given the consecutive sequence (CISD followed by CISM), lack of random assignment to groups, and lack of a control group. And it is impossible to attribute the reductions over time to either CISD or to CISM; most people who witness traumatic events, such as robberies, will improve regardless of any intervention.

Leeman-Conley (1990) described a crisis intervention program established in Australia to provide counseling and support for bank employees who are present during bank robberies. After a holdup, each employee is seen individually by a counselor, and then participates in a group session the next day. Employees took 60% fewer sick days off work during the first 2 years after the program's implementation than they had prior to program implementation, and associated worker compensation costs declined 66% as well. Unfortunately, there was no control group, making it impossible to attribute declines in sick days and compensation payments to the program. Moreover, Leeman-Conley reported no data on any psychological symptoms (e.g., anxiety, PTSD). Although she did not cite the work of Mitchell and Everly or mention CISM, Mitchell (2002) nevertheless described this program for direct victims as a "CISM program" (p. 37).

Using Mitchell's (1983) CISD intervention, Bohl (1995) reported that firefighters ($n = 30$) who had received debriefings scored significantly lower on "psychological measures" (p. 126) of depression, anger, anxiety, and "long-term stress symptoms" (p. 126) than did firefighters ($n = 35$) who had not been debriefed. Although Bohl said that she administered "objective tests" (p. 125) 3 months posttrauma, she did not say what those tests were (e.g., questionnaires?). Hers was not a randomized trial; rather, she compared responses of firefighters whose firehouse debriefed its employees with those of firefighters from a firehouse where debriefings did not occur. Moreover, it is unclear whether firefighters were debriefed in groups or one-on-one.

Jenkins (1996) administered questionnaires to 36 emergency medical workers 8 to 10 days after they had worked at the scene of a mass shooting and again about 1 month later. Shortly after the incident, 52% of the workers attended at least one CISD session, but participants were not randomly assigned to receive debriefing or not. Unfortunately, Jenkins merely correlated self-report symptoms with other measures (e.g., of social support). She did not compare psychological outcomes in debriefed versus nondebriefed participants. Nevertheless, she concluded that the study indicates "the apparent usefulness of CISD for reducing symptoms of depression and anxiety over the month after the incident" (p. 488).

Finally, debriefing advocates (Everly & Mitchell, 1999; Mitchell, 2002) have cited Flannery's Assaulted Staff Action Program (ASAP) as support for the efficacy of CISM (Flannery, 1999, 2001; Flannery, Fulton, Tausch, & DeLoffi, 1991; Flannery, Hanson, Penk, Flannery, & Gallagher, 1995; Flannery et al., 1998; Flannery, Penk, & Corrigan, 1999; Flannery, Stone, Rego, & Walker, 2001). Concerned that staff working on psychiatric units are at risk for developing PTSD symptoms after being attacked by mental patients, Flannery instituted a CISM-like program at an inpatient facility in Massachusetts (Flannery et al., 1991). Whenever a staff member is assaulted by a patient, a trained ASAP clinician immediately conducts a one-on-one debriefing with the staff victim. The debriefer assesses the victim's sense of emotional control, social supports, and ability to make sense out of the incident, and then contacts the victim again 3 and 10 days later. If further action is warranted, the victim is referred to a support group comprising staff members who are trying to cope with having been attacked by patients. Referrals for private counseling and family counseling are provided as needed. Occasionally, the assault is so severe that a group debriefing for all ward staff occurs.

Flannery has yet to publish any questionnaire or interview data regarding ASAP's capacity to attenuate assault-related stress symptoms, nor has he conducted an RCT of ASAP's efficacy. Interestingly, in six settings (four residential programs, two community mental health centers), Flannery (2001) has documented a significant reduction in the frequency of assaults on staff following the implementation of ASAP. This decline ranged from 25% to 62%. Flannery acknowledges that these findings are uncontrolled—another variable correlated with the implementation of ASAP might explain the sudden drop in violence against staff. However, ASAP may itself have altered staff behavior, which, in turn, may have inadvertently reduced the likelihood of patients becoming violent.

In summary, the studies we have just described constitute the evidence adduced in support of psychological debriefing (Everly & Mitchell, 1999, pp. 107–129). Because of their methodological limitations, these studies fail to provide a convincing case for the efficacy of debriefing to mitigate distress and prevent posttraumatic psychopathology.

Studies Adduced as Showing Null or Adverse Effects of Psychological Debriefing

Critics of debriefing have cited studies showing that debriefed participants fared no better—or even worse—than their

nondebriefed counterparts on measures of posttraumatic symptoms. In their discussion of the evidence, critics have concentrated on RCTs (e.g., Rose et al., 2001).[4]

Following Mitchell's (1983) model of CISD, Rose, Brewin, Andrews, and Kirk (1999) randomly assigned 157 adult crime victims (118 male) to either psychological debriefing, an educational intervention, or assessment only. All had been victimized during the previous month; 94% had been physically assaulted, and the others had been victims of attempted or completed sexual assault, physical assault, or robbery. One third had suffered severe injuries (e.g., broken bones). The number of participants, however, was only a fraction of those who might have qualified. That is, Rose et al. contacted 2,161 crime victims who potentially qualified for inclusion, but only 243 replied, and of those who replied, only 157 turned out to be eligible for the study.

The debriefing lasted about 1 hr, and was delivered in an individual format. Subjects were urged to describe their traumatic experience in detail, including the facts of the crime and their thoughts and feelings during it. They were encouraged to express negative emotions, such as guilt, shame, and fear, that might otherwise have gone unexpressed. In addition, they received information about common reactions to traumatic events and where to get further psychological help, if necessary. All debriefing sessions were reviewed by the research team to ensure that protocol was being followed. Subjects randomly assigned to the 30-min educational intervention received only information about common reactions, whereas those in the assessment-only condition received only the assessment measures.

Assessments occurred before the intervention, and again 6 and 11 months later. At the 6-month assessment, rates of PTSD were 26%, 23%, and 11% in the assessment-only, debriefing, and educational groups, respectively. Although the rate of PTSD was nearly twice as high in the group that received debriefing as in the group that received only education, there were no statistically significant differences among the groups in rates of PTSD. By the 11-month assessment, rates of PTSD had dropped so low in all groups that statistical analysis was pointless. All groups exhibited marked and significant improvement over time on questionnaire measures of both PTSD and depression. But the groups did not differ significantly in their improvement on these measures. The authors concluded, "No evidence was found to support the efficacy of brief, one-session interventions for preventing post-traumatic symptoms in individual victims of violent crime" (Rose et al., 1999, p. 793).

Conlon, Fahy, and Conroy (1999) randomly assigned survivors of motor vehicle accidents to either psychological debriefing ($n = 18$) or an assessment-only control group ($n = 22$). None of the subjects required hospitalization, and the 30-min debriefing session immediately followed the baseline assessment, occurring an average of 7 days after the accident. The debriefing protocol included information on common responses to traumatic events, and it encouraged subjects to express the thoughts and feelings they had experienced during the trauma. Advice on coping strategies and seeking further help was provided. Both groups improved markedly on both self-report and clinician-rated measures of PTSD symptoms, but there were no significant differences between the groups 3 months after the intervention. Conlon et al. concluded that they failed "to show any prophylactic benefit of PD [psychological debriefing] in trauma victims" (p. 43).

Debriefing advocates often emphasize that CISD is not a "one-off, stand-alone" method, meaning that it is not intended to be administered in a single session. Therefore, to test whether more extensive intervention might help, Carlier et al. (2000) administered three successive debriefing sessions, based on Mitchell's (1983) method, delivered 24 hr, 1 month, and 3 months posttrauma. In fact, regulations in The Netherlands, where this study was conducted, require that three debriefing sessions be offered to any police officer who has been exposed to a critical incident. The research team formed an "external control group" comprising 75 police officers who had been exposed to trauma prior to the introduction of debriefing. The "internal control group" comprised 82 police officers who

4. In their review, Rose et al. (2001) included not only studies on debriefing of individuals exposed to trauma, but also studies on debriefing of individuals who experienced other kinds of misfortunes. For example, Lee, Slade, and Lygo (1996) randomly assigned women who had recently had a miscarriage to either debriefing or no treatment. Two days after the miscarriage, the women were sent questionnaires assessing symptoms of anxiety, depression, and PTSD, as well as reactions to miscarriage. The women in the debriefing group received a 1-hr session in their homes about 2 weeks after the miscarriage. The female psychologist who did the debriefing used a protocol based on Mitchell's (1983) method. In both groups, PTSD symptoms were initially elevated, but dropped sharply by the 4-month assessment. Depression scores were not elevated at either assessment, whereas anxiety scores were elevated at both time points. Although debriefing had no effect on emotional adjustment, debriefed subjects reported it to be helpful.

Although a routine, uncomplicated birthing does not qualify as a traumatic stressor, some researchers have tested whether debriefing new mothers might reduce the incidence of postpartum maladjustment. Lavender and Walkinshaw (1998) randomly assigned women who had just given birth to either a session of debriefing ($n = 56$) or no intervention ($n = 58$). The women had experienced a normal labor and delivery. The debriefing session lasted between 30 and 120 min, and consisted of the mother discussing the labor and delivery, getting any questions answered, and exploring her feelings about the experience. Three weeks later, debriefed mothers were significantly less likely than nondebriefed mothers to score high on a self-report measure of anxiety and depressive symptoms.

Another research group randomly assigned 1,041 women who had just given birth via an operative method (caesarian, forceps, or vacuum extraction) to either debriefing or usual postpartum care (Small et al., 2000). At 6 months postpartum, the proportion of debriefed women who were depressed (17%) was nonsignificantly higher than the proportion of nondebriefed women who were depressed (14%). Scores on seven of eight scales measuring general health status indicated that the debriefed women were doing worse than the nondebriefed women, although the difference was significant on only one scale (emotional role functioning). Despite the failure of debriefing to reduce postpartum depression, 51% of the debriefed subjects endorsed the debriefing session as "helpful," and another 43% rated it as "very helpful."

had refused debriefing either because they regarded the incident as too trivial or because they lacked the time to participate. These two groups were compared with a debriefed group of 82 officers. The sessions were delivered one-on-one by a trained peer debriefer.

Ninety-eight percent of debriefed participants expressed satisfaction with the first two sessions, and the remaining 2% reported some satisfaction. However, ratings of satisfaction were unrelated to participants' number of psychological symptoms, number of sick days off work, or resumption of regular duties (Carlier et al., 2000).

Carlier et al. (2000) assessed participants shortly before the start of debriefing (i.e., pretest), shortly after the first debriefing session (i.e., 24 hr posttrauma), 1 week posttrauma, and 6 months posttrauma (i.e., after the second and third debriefing sessions). One week after the trauma, debriefed participants reported significantly more PTSD symptoms than did nondebriefed participants. There were no differences among the groups at the 24-hr and 6-month assessments. In fact, rates of PTSD symptoms were very low across all groups.

Two studies suggest that debriefing may impede natural recovery from trauma. Bisson, Jenkins, Alexander, and Bannister (1997) randomly assigned hospitalized burn victims to either debriefing ($n = 57$) or an assessment-only control condition ($n = 46$). The debriefing session occurred between 2 and 19 days after the accident, and it lasted between 30 and 120 min. Mitchell's (1983) protocol was used. The partners (usually a spouse) of 16 of the 57 debriefed individuals attended the debriefing sessions, whereas the remaining subjects were debriefed alone. Bisson et al. compared the groups on 10 measures of trauma severity (e.g., percentage of body burned, pain). The debriefing group scored significantly ($p < .05$) higher than the control group on only 1 measure (whether other people were involved in the accident), although nonsignificant trends for 4 other severity measures (ps = .05, .11, .11, and .12) suggested more severe trauma in the debriefing group (i.e., greater percentage of body burned, stressfulness, perceived life threat, and number of days hospitalized, respectively). The authors failed to correct for multiple comparisons, thereby increasing the likelihood of their finding a spurious significant result.

At the initial assessment, although the to-be-debriefed group tended to score slightly higher than the control group on questionnaires measuring anxiety, depression, and posttraumatic stress, the differences were not significant (ps = .77, .43, and .79, respectively; Bisson et al., 1997). At the 3-month assessment, the rate of PTSD, based on clinical interviews, was nonsignificantly higher in the debriefed group than in the control group (21% vs. 15%). At the 13-month assessment, however, the rate of PTSD was significantly higher among debriefed subjects than among control subjects (26% vs. 9%). Moreover, at 13 months, the debriefed group had significantly higher scores on self-report measures of PTSD, anxiety, and depression than did the control group. These significant differences remained even when the researchers controlled statistically for baseline severity of PTSD, anxiety, and depression. Worse outcomes were associated with longer debriefing sessions, more severe burns, higher initial scores on a psychopathology questionnaire, and a shorter period of time between the burn trauma and the debriefing session. Nevertheless, 52% of the participants endorsed debriefing as "definitely useful" (Bisson et al., 1997, p. 79). The authors, however, concluded that even if debriefing is merely ineffective, rather than harmful, "its routine use should be discontinued" (p. 81). That is, continued use of an inert (and possibly harmful) intervention is a waste of time and resources, and impedes discovery of early interventions that actually do reduce risk of subsequent psychopathology.

In the second study showing that debriefing may have harmful effects, Hobbs, Mayou, Harrison, and Worlock (1996) randomly assigned victims of road traffic accidents to either a single debriefing session ($n = 54$) or an assessment-only control condition ($n = 52$). Victims received individual, one-on-one debriefing, not group debriefing. Despite randomization, the debriefing group had a higher score on an index of injury severity. The groups did not differ in terms of baseline severity of PTSD or other psychiatric symptoms. The 1-hr debriefing occurred between 24 and 48 hr after the accident, and involved providing information about common emotional reactions, reviewing the trauma, encouraging emotional expression, and suggesting gradual return to normal travel. At the 4-month assessment, neither group evinced a reduction in symptoms of PTSD, anxiety, or depression. Moreover, the debriefing group had significantly worse scores than the control group on two subscales of the Brief Symptom Inventory (a questionnaire tapping symptoms of emotional disorder). However, 4-month follow-up data were unobtainable for 22% of the debriefing group and 6% of the control group.

This research team reassessed the patients from the original study (Hobbs et al., 1996) 3 years later (Mayou, Ehlers, & Hobbs, 2000). Relative to the control group, the debriefing group reported significantly more PTSD symptoms, general psychiatric symptoms, fear of traveling as a passenger, pain, physical problems, and financial problems. Further analyses indicated that those participants who had initially scored high on the measure of PTSD and were not debriefed improved markedly by the 3-year follow-up. However, those who had originally scored high on the PTSD measure and were debriefed remained highly symptomatic at follow-up. Therefore, debriefing appeared to impede natural recovery from acute PTSD symptoms. Controlling statistically for differences in injury severity, Mayou et al. found that this could not account for the significantly worse outcome of debriefed subjects with high PTSD scores at baseline. This study suggests that individual debriefing may have long-term adverse effects, although the study had limitations, such as a very early intervention, an inability to contact and assess all participants at follow-up, and differences in injury severity between the debriefed and nonde-

briefed groups. Between-group differences in injury severity complicate interpretation of the results. That is, despite the authors' attempt to statistically control for the effect of differential injury severity, it is difficult to rule out completely that the greater injury severity contributed to the poorer outcome in the debriefed group.

Although the RCT is the gold standard for assessing debriefing or any other intervention, RCTs conducted to date have all concerned one-on-one debriefing, not group debriefing. In the studies we review next, although researchers were unable to assign participants to conditions randomly, chance factors usually determined whether a person received debriefing or not (e.g., whether the person was present at work on the day of the debriefing). These non-RCTs do have the virtue of testing a group debriefing format versus a control (no-treatment) condition.

Hytten and Hasle (1989) assessed 39 volunteer firefighters who had undergone a formal debriefing after having fought a major blaze in a hotel in Norway. All but 1 firefighter considered the intervention helpful. There was no significant difference in self-reported PTSD symptoms between the debriefed group and nondebriefed volunteers who opted to discuss their experience informally with their peers. However, the level of PTSD symptoms was low overall.

Following the Newcastle earthquake in Australia, Kenardy et al. (1996) assessed disaster workers and volunteer helpers who had either been debriefed ($n = 62$) or not ($n = 133$). The groups did not differ in either self-reported exposure to threat or postintervention PTSD symptoms. Although 80% of the debriefed group found the process helpful, responses on a general health questionnaire indicated that this group was significantly more symptomatic than the nondebriefed group. Unfortunately, there was no random assignment to groups and no verification that the debriefing protocol was implemented properly.

Using Mitchell's (1983) protocol, Carlier, Lamberts, van Uchelen, and Gersons (1998) provided a single session of group debriefing to 46 police officers who had responded to a plane crash. Each group session had a maximum of 10 participants. The debriefers had received formal training in the Mitchell method, and fidelity to protocol was checked. No predebriefing measures were taken, but structured interviews were done 8 and 18 months after the debriefing. A group of 59 police officers who, by chance, had been unable to attend a debriefing session because of schedule conflicts were designated as the comparison group and assessed as well. Although assignment to groups was not random, the groups did not differ in age, sex, history of previous traumas, activities at the disaster site (e.g., body handling, rescue operations), or desire for debriefing. Although 7% of the total sample had developed acute PTSD shortly after the disaster, by the 8-month assessment, only 2 participants had PTSD; 1 had been debriefed and the other had not. Assessment of PTSD symptoms showed no significant differences between the groups at 8 months. At 18 months, debriefed participants had significantly more PTSD hyperarousal symptoms than did nondebriefed participants, but otherwise the groups were indistinguishable in terms of symptoms.

Debriefing Advocates Respond to the Negative Studies

Studies adduced in support of debriefing are marred by serious methodological flaws. And almost all well-designed studies have failed to confirm the efficacy of debriefing as a means of preventing posttraumatic psychopathology. However, advocates of debriefing argue that these negative studies are characterized by fatal flaws that undermine their probative import. We next scrutinize the merits of these critiques.

Group versus individual debriefing

RCTs have failed to demonstrate that debriefing reduces subsequent psychopathology, and two trials have shown that debriefing may impede natural recovery from trauma. However, none of these RCTs employed group debriefing. According to Everly and Mitchell (1999), such studies do "not warrant consideration" (p. 125).

There are several problems with Everly and Mitchell's dismissal of RCTs on individual-format debriefing. First, "in clinical practice individual debriefing is the rule rather than the exception" (van Emmerik et al., 2002, p. 767). Second, debriefing advocates (e.g., Mitchell, 2002) approvingly cite one-on-one debriefings if they regard the results as favorable (e.g., Campfield & Hills, 2001; Flannery et al., 1991). It is unclear why these studies are deemed probative, whereas individual-format studies yielding disappointing results are not. Third, it is unclear why an intervention would be inert (or harmful) when delivered in a one-on-one format, but efficacious if delivered in the presence of other people. A group may mobilize social support, but it may also inhibit frank expression of thoughts and feelings, especially among one's coworkers. A further risk in group debriefing is the possibility of vicarious traumatization. Listening to the gruesome details of the event experienced by coworkers may worsen group members' own distress. Fourth, if group debriefing were, in fact, efficacious, then nonrandomized studies comparing groups of debriefed subjects with matched groups of nondebriefed subjects should confirm the superiority of group debriefing. However, such studies have indicated either no effect on PTSD symptoms (e.g., Hytten & Hasle, 1989) or slightly adverse effects (e.g., Carlier et al., 1998; but see Wee et al., 1999, for a possible exception).

Debriefing as a stand-alone technique

Everly and Mitchell (1999) have claimed that CISD was never "meant to be a stand alone technique" (p. 93). There are several problems with this criticism, however.

First, according to Mitchell (1983), a single CISD session "will generally alleviate the acute stress responses which ap-

pear at the scene and immediately afterwards and will eliminate, or at least inhibit, delayed stress reactions" (p. 36). Indeed, after conducting a meta-analysis of several nonrandomized studies (all included in our previous discussion) in which CISD was a single, stand-alone intervention, Everly et al. (1999) affirmed "the power of the psychological debriefing technology to mitigate symptoms of psychological distress" (p. 232). Second, as Rose et al. (1999) observed, debriefing is almost always applied as a single-session intervention rather than as a component in some larger package. Third, Carlier et al. (2000) administered three spaced debriefing sessions, and found small, but adverse, effects. Fourth, debriefing advocates often claim that researchers must test CISM rather than merely CISD—one of its components. However, CISM is not an intervention at all. It is an umbrella term embracing diverse techniques that are relevant to different contexts. CISM includes activities ranging from in-service continuing education programs to outreach support programs for families whose loved ones perished in the line of duty to programs providing coffee and doughnuts to relief workers at disaster sites. Hence, unlike CISD, CISM is not a crisis intervention itself, but rather an administrative framework.

Debriefed people appreciate debriefing

Debriefing advocates correctly state that most debriefed people appreciate the experience. This is true even in studies showing adverse effects of debriefing (e.g., Bisson et al., 1997; Carlier et al., 2000). Therefore, we believe that consumers' satisfaction ratings apparently reflect polite expressions of gratitude rather than intervention efficacy.

Furthermore, there are other possible explanations for the perceived helpfulness of debriefing. For example, people usually feel better at follow-up and may attribute this to the debriefing, not knowing that, on average, they likely would have been just as well if they had not been debriefed. If they had this information, their helpfulness ratings may well be different. Also, justification of effort is a well-known psychological phenomenon. When people are made to comply with unpleasant tasks, they later rate those tasks more positively the less external justification they had for doing them.

Researchers have used the wrong measures to evaluate debriefing

According to Mitchell (2002), researchers who have reported negative outcomes have made the "egregious" (p. 18) error of "using treatment outcome measures (dependent variables such as reductions in clinical depression and symptoms of Posttraumatic Stress Disorder) instead of crisis intervention outcome measures (adaptive function, return to work, lower sick time utilization). That very fact indicates that the researchers have confused crisis intervention and psychotherapy" (p. 18).

There are several problems with this objection. First, using measures of PTSD to evaluate the preventive impact of a crisis intervention does not mean that researchers have confused crisis intervention with psychotherapy. (Ironically, in view of this objection, Mitchell, 2002, himself cited as supportive of debriefing studies of treatments that fail to count as crisis interventions. For example, the intervention studied by Chemtob et al., 1997, was delivered long after the trauma, and the intervention studied by Busuttil et al., 1995, was integrated into a residential group psychotherapy program). Second, elsewhere he stated that the goal of his approach is "to reduce the incidence, duration, and severity of, or impairment from, traumatic stress" (Everly & Mitchell, 1999, p. 72). Accordingly, researchers have assessed PTSD symptoms as an index of impairment from traumatic stress. If crisis intervention following traumatic stressors is a form of secondary prevention, then it ought to prevent the emergence of PTSD. Third, Mitchell (2002) does not object to using PTSD symptoms as an outcome measure when he believes that the study shows favorable effects of debriefing (e.g., Chemtob et al., 1997; Wee et al., 1999). It is unclear why measurement of PTSD symptoms is appropriate when the results appear favorable, but inappropriate when the results appear unfavorable.

Inappropriate participants

Mitchell (2002) criticized certain debriefing studies for including primary victims of trauma (i.e., not emergency service personnel), but he directed this critique at studies with negative outcomes only and cited studies on primary victims approvingly when he believed the results supported his approach (e.g., Amir et al., 1998; Campfield & Hills, 2001; Wee et al., 1999). This seems inconsistent: If debriefing advocates believe that primary victims should be excluded from studies on the efficacy of debriefing, then they should not support the efficacy of debriefing by citing studies on primary victims that yielded results they deem favorable.

Other departures from protocol

When researchers fail to confirm the efficacy of debriefing in controlled studies, debriefing advocates reply that proper protocol was not followed. Negative studies can then be dismissed as irrelevant.

There are several problems with this criticism. First, it presupposes the efficacy of the specific protocol. To say that a departure from recommended protocol, such as failing to debrief in groups, is responsible for the null results, one must first document that the specific protocol is, indeed, effective. Debriefing advocates seemingly believe that one is entitled to assert the efficacy of debriefing until scientists "prove" that it does not work. This logic is exactly backwards: The burden of proof lies squarely on the shoulders of those claiming the efficacy of a specific protocol. Only when a specific protocol has been shown to be effective is one entitled to complain when researchers depart from it.

Second, in most of the negative studies, the researchers did use the "Mitchell model" of debriefing, albeit often with one-on-one debriefings. And some of the studies cited in support of CISD depart even more dramatically from the recommended protocol (Busuttil et al., 1995; Chemtob et al., 1997) than do the RCTs on individual debriefing (Bisson et al., 1997).

CONSIDERATIONS FOR CRISIS INTERVENTION

The Right Time to Talk About the Trauma

Studies showing null effects for psychological debriefing motivate reexamination of a belief shared by many trauma specialists: that expressing thoughts and feelings about the trauma hasten healing, and that "bottling up" these feelings will impede recovery. Some evidence supports this view. Pennebaker and his colleagues have found that repeated writing about one's thoughts and feelings concerning a very upsetting personal event has positive long-term effects on one's mood and health (e.g., Pennebaker & Beall, 1986). Conversely, attempts to avoid thinking about one's trauma and to avoid reminders of trauma are associated with persistent PTSD symptoms (e.g., Ehlers et al., 1998). Furthermore, most trauma therapies emphasize the importance of talking about one's feelings and thoughts about the trauma.

These research findings seem to suggest that helping people ventilate their emotions soon after a critical event will hasten recovery from posttraumatic stress. However, the problem with this inference is that this research was done weeks, months, or years after the trauma, and thus may not apply to the immediate aftermath of an event. Indeed, as Pennebaker (2001) emphasized, his research focused on the psychobiological benefits of writing about traumatic events that had remained undisclosed for months or years. Hence, Pennebaker's work cannot be adduced in support of psychological debriefing that occurs shortly after the traumatic event.

What do people (most of whom will recover on their own) actually do to process a traumatic event? They appear to alternate between phases of avoidance and phases of processing (e.g., Horowitz, 1986; Pennebaker & Harber, 1993). Furthermore, if given a choice, only about 10% of trauma survivors seek to discuss their experience with mental health professionals (e.g., Rose et al., 1999). In the days and even weeks after a traumatic event, "an individual may or may not be in a state in which he or she wishes, or is prepared, to discuss what has happened" (Raphael, Wilson, Meldrum, & McFarlane, 1996, p. 466).

Professionals working with trauma survivors may have too quickly concluded that the initial disinclination of survivors to discuss their trauma constitutes a form of dysfunctional avoidance likely to hinder recovery. The intermittent processing favored by most survivors may adaptively enable them to begin rebuilding their lives and to concentrate on the practical problems they face, and thereby help them to put the event in the past. Furthermore, memories tend to fade with time, and it remains untested whether very early exposure to traumatic memories promotes or retards this process. Research has shown that certain conditions are necessary to facilitate emotional processing of distressing material: "The material, especially in the early stages of treatment, should be made predictable, controllable, presented in small chunks, and tackled in a progressive but gradual way" (Rachman, 2001, p. 166). These conditions are seldom met in the immediate aftermath of trauma. Thus, encouraging survivors to discuss their thoughts and feelings right away may increase the risk that they will be overwhelmed by the experience, which will be counterproductive. Furthermore, as Rachman (2001) has pointed out, there are several routes to emotional processing, and the activation of the trauma memory by reliving the experience may be only one of them.

Thus, contrary to a widely held belief, pushing people to talk about their feelings and thoughts very soon after a trauma may not be beneficial. Perhaps systematic exposure to the trauma memories should be reserved for people who fail to recover on their own. Similarly, Brewin (2001) concluded that

> any intervention that is carried out within two or three days following a mild trauma, or within a month following a severe trauma, is probably coinciding with natural recovery processes. An obvious concern is that the intervention should interfere as little as possible with these processes, at least until some hindrance of recovery is evident. (p. 166)

Thus, clinicians working with trauma survivors soon after the event face a dilemma. On the one hand, any intervention they offer should not interfere with natural recovery. On the other hand, they will want to offer treatment as soon as possible to those survivors who are unlikely to recover on their own, to shorten their suffering and to prevent the development of secondary problems such as job loss, problems with relationships, or substance abuse. In Identification of Individuals at Risk for Chronic PTSD, we address how best to identify trauma survivors who are unlikely to recover on their own.

The practice of talking about a traumatic event shortly after its occurrence has a long historical tradition in military settings. The principles of proximity, immediacy, and expectancy (PIE) have often governed early intervention in the military (Artiss, 1963). Distressed soldiers are treated close to the battlefield (proximity), as soon as possible (immediacy), and with full expectation that they will return to duty (expectancy). The treatment seldom involves more than providing food, rest, and reassurance that they will be feeling better soon. Although these principles have often been accepted as useful in military contexts, the PIE approach has seldom been evaluated rigorously. In one ambitious study, Solomon and Benbenishty (1986) studied troops involved in the Lebanon War. Some were managed according to PIE principles, and others were treated some distance from the battlefront. Solomon and Benbenishty reported that troops managed according to the principles of PIE