Evaluating in traumatic contexts: Considering the contextual, ethical, emotional, and political aspects

Carol Mutch and Annie Weir

This article is drawn from the personal experience of two evaluators working in a post-disaster context. While the evaluators, also the authors of this article, were both experienced in a range of evaluation settings, they found that little in the evaluation literature prepared for them for working in such a physically exhausting and emotionally draining context. The article is set out in three parts. First, the authors synthesise the literature on evaluating in disaster and other traumatic contexts. Secondly, they outline the evaluation they undertook for The Salvation Army in the wake of the 2010/2011 Canterbury earthquakes. Finally, the authors reflect on their experiences by discussing the evaluation context, design, eth-ics, implementation and outcomes in order to add to our growing understanding of evaluating in traumatic contexts.

Evaluating in traumatic contexts: Considering contextual, ethical, emotional, and political aspects

In the past few decades, there has been a rise in the number of evaluations conducted in emergency contexts, such as post-disaster or post-conflict settings. Those involved in such evaluations note the importance of the wider evaluation community learning from these experiences in order to provide vulnerable populations with more effective and accountable services (Alexander, 2005; Few, McAvoy, Tarazona, & Walden, 2014; Puri, Aladysheva, & Iversen, 2015; Ritchie & MacDonald, 2010a, 2010b). Many of the evaluations described in the literature are conducted by supranational organisations or global aid agencies with many years' experience in the development context who arrive in the emergency setting with explicit protocols to guide their approach (for example, the United Nations, World Bank, or Red Cross). More recently, however, traumatic events arising from both natural and human causes have impacted increasingly on developed nations (Ferris & Petz, 2012). In these cases, first responders and disaster recovery agencies are first mobilised from *within* the disaster context. Conducting evaluations in traumatic contexts is difficult at the best of times, but when the responders are also the victims, it adds an extra layer of complexity. The 2010/2011 earthquakes in and around the city of Christchurch and the Canterbury region of New Zealand are a case in point.

The Salvation Army commissioned a team, including the authors, to evaluate their 2010/2011 Canterbury earthquakes response and recovery efforts. The Salvation Army had already been active in the community, especially in the hard-hit eastern suburbs of Christchurch, but they went on to play a significant role in post-earthquake recovery across Canterbury. As evaluators, we found that the existing guidelines and protocols for evaluations in disaster contexts did not

prepare us for the layers of contextual, ethical, emotional, and political complexity we would face when conducting the evaluation. The guidelines focus on what evaluators *should* do. It was difficult to find discussion of how these guidelines actually work in practice. In particular, the literature had very little to say on the impact of evaluating in traumatic contexts for the two sets of key participants—those being evaluated and those conducting the evaluation. Our article sets out to begin to redress this concern, especially for the latter.

First, we canvass the relevant literature on evaluating in traumatic contexts. Next, we share the evaluation of The Salvation Army as a contextual case study, outlining the purpose, methodology, implementation and key findings. The third section of the article is our post-evaluation reflection on what made this evaluation different to those we normally conduct in non-traumatic contexts. Our reflective discussion revolves around: (a) understanding the context; (b) designing the methodology; (c) considering the ethics; (d) navigating the evaluation; (e) negotiating the outcome; and, finally, (f) reflecting on what it means to be an evaluator in an emotionally fraught, ongoing, and uncertain traumatic setting.

Literature review

There is a wealth of literature on how to undertake rigorous evaluations but our interest was in the more limited pool that focused on conducting evaluations in traumatic contexts. Within this pool there is very little literature that reflects on or theorises what happens in the field. The 2010 issue of *New Directions for Evaluation*, edited by Ritchie and MacDonald, is one exception. Thus, to understand this emerging field it was necessary to also explore the broader disaster literature, including the grey literature (guidelines and reports from supranational organisations and international aid agencies), and the literature on researching in sensitive contexts.

An analysis of the literature produced six themes that are relevant to this article: the nature of disaster and emergency contexts; the challenges of working in these contexts; the limited pool of appropriate methodologies; the importance of conducting these evaluations; the need to act in a heightened ethical manner; and a call for more collaboration and sharing of best practice.

The nature of disaster contexts

The United Nations (2008, p. 3) defines a disaster as: "A serious disruption of the functioning of a community or a society causing widespread human, material, economic or environmental losses which exceed the ability of the affected community or society to cope using its own resources." Mutch (2014, p. 6) outlines common themes across a range of disaster definitions as: "suddenness, unexpectedness, lack of preparedness, size of the event and ensuing damage, inability of existing systems to cope, large-scale death or dislocation, and often lack of immediate access to food, water, shelter and medical aid." The disaster evaluation literature notes that the size and scope of disaster and emergency events are often unprecedented, while contexts are unpredictable and volatile. As Ritchie and MacDonald (2010a, p. 4) note: "Disaster and emergency preparedness, response and recovery are complicated at the best of times, often chaotic, and driven by countless impulses and requirements. There is no precise blueprint for organising the many facets of these activities."

The challenges of disaster contexts

Evaluators can find themselves "working in dynamic situations with fluctuating insecurity and population movements", and thus "the evaluation of such dynamic, context-specific responses is therefore complex" (Bornemisza, Griekspoor, Ezard, & Sondorp, 2010, p. 24). There are challenges in setting up, implementing, and completing evaluations in these large-scale and fluid contexts. Commissioners

need to scope what is possible, time the evaluations carefully, and recruit skilled evaluators. Evaluators need to select and implement appropriate evaluation methods, ensuring that they engage affected populations in culturally and ethically responsive ways. Finally, decisions need to be made about responsibility for follow-up and ongoing monitoring (Bornemisza et al., 2010; Few et al., 2014; Steinke-Chase & Tranzillo, 2010; Walden, 2014).

Many of the issues are both logistical and ethical (Puri et al., 2015). The need to deliver aid might take precedence over data collection. Access might be difficult and place an extra burden on overstretched resources such as transport and accommodation. Evaluators are also faced with the ethical dilemma of approaching traumatised communities who might still be searching for or burying their loved ones, or trying to salvage what they can from destroyed homes and businesses (Rencoret, Stoddard, Haver, Taylor, & Harvey, 2010; Walden, 2014).

Each disaster context is also unique and contains varying "social, demographic, cultural, legal, environmental, and technological challenges" (Ritchie & MacDonald, 2010a, p.108). The literature stresses that there is no one way to conduct an evaluation in a development or emergency context (International Organisation for Migration, 2002; Ministry of Foreign Affairs of Denmark, n.d.; Puri et al., 2015; Rencoret et al., 2010; Todd & Todd, 2011). What is possible and meaningful must also be weighed against the trade-offs and compromises needed in such volatile and shifting contexts, for example, "the value of local participation may cut across the benefits of speedy action" (Todd & Todd, 2011, p.vii).

Evaluating in disaster contexts

The range of evaluation approaches in post-disaster contexts is similar to those in regular circumstances. The International Federation of Red Cross and Red Crescent Societies (IRCF, 2011), for example,

use a range of evaluations in disaster contexts focusing on process, outcomes, or impact. The evaluations might examine individual organisations, clusters of organisations, sectors, or cross-sector responses. They might be at the qualitative, thematic analysis, or empowerment end of the scale. They might be more quantitative, including evaluation syntheses or meta-analyses. How these approaches play out in difficult post-disaster contexts is what differs (Buttenheim, 2009; Few et al., 2014; Puri et al., 2015; Ritchie & MacDonald, 2010b; Spence & Lachlan, 2010; Steinke-Chase & Tranzillo, 2010; Walden, 2014). Spence & Lachlan (2010, p. 104) explain: "Unpredictable circumstances may necessitate the use of atypical practices in terms of design, data collection and analysis, and certain conventions concerning collection and analysis may be called into question." Baseline data is not often available for comparative purposes and, as the context changes, planned activities may become redundant. Transient populations make large-scale surveys difficult and it is not often possible or ethical to create control groups for experimental designs. Nor does the unprecedented and fluid nature of the situation lend itself to being measured easily against benchmarks and indicators (Bornemisza et al., 2010; Brusset, Cosgrave, & MacDonald, 2010; Buttenheim, 2009; Few et al., 2014; Janis, Stiefel, & Carbullido, 2010; Puri et al., 2015; Ritchie & MacDonald, 2010b; Spence & Lachlan, 2010; Steinke-Chase & Tranzillo; Walden, 2014).

Evaluations in disaster contexts can be ex-ante, formative, real-time, summative, or ex-post. Ex-ante and early formative evaluations, such as vulnerability or risk assessments and disaster preparedness reviews, can provide useful information, but because they deal with the hypothetical they do not accurately forecast the complexity and unpredictability of the event when it happens (Bornemisza et al., 2010). Ongoing formative and real-time evaluations give more immediate feedback to those on the ground to inform decision making,

but they too have limitations, described by Brusset et al. (2010) as: staff who already have heavy workloads; little available documentation; the need for speedy data gathering; and limited time for a thorough analysis.

Most evaluations in disaster contexts, however, are conducted after the event. Bornemisza et al. (2010) recommend these should take place within 3–6 months of an acute major crisis or after 12–24 months in an ongoing or chronic situation. Most common approaches are formal summative evaluations, sometimes called After-Action Reviews, in which those involved in the response, often with the assistance of an outside facilitator, analyse what happened, how and why, with a view to improving systems and processes (Brusset et al., 2010). Common data-gathering methods are: key informant interviews (both audio and video); document, visual, and electronic source analysis; focus groups or community discussions; mobile methods, such as transect walks or sector mapping; observations; and, where possible, surveys using questionnaires or relevant assessment tools (Brusset et al., 2010; Few et al., 2014; Janis et al., 2010; Morris, 2014; Puri et al., 2015; Rencoret et al., 2010; Spence & Lachlan, 2010). Ex-post evaluations are those that are undertaken some considerable time later to assess ongoing impact or sustainability.

Evaluation ethics in disaster contexts

Because of the physical, emotional, social, cultural, and emotional vulnerability of disaster victims, evaluators must take even more care to act in a safe and ethical manner (van Zijll de Jong et al., 2011). The National Institute of Mental Health (2007) suggests evaluators consider the vulnerability of the potential participants and their capacity to make sound decisions or give informed consent. Evaluators also need to consider the risks and benefits of the disaster survivor's participation.

While physical, emotional, and psychological safety appear the most obvious concerns, evaluators also need to consider participants' age, gender, religion, socio-economic and social status, cultural or ethnic affiliation (including indigeneity), languages spoken, and their ability or disability (Puri et al., 2015; van Zijll de Jong et al., 2011; World Bank, n.d.). There is also a political dimension to consider, most obviously in post-conflict situations or zones with high political tension (Morris, 2014; van Zijll de Jong et al., 2011), but also the politics of aid distribution and the day-to-day power politics of communities and organisations. Discussing their post-disaster situation, van Zijll de Jong et al. (2011) outline some of the complexities: "We understood that we would be addressing fundamental issues of violence, death, destruction, fear, grief, loss of identity, loss of community, foreign interventions, and financial aid, the negation of truth and power relations in [this] post-disaster context" (p. 181).

The importance of evaluations in disaster contexts

The literature stresses that, despite the difficulties, it is important to conduct evaluations to ensure accountability, transparency, relevance, and effectiveness of disaster or emergency response and recovery processes. As Ritchie and MacDonald (2010a) state: "At stake are millions of vulnerable people victimized by disaster, both in developing and developed countries. People should be able to rely on first responders and humanitarian assistance systems to deliver in a timely, effective and appropriate manner" (pp. 3–4). The Australian government's disaster recovery evaluations, as an example, are guided by principles that cover the usefulness of the evaluation to all parties involved and the feasibility of conducting evaluations in difficult conditions. They stress the importance of approaching evaluations ethically, rigorously, and in a participatory manner, "recognising that evaluation should occur 'with' participants and stakeholders rather

than just being done 'to' them" (Institute of Child Protection Studies (2007, p.8). Finally, learning should be widely shared.

As Ritchie and MacDonald (2010b) conclude:

> Difficult and dangerous though it may be to conduct an evaluation of disaster assistance either during a relief effort or soon after, it is a critical component of emergency management. Without assessing the effectiveness of aid, funders, providers, and people on the ground will be unable to make informed decisions about efficient delivery of assistance in the current crisis, or learn lessons in anticipation of the next disaster. (p. 108)

Disseminating best practice

In reviewing emergency and disaster programmes to date, many writers highlight the need to learn from the many evaluations that have been undertaken over the years, and to record and disseminate best practice. Alexander (2005) recommends a comprehensive review of emergency planning as any errors "will be paid for in avoidable casualties and damage" (p. 160). He argues for common standards to better match resources with need, anticipate and tackle foreseeable contingencies, improve clarity and articulation, avoid inefficiencies, and ensure that systems do not become dysfunctional. The IFRC (2011) is one agency that has taken up this call and produced a set of criteria, standards, and processes to guide their evaluation work.

Alexander (2005) notes the need for multi-organisational plans with "minimum levels of accountability, compatibility and functionality" (p.160). Other writers also argue for less fragmentation and duplication and call for increased interagency collaboration (Bornemisza et al., 2010; Brusset et al., 2010; National Institute of Mental Health, 2007; Ritchie and MacDonald, 2010a; Wulf, 2012).

Going beyond the evaluation literature

While there is a growing body of advice for conducting evaluations

in traumatic contexts, our experience after the Canterbury earthquakes was that nothing in the literature prepared us as evaluators for the ethical dilemmas and emotional impacts of doing this work. The literature tends to take the stance that an evaluator is neutral and objective (that is, detached and unemotional), and evaluations are conducted *on* or *for* abstract participants or populations. We needed to turn to the qualitative research literature on working in sensitive contexts to get an understanding of how to navigate our way through a volatile setting where the ethics and politics of conducting high-stakes evaluations were fraught with emotion and tension. We note, in hindsight, that there are other bodies of literature that do exist which we could have consulted in this context, such as trauma counselling or social work. This is an area for further exploration in the future.

Van Zijll de Jong et al. (2011) note how little discussion on the realities of working in disaster zones appears in the scientific and social-sciences literature. They raise questions about how researchers can be supported as they negotiate topics such as grief, loss, destruction, mental and physical damage, and loss of community. They suggest that researchers need relevant experiences of proactive approaches to responsible research. These should include examples of responsible behaviour, standards of integrity, and procedures for reducing risk in researching sensitive topics. Researchers need to consider strategies for meeting the safety needs of all participants, including the researchers.

There are clearly established ethical protocols for protecting participants from any harm that might ensue because of their engagement in research (Dickson-Swift, James, Kippen, & Liamputtong, 2007; Dickson-Smith, James, & Liamputtong, 2008; McCosker, Barnard & Gerber, 2001). Textbooks abound with discussions of such matters; ethical clearance committees scrutinise proposals with

care (McCosker, Barnard & Gerber, 2001). Writers note that doing sensitive research involves more than ethical concerns in relation to the participants, it also presents dilemmas of selection of participants, gaining access, entering the field, conducting the research, leaving the field, analysing data and disseminating findings (Connolly & Reilly, 2007; Dickson-Swift et al., 2007; Dickson Swift et al., 2008; Watts, 2008). Sensitive research also impacts strongly on the researcher, and writers claim this is not given due attention (Connolly & Reilly, 2007; Dickson-Swift et al., 2007; Dickson Swift et al., 2008; McCosker, Barnard & Gerber, 2001; van Zijll de Jong, 2011; Watts, 2008).

These writers note some of the challenges researchers face as understanding what it means to enter into the lives of others, maintaining boundaries, and dealing with disclosure beyond the needs of the research. It also requires awareness of social and cultural nuances and an ability to develop rapport. Often researchers are left feeling a sense of responsibility beyond the research as well as coping with physical and emotional exhaustion. As Watts highlights: "Research that has the added dimension of profound sensitivity may also present the researcher with the challenge of managing the impacts on them of emotional stress caused by watching people's discomfort and suffering" (Watts, 2008, p. 3).

When engaged in this kind of work, Connolly and Reilly (2007) found their identities were constantly shifting. They needed to be aware of the stresses on both parties. They felt the need to keep to the adage "do no harm", but Connolly notes how it was important for her participants to talk about their experiences to make sense of the event but then she became the "repository for the participant's feelings and emotions". She continues, "Unlike a psychotherapist who first hears and then assists the traumatized victim navigate through his or her recovery process, I heard their experiences and then was left to hold or bear their stories" (p. 529).

Some writers even note inherent dangers, such as personal safety, forming attachments, compassion fatigue, or becoming desensitised (Dickson-Swift et al., 2007; Watts, 2008). Yet, these writers also stress that building an empathetic rapport is essential to sensitive research. Such a rapport values the experience of the participant and enables the researcher and researched to make an emotional connection, which acts "as a 'doorway' to the inner terrain of people" (Watts, 2008, p. 8). This leads to more authentic data (Watts, 2008). Watts (2008) notes, "Emotional difficulties experienced by participants in telling their stories with, for example, tearful episodes and problems with fragmented discussion of events, have alerted me to significant aspects of accounts that I may have overlooked without this emotional engagement" (p. 8).

When the data gathering involves the recalling and possible reliving of traumatic incidents, and takes place in an insecure and uncertain environment such as ongoing conflict or disaster recovery, the usual understanding of the roles of evaluator and participant become fluid and repeatedly negotiated (van Zijll de Jong et al., 2011). Add outlining the evaluation purpose, reviewing the evidence, and coming to an eventual judgement in a politically charged context, and the evaluation relationship has further layers of complexity.

Advice on what we could expect to encounter when conducting an evaluation in a traumatic context was largely absent from the evaluation literature. We felt there was a need to articulate what we had faced—the emotional exhaustion of participants and evaluators, the constant renegotiation of roles between counsellor and confidante to data-gatherer and evaluator, and the need for contextual awareness and political astuteness.

The remainder of this article sets the scene for our disaster-evaluation experience and then provides the opportunity for us to share this experience so that others can be better prepared for what they might face in similar situations.

The Salvation Army evaluation case study

The evaluation context

Much is now written about the devastating sequence of earthquakes that hit the city of Christchurch and surrounding districts of Waimakariri and Selwyn in New Zealand's Canterbury region in 2010 and 2011 (see, for example, Aydan, Ulusay, Hamada, & Beetham, 2012; Canterbury Earthquakes Royal Commission, 2012). The earthquakes severely damaged the transport, water, power, and sewerage infrastructure, ruined the city's centre, and damaged over 100,000 homes and 1200 inner city multi-storey buildings. The toll on people was immense, with 185 people losing their lives, over 4000 seriously injured, and many families and businesses dislocated for several years. While the effects were far-reaching across the region, the eastern suburbs of Christchurch, where many of the lower socio-economic suburbs are located, were particularly hard-hit. It is in this area of the city that much of The Salvation's Army's work was regularly undertaken before the earthquakes.

The Salvation Army's response to the earthquakes, in particular, the September 2010 and February 2011 Christchurch events, was an enormous and complex logistical undertaking. It involved the management of over a thousand staff and volunteer personnel, the development of extensive disaster response and recovery services, and the administration of millions of dollars of contracted and fundraised resources.

The formation, implementation, and maintenance of earthquake responses and recovery services, particularly after the February 2011 earthquake, extensively stretched The Salvation Army's organisational capacity. Managing the aftermath of the earthquakes was a steep learning curve for everyone involved and there were valuable lessons that were gained from reviewing The Salvation Army's response.

The evaluation

To achieve an objective and fair assessment of their response to the earthquakes, The Salvation Army commissioned an independent evaluation.

The evaluation contract was awarded to Impact Research New Zealand, who then worked in a collaborative manner with senior Salvation Army staff to formulate the overall evaluation design, and to identify the necessary documentation and interview participants.

The purpose was to investigate what was done in the case of the earthquakes and what could be done to ensure that The Salvation Army's response to future disasters would be appropriate and effective. At all times the approach needed to be consistent with its mission statement: "Caring for people, transforming lives and reforming society through God in Christ by the Holy Spirit's power".

The evaluation had seven objectives:

· To record the activities that formed The Salvation Army's immediate, medium- and long-term earthquake response.

· To assess The Salvation Army's disaster response and recovery strategy.

· To review the effectiveness of the governance arrangements that monitored the implementation of The Salvation Army's disaster response and recovery strategy.

· To assess The Salvation Army's quality of connectivity with other agencies.

· To comment on The Salvation Army's communications, externally and internally, following the two earthquakes.

· To assess the financial performance and outcomes which resulted from the expenditure of The Salvation Army Earthquake Fund.

· To provide suggestions that would support The Salvation Army's response to future disaster events.

A literature review was undertaken to identify key research findings in relationship to best-practice disaster management. The evaluation project then proceeded with two phases of data collection. In Stage 1, a document review and analysis was undertaken assessing a range of earthquake-response documents from internal memos and e-mails to formal reports and financial documents.

In Stage 2, a mixture of face-to-face and telephone interviews were undertaken with ten senior Salvation Army staff and three key informants selected from government, local government and NGO personnel involved in the Canterbury earthquakes. All interviews were audio-recorded and transcribed. A thematic pattern analysis was performed on all interviews and emerging themes were then organised into broad categories. Selected quotes from respondents' accounts were used to illustrate the themes. Care was taken to de-identify all quotes to ensure interviewee confidentiality. The interview data were then triangulated with the document review data and literature review to confirm the emergent theme findings.

Evaluation findings

The relevant literature is clear that the provision of professional disaster response and recovery services requires a pre-established disaster management plan, strong relationships with other disaster response organisations, and a clear and well understood disaster governance structure. To prepare for events there needs to be pre-existing policies and protocols to guide the recruitment and behaviour of staff and ongoing staff capacity building. Finally, in the event of a disaster, there need to be organisational management practices that support staff working or living, or both, in a disaster situation (Simpson, Clegg, & Pina e Cunha, 2013).

Since the first earthquake struck the Canterbury region in September 2010, The Salvation Army has been a key non-government

organisation (NGO) provider of support services and resources to the people of Christchurch and the outlying regions. The strengths of The Salvation Army's disaster response and recovery work included its ability to respond to the two earthquakes in a quick and timely manner, the capacity to supply extensive on-the-ground staff and volunteer personnel at short notice and over an extended period of time, and the logistical capacity to organise extensive support services and resources at short notice.

An internal review of The Salvation Army's response to the initial 2010 earthquake revealed that the organisation responded (at the request of the Ministry of Civil Defence & Emergency Management) by providing over 20,000 meals for earthquake victims and disaster response personnel. At the same time The Salvation Army transformed its local counselling service into an earthquake-orientated psychosocial service (in response to a request from the Ministry of Social Development), established a strategic partnership with World Vision and the Westpac Bank to support fundraising initiatives, and expanded its local Community Ministries service into a large-scale provider of food packages, vouchers, and general welfare services.

When the second earthquake struck in February 2011, with its significant consequences of loss of life and extensive damage to property and infrastructure, The Salvation Army again initiated its provision of the earthquake response services (as an expansion of the September 2010 services). In total, The Salvation Army's response included the provision of 27,000 meals, 6500 care packages, 11,000 food parcels, 8600 vouchers and debit cards, 250 respite holidays for stressed families, and the mobilisation of over a thousand Salvation Army personnel who visited tens of thousands of affected households providing social, spiritual, and material support.

As the Canterbury earthquake effort progressively evolved into a recovery and rebuilding operation, The Salvation Army modified its

earthquake services to include the formation of a range of community cohesion initiatives, including the establishment of neighbourhood support groups in areas where demolitions and migration of residents had brought about social dislocation and isolation, and the expansion of a Schools Support Programme to assist primary school children and their wider communities. The Salvation Army Employment Plus[1] also established an employment initiative, U-Build-4-the-Rebuild, to provide construction training and to help place unemployed individuals into full-time work, and partnered with other social service agencies and the Christchurch City Council to fund an affordable housing project for earthquake-affected families and the elderly.

The evaluation findings demonstrate that, throughout the traumatic and chaotic events, The Salvation Army remained a learning organisation. They had a strong desire to improve their practice in order to be the best they could be in responding to disasters. Their members, who are strongly bonded by their Christian beliefs, displayed a sense of great dedication to humanity. As such, the community could trust and rely on them in times of disaster. This made a profound difference in Canterbury. The Salvation Army modelled practical compassion, and the teams supporting those in greatest need made a real difference.

While the lessons learned by and about The Salvation Army are of interest in themselves to those who wish to know more about resilient organisations, the focus of this article now turns to reflecting on the process of conducting evaluations in traumatic circumstances.

Our reflections

Ongoing debriefing was an important aspect of the evaluation, but the opportunity to reflect more systematically came when we decided

---

1   In April 2014 Salvation Army Employment Plus changed its name to Salvation Army Education and Employment.

to present at the Aotearoa New Zealand Evaluation Association conference. After an initial brainstorm of ideas, observations, emotions, and reflections, we were able to distil our comments into relevant themes. Feedback from the presentation and further reflection has reduced those wider themes into those that follow: (a) understanding the context; (b) designing the methodology; (c) considering the ethics; (d) navigating the evaluation; (e) negotiating the outcome; and (f) reflecting on what it means to be an evaluator in a traumatic context.

Understanding the context

It is an expectation that evaluators will take time to understand the context in which they are working. However, the context in which we found ourselves was so unprecedented in our experience—and the experience of our participants—that there was no blueprint to guide us. Other factors included the constantly changing and volatile situation. Earthquakes differ from many other disasters in that there is no warning and no clear endpoint. The main events themselves were traumatic enough without the constant aftershocks contributing to the constant state of hyper-alertness and anxiety. A further complexity was that, in order to understand what had happened, why the organisation had responded in the way it had, and what had driven their decision making, we needed our participants to revisit the emotionally charged events of the disaster, their immediate aftermath, and the ongoing state of recovery. The Salvation Army had taken on a wide range of familiar and unfamilar tasks, and the people responding on the ground were often victims of the earthquakes themselves. Allied to this, The Salvation Army has a hierarchical structure, which works well in normal times when there is time for consultation and delegation but, in the volatile post-earthquake setting, this structure created tensions between local, regional, and national decision

making. This meant there was even more complexity in the way participants responded to us and the way we engaged as evaluators.

One aspect that helped us to understand the context as best we could was the careful selection of the team. The team included an "insider" (who had experienced the earthquakes) and "outsiders" (who had not, but had links to Christchurch in a way that heightened their sensitivity). In this way the team had credibility with the participants and could examine the evidence with empathy and objectivity. It was also important that the team were skilled evaluators with a range of experiences to draw on, as it was not the context in which to induct novices. Another important factor was the time the lead evaluator took to understand the organisation, its ethos, its way of working, and its layers of operation. Holding preliminary discussions on the purpose and design of the evaluation also contributed to building these understandings.

Designing the methodology

Designing the methodology was a collaborative activity between The Salvation Army, as the commissioners, and Impact Research New Zealand, as the evaluators. The Salvation Army had already signalled what they hoped to achieve in their call for tenders. In selecting our proposal they already had confidence in our skills and experience. From there it was important to ensure that the methodology would help them achieve their aims. They understood the need to benchmark their activities against best practice in disaster response, and the literature review afforded that opportunity. Next, they made a range of documents available for analysis—daily field reports, email communications between staff, quarterly, and annual board reports, financial reports related to earthquake fundraising and expenditure, website press releases, and two low-level review reports completed by the Emergency Services. This enabled the external evaluation team

to provide them with a comprehensive review of their response and recovery decision making. They also wanted to hear from key personnel within the organisation and from outside who had worked closely with them throughout the earthquakes. The best approach to gain a range of opinions on how well they responded and what they could do to improve was to undertake in-depth interviews. An interview protocol was agreed upon to provided consistency across the interviews but with the understanding that there could be flexibility if the need arose. The Salvation Army identified both internal personnel (senior managers) and external stakeholders (e.g. Civil Defence) who held leadership roles during the response, recovery, and rebuild phases. Ten key Salvation Army senior staff were interviewed. Five were based in Christchurch and five were based nationally. In addition, two key senior stakeholders from the head office of the Ministry of Social Development and another from the Ministry of Civil Defence were interviewed.

An important aspect of the methodology was triangulating the data. The international literature provided information on best practice in disaster response, highlighting key concepts around communication, relationships, and authority in relation to disaster management. The document review and the in-depth interviews of senior Salvation Army staff and external stakeholders could then be matched against this best practice and one another to highlight strengths and pinpoint weaknesses.

Considering the ethics

As many writers suggest, conducting an ethical evaluation requires making ethical judgements throughout the process. This process includes responding to the commission, designing the approach, implementing the evaluation, analysing the evidence, reaching conclusions, and presenting the findings (Simons, 2006). In the case of

this evaluation, The Salvation Army was keen to learn how well they had undertaken their role and how they could improve in the future. A heartfelt sense of dedication to genuine improvement meant that they willingly shared what had worked well and where they could improve. They trusted us as evaluators to use their experiences and concerns to make recommendations that could drive change within the organisation. As evaluators, we still needed to balance the different individual organisational perspectives to produce a comprehensive but fair outcome—one that took the nature of the context into account, yet provided sound advice for moving forward.

The lead evaluator focused on establishing and sustaining trust between the team and the organisation. This required shared understanding of the purpose, structure, culture, and nuances of the organisation, including the spiritual basis for their work. The evaluators brought credentials in terms of contextual awareness, their skill as evaluators, and their willingness to engage in a participatory process. The Salvation Army brought a genuine commitment to their mission and to their role as a learning organisation.

Once the evaluation approach was agreed and the list of possible participants established, contacting the interviewees could begin. Initial contact by phone, followed by a more detailed email, set the scene for clarifying the purpose, explaining the approach, and assuring the participants that the interviews would be conducted in an emotionally safe and sensitive manner.

The interviews themselves were emotionally charged affairs. Participants who had bottled up their emotions, frustrations, and memories found the telling of their stories very cathartic. As some participants shared their stories they began to unpack the depth and layers of their own experiences, and those of their colleagues and those whom they were serving. They also reflected on how their roles had expanded beyond what they had been "called" to do in the past,

extending them well beyond their comfort zones. They also commented that the earthquake experiences had a profound effect on their lives. As responders whose mission was to serve others, they had suppressed their own feelings. When given someone else's full attention, they poured out their suppressed emotions. It was not our job to be counsellors, but we did have to listen empathetically without being drawn too deeply into their emotions. This often meant that they opened up about frustrations with the wider organisation. As evaluators we had the ethical dilemmas of what was "on the record" or "off the record". We needed to acknowledge each person's perspective. The participants wanted their story to be valued and our interest to be genuine. From the individual recollections, we would aim to synthesise a fair and realistic portrayal of the issues. We were stretched beyond our normal sense of ethics to a heightened sense of a personal duty of care, but no matter how much individuals' stories remained with us, we needed to step back and see where they fitted with the bigger picture that we were assembling.

Navigating the evaluation

Given the complexity of the evaluation, we were constantly negotiating our way through the implementation of the evaluation. The nature of the roles of many participants, within and outside the organisation, meant getting a response, gaining access, and agreeing on a time and location was somewhat haphazard. The interviews were not able to be conducted in any particular order, thus sense-making was ongoing and iterative with new themes being constantly added and reviewed.

As these interviews were emotionally charged they required the interviewer to balance empathy with objectivity—which we described as "stepping into their shoes and then out again". How did we make sense of the individual's story and how did this fit into a

larger web of emerging themes? During both the interviews and the analysis we needed to keep peeling back the layers, looking for the obvious and the not-so-obvious, the said and the unsaid, the surface explanation and the deeper meaning. Being sensitive to the post-disaster tensions within the organisation and allowing each participant to feel they were listened to in a non-judgemental way also required careful navigation.

Much of the sense-making relied on the increasing understanding of the political context. The political tensions were apparent at all levels, not just those within the organisation. There were tensions between national and local government over who managed the earthquake recovery—CERA (Canterbury Earthquake Recovery Authority) or the Christchurch City Council. There were jurisdiction tensions between all the agencies working in earthquake-affected Canterbury—first-response agencies, relief agencies, government departments, NGOs, and newly emerging grassroots organisations. We needed to understand The Salvation Army's response within this larger web of complexity. We became more aware of the taken-for-granted, the nuances, and the pieces of the puzzle that did not always seem to fit easily with the emerging story. There was a need to keep true to what was developing as the essence of the story but to also keep reflecting with one another and debriefing with the organisation.

Throughout this time, the Christchurch-based interviewees continued to live with the earthquakes. We felt for them on a very personal level, especially as we came to see that many of the systemic issues were outside their control. It was also important to acknowledge the toll on the evaluation team—providing logistical support, managing schedules, and debriefing interviews all contributed to helping us keep our task in perspective.

Negotiating the outcome

Despite being in awe of what The Salvation Army had achieved, we had a job to do. We needed to prepare a report that gave a fair and balanced account of the strengths and weaknesses of what we found. The report had to honour the voices of the participants and clearly articulate their experiences, but we had to take care not to further strain relationships within the organisation. The issues needed to be clearly outlined and framed in a way that would have impact yet be accepted and acted upon. We needed to get to the heart of the matter without simplification. The audit trail of evidence leading to our findings needed to be clear and defensible. The tone needed to convey our understanding of the context in which they worked, acknowledge the outstanding contribution The Salvation Army had made to the Canterbury earthquake response and recovery efforts, yet suggest ways forward for improvement. We were also conscious that the report might be read by multiple audiences inside and outside the organisation and therefore had to take care to express issues of concern in a non-judgemental and future-focused manner.

The key to being able to produce a report that met the many intended goals, criteria, and audiences was to maintain the participatory relationship. This meant touching base regularly, updating progress, asking questions and clarifying concerns, engaging the organisation in sense-making, and discussing tentative findings.

The success of the evaluation overall has been in what has happened since. The Salvation Army has accepted the report and its findings. The recommendations are now being implemented nationally.

Concluding reflections on evaluating in a traumatic setting

In evaluation work it is important to have a general understanding of the context of any intended evaluation. The evaluation literature did not adequately prepare us for evaluating in a traumatic context. We

are now more aware that our experiences resonate with the limited literature on this topic. These situations are fluid and volatile. They present a range of contextual and practical problems. The context is highly political, with power struggles over funding, resources, roles, and responsibilities. Any work is emotionally draining and ethically challenging. We have presented our reflections so that others might be better prepared for such situations although, as the literature suggests, each context is unique and will present particular challenges.

In summary, we found conducting an evaluation in a traumatic context, such as the aftermath of a disaster, takes more time and needs more flexibility than in regular settings. It is difficult to quantify this: as evaluators we were constantly reviewing the situation and adjusting as we went. Because of the emotional, contextual, and political sensitivity of our context, the evaluation required careful navigation. While projecting an air of compassionate neutrality, we were constantly balancing empathy and objectivity. We were trying to understand the context from very different perspectives at the same time as testing our emerging findings. We feel a traumatic context is not a place for inexperienced evaluators. It needed all the knowledge, skill, and facilities that we had amassed over the years. To be able to build relationships, call on experience, solve problems quickly and effectively, or find practical solutions was imperative in the fluid and uncertain context. Finally, we found it was exhausting for all parties. Consideration of the toll that working in these contexts can take needs to be built in to the design through ongoing debriefing, reflection, and review.

Yet despite the difficulties we faced, we learned so much. We learned about the strength of the human spirit; we learned about ourselves. We feel privileged to have been able to give acknowledgement to The Salvation Army's extraordinary humanitarian response to the most vulnerable in Christchurch and beyond. They went above and

beyond their call to duty to walk beside their communities in their darkest days. We are honoured to tell their story.

References

Alexander, D. (2005). Towards the development of a standard in emergency planning. *Disaster Prevention and Management: An International Journal 14* (2) 158–175. http://dx.doi.org/10.1108/09653560510595164

Aydan, O., Ulusay, E., Hamada, M., & Beetham, D. (2012). Geotechnical aspects of the 2010 Darfield and 2011 Christchurch earthquakes, New Zealand, and geotechnical damage to structures and lifelines. *Bulletin of Engineering Geology and the Environment*, *71*, 637–662. http://dx.doi.org/10.1007/s10064-012-0435-6

Bornemisza, O., Griekspoor, A., Ezard, N., & Sondorp, E. (2010). The Interagency Health and Nutrition Evaluation initiative in humanitarian crises: Moving from single-agency to joint, sectorwide evaluations. *New Directions for Evaluation, 126*, 21–35. http://dx.doi.org/10.1002/ev.327

Brusset, E., Cosgrave, J., & MacDonald, W. (2010). Real-time evaluation in humanitarian emergencies. *New Directions for Evaluation*, *126*, 9–20.

Buttenheim, A. (2009). *Impact evaluation in the post-disaster setting: A conceptual discussion in the context of the 2005 Pakistan earthquake.* Retrieved from: http://www.3ieimpact.org/

Canterbury Earthquakes Royal Commission. (2012). *Final report* (Vols 1–7). Christchurch: Canterbury Earthquakes Royal Commission.

Connolly, K. & Reilly, R. C. (2007). Emergent issues when researching trauma: a confessional tale. *Qualitative Inquiry*, *13*(4), 522–540. http://dx.doi.org/10.1177/1077800406297678

Dickson-Swift, D., James, E., Kippen, S., & Liamputtong, P. (2007). Doing sensitive research: what challenges do qualitative researchers face? *Qualitative Research*, *7*(3) 327–353. http://dx.doi.org/10.1177/1468794107078515

Dickson-Smith, V., James, E., & Liamputtong, P. (2008). *Undertaking*

sensitive research in the health and social sciences: Managing boundaries, emotions and risks. Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511545481

Ferris, E., & Petz, D. (2012). *The year that shook the rich: A review of natural disasters in 2011*. London, UK: The Brookings Institution, London School of Economics.

Few, R., McAvoy, D., Tarazona, M., & Walden, V.M. (2014). *Contribution to change: An approach to evaluating the role of intervention in disaster recovery*. Rugby, UK: Practical Action Publishing. http://dx.doi.org/10.3362/9781780448114

Institute of Child Protection Studies. (2007) *Project report. Evaluation framework: Australian Government Disaster Assistance Package*. Dickson, ACT: Australian Catholic University.

International Federation of Red Cross and Red Crescent Societies. (2011). *IFRC framework for evaluation.* Retrieved from www.ifrc.org

International Organisation for Migration. (2002). *Monitoring, evaluation and audit in emergencies: Technical reference paper.* Geneva, Switzerland: Office of the Inspector General (OIG) and the Emergency and Post Conflict Division (EPC).

Janis, A., Stiefel, K. M., & Carbullido, C. C. (2010). Evolution of a monitoring and evaluation system in disaster recovery: Learning from the Katrina Aid Today National Case Management Consortium. *New Directions for Evaluation*, *126*, 65–77. http://dx.doi.org/10.1002/ev.330

McCosker, H., Barnard, A., & Gerber, R. (2001). Undertaking sensitive research: Issues and strategies for meeting the safety needs of all participants. Forum: *Qualitative Social Research/SozialForschung*, *2*(1). Retrieved from http://www.qualitative-research.net/index.php/fqs/index

Ministry of Foreign Affairs of Denmark. (n.d.). *Evaluation guidelines*. Retrieved from http://www.netpublikationer.dk/um/7571/index.htm.

Morris, B. (2014). Creating contextualised standards for emergency programming: Lessons learnt. *Ontrac. The newsletter of INTRAC:*

International NGO Training and Research Centre 58 (November), 5–6.

Mutch, C. (2014). The role of schools in disaster preparedness, response and recovery: what can we learn from the literature? *Pastoral Care in Education*, *32*(1), 5–22. http://dx.doi.org/10.1080/02643944.2014.880123

National Institute of Mental Health. (2007). *Ethical issues to consider in developing, evaluating, and conducting research post-disaster.* Retrieved from http://www.nimh.nih.gov/funding/grant-writing-and-application-process/ethical-issues-to-consider-in-developing-evaluating-and-conducting-research-post-disaster.shtml

Puri, J., Aladysheva, A., & Iversen, V. (2015). *What methods may be used in impact evaluations of humanitarian assistance?* Bonn, Germany: Forschungsinstitut zur Zukunft der Arbeit/Institute for the Study of Labor.

Rencoret, N., Stoddard, A., Haver, K., Taylor, G., & Harvey, P. (2010). *Haiti earthquake response: Context analysis.* London, UK: ALNAP.

Ritchie, L. A., & MacDonald, W. (2010a). Enhancing disaster and emergency preparedness, response, and recovery through evaluation. *New Directions for Evaluation*, *126*, 3–7. http://dx.doi.org/10.1002/ev.325

Ritchie, L. A., & MacDonald, W. (2010b). Evaluation of disaster and emergency management: Do no harm, but do better. *New Directions for Evaluation, 126*, 107–111. http://dx.doi.org/10.1002/ev.333

Simons, H. (2006). Ethics in evaluation. In I. Shaw, J. Greene, & M. Mark, (Eds.), *Handbook of Evaluation: Policies, programs, practices.* Thousand Oaks, CA: Sage. http://dx.doi.org/10.4135/9781848608078.n11

Simpson, V., Clegg, S., & Pina e Cunha, M. (2013). Expressing compassion in the face of crisis: Organizational practices in the aftermath of the Brisbane floods of 2011. *Journal of Contingencies and Crisis Management, 21*(2), 115–124. http://dx.doi.org/10.1111/1468-5973.12016

Spence, P. R., & Lachlan, K. A. (2010). Disasters, crises, and unique populations: Suggestions for survey research. *New Directions for Evaluation, 126*, 95–106. http://dx.doi.org/10.1002/ev.332

Steinke-Chase, M., & Tranzillo, D. (2010). Save the Children's approach to

emergency evaluation and learning: Evolution in policy and practice. *New Directions for Evaluation*, *126*, 37–49. http://dx.doi.org/10.1002/ev.328

Todd, D., & Todd, H. (2011). *Natural disaster response: Lessons from evaluations of the World Bank and others.* Washington, DC: Independent Evaluation Group, The World Bank.

United Nations. (2008). *Disaster preparedness for effective response: Guidance and indicator package for implementing priority five of the Hyogo framework.* New York/Geneva: United Nations.

Van Zijll de Jong, S., Dominey-Howes, D., Roman, C. E., Calgaro, E., Gero, A.,Veland, S., … Afioga, T. L. (2011). Process, practice and priorities — key lessons learnt undertaking sensitive social reconnaissance research as part of an (UNESCO-IOC) International Tsunami Survey Team. *Earth-Science Reviews*, *107*, 174–192. http://dx.doi.org/10.1016/j.earscirev.2011.03.001

Walden, V.M. (2014). Contribution to change: An approach to evaluating the role of intervention in the disaster recovery process. *Ontrac: The newsletter of INTRAC: International NGO Training and Research Centre 58* (November) 2–4.

Watts, J. H. (2008). Emotion, empathy and exit: Reflections on doing ethnographic qualitative research on sensitive topics. *Medical Sociology Online, 3*(2), pp. 3–14.

World Bank. (n.d.). *Gender informed monitoring and evaluation in disaster risk management*. Washington, DC: The World Bank.

Wulf, J. (2012). *A balanced scorecard for the humanitarian sector? Adaptability of the balanced scorecard model to sector-wide performance management in humanitarian aid feasibility and implications*. Retrieved from: http://www.rub.de/ifhv/6-publications/6-workingpapers.html

The authors

Dr Carol Mutch: Associate Professor and Head of School, Critical Studies in Education, Faculty of Education and Social Work, The University of Auckland.
Email: c.mutch@auckland.ac.nz

Dr Annie Weir: Honorary Research Fellow, School of Critical Studies in Education, Faculty of Education, University of Auckland. Director, Impact Research NZ.
Email: a.weir@auckland.ac.nz