ACADEMIC ARTICLE Evaluation Journal of Australasia Vol 17 | No 2 | 2017 | pp. 29–36

RALPH RENGER | JIRINA FOLTYSOVA | SKYLER IENUSO | JESSICA RENGER | WAYNE BOOZE

Evaluating system cascading failures

This article shares methods used to evaluate system cascading failures. A cascading failure occurs when a problem is passed from one subsystem to a downstream subsystem creating a domino effect that undermines system efficiency and effectiveness. First, the basics of system evaluation theory (SET) are reviewed. Then drawing on different examples from the evaluation of emergency response systems the article describes how a) standard operating procedures (SOPs) can be used to locate possible system cascading failure trigger points, and b) mock exercises and secondary data are used to evaluate these trigger points. The discussion highlights the need to expand SET's conceptualization to include within subsystem cascading failures in addition to between subsystem cascading failures. The extent to which program evaluation methods can be adapted for use in system evaluation is also discussed.

Methods for evaluating system cascading failures

Many evaluators are exploring the value of system thinking to improve program evaluations (Renger, Wood, Williamson & Krapp, 2011; Williams & Hummelbrunner, 2010). The basic premise is systems thinking is a way to address the artificialities of many theory driven program evaluation approaches (Williams & Hummelbrunner, 2010). For example, system thinking is thought to better capture the complex context in which a program operates thereby addressing the limitations of oversimplified program logic models (Gamel-McCormick, 2011). The goal of evaluators applying system thinking is to produce more meaningful and usable program evaluations (Patton, 2008; Williams & Hummelbrunner, 2010).

However, Renger (2015; 2016) notes another evaluation branch is emerging within theory-driven evaluation that is using systems thinking to evaluate modern day systems. Renger (2015) published the SET to guide evaluating modern day systems.[1] SET employs both system thinking *and system theory* to meet the emerging stakeholder demand to evaluate entire systems, of which a program may be one component.

SET suggests evaluators follow three basic steps in conducting a system evaluation: i) define the system, ii) evaluate system efficiency, and iii) evaluate system effectiveness. Defining the system also consists of three steps to capture the detail necessary for evaluating system efficiency and effectiveness. The first step is defining the system boundaries (Renger, 2015; Williams & Hummelbrunner, 2010). This step is critical in establishing the evaluation scope and resources (e.g. stakeholders to include) (Williams & Hummebrunner, 2010). Once system boundaries are established it is then necessary to identify subsystems which are contributing to the common system goal(s). The third step then details within and between subsystem relationships, also known as the SOPs (Nickols, 2000).

Explicit documentation of the SOPs is critical to understanding how each component of the system and interaction between them is supposed to operate.

Once the system is defined, the second SET step is to define system efficiency. Central to this step is operationalizing system efficiency. All system components should work toward the same efficiency goal. For example, Renger (2017) applied SET to evaluating points of dispensing (POD). In this example, the purpose of a POD is to deliver mass immunizations/vaccinations in a public health crisis. A POD consists of numerous interdependent stations (e.g. registration, medical dispensing) all designed

Ralph Renger is a Professor and Director of Evaluation at the Center for Rural Health, University of North Dakota.
Email: ralph.renger@med.und.edu

Jirina Foltysova is an Assistant Professor of Population and Rural Health at Walden University.
Email: kitkafoltysova@gmail.com

Skyler Ienuso is a Graduate Research Assistant at the Center for Rural Health, University of North Dakota.
Email: Skyler.ienuso@und.edu

Jessica Renger is a Research Assistant at the University of Arizona.
Email: Jessica.renger@hotmail.com

Wayne Booze is a Program Officer at Rural Healthcare, Helmsley Charitable Trust.
Email: WBooze@helmsleytrust.org

to process the public as quickly as possible, but at the same rate. In a POD, efficiency is operationalized as the patient/minute ratio. If one station processes patients at a different rate, then this compromises overall system efficiency by creating system backups and surges.

SET also identifies four system attributes affecting system efficiency: leadership; competent and capable system actors; a functional informational technology infrastructure; and a cohesive commitment by all to the system goals (i.e. a shared culture). If any of these system attributes are not in place, then system inefficiencies occur.

SET notes system efficiency also depends on the timely, sufficiently frequent, specific, relevant, and credible sharing of information within and between subsystems via feedback mechanisms. There are two basic types of feedback mechanisms, intra- (i.e. within subsystem) and inter-feedback (i.e. between subsystems) mechanisms. Recently, Renger (2016) published a case example of how to identify and evaluate both types of system feedback mechanisms.

The final SET step is to evaluate system effectiveness. As with the system efficiency goal, all system components should work toward a common effectiveness goal. Of importance to this step are the principles of interconnectedness and wholeness. In the POD example, the overall goal is to maximize patient throughput. All POD stations are dependent on each other in meeting the system goal.

The focus of this article is on the evaluation of one SET system principle affecting both system efficiency and effectiveness: cascading failures (Parsons, 1961; Peters, Buzna & Helbing, 2008). Because a system consists of interconnected subsystems, failure in one subsystem can be passed on (i.e. cascades) to other parts of the system (Ericson, 2011). This domino effect reduces system efficiency and system effectiveness. Using examples from evaluations of emergency response systems we will illustrate some methods for locating, evaluating, and depicting cascading failures. In so doing we hope to meet some of the emerging demand in the evaluation community for methods to bridge SET into practice (Alderman, 2016; Renger, 2016).

Identifying trigger points for possible cascading failures

To identify cascading failures requires an understanding of how subsystems are supposed to function and interrelate to each other. We refer to these processes and procedures as SOPs. System efficiency depends, in part, on the ability to achieve, maintain, and streamline SOPs. SET describes two types of SOPs: those detailing within subsystem processes and those detailing between subsystem processes.

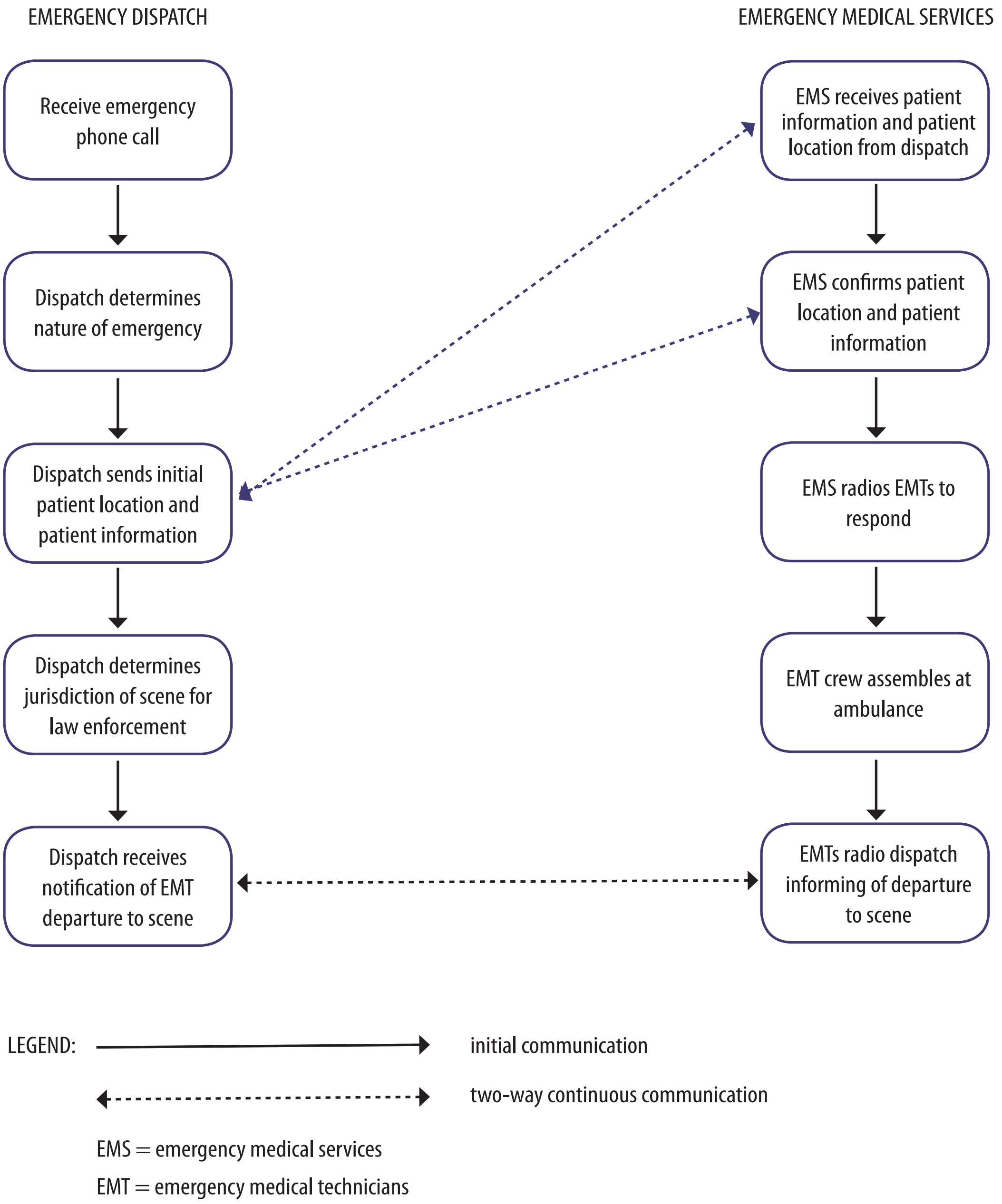
SOPs are important for several reasons. First, they are the basis for training. Often system processes must be completed by human beings, or system actors. SET notes when system actors are capable they are more likely to execute necessary processes efficiently and make meaningful contributions to quality improvement processes to better streamline the SOPs.

Second, SOPs help understand critical points for information technology interfaces. SET notes communication within and between subsystems is often technology dependent, whether it be phones, email, internet, and so forth. The permeation and dependence of information technology for virtually all modern day system communications and functioning is undeniable and inescapable. SOPs help pinpoint where information technology is critical for efficient system functioning.

Third, SOPs enable system continuous quality improvement (Renger, 2016). SOPs describe the *what and how* of system functioning. In practical terms making changes to SOPs is how recommended improvements in system efficiency become operationalized.

Fourth, and key for the system evaluation, SOPs note key points of interaction between subsystems. The efficiency of these interactions can be either human or technology dependent. For example, Figure 1 shows an emergency medical services (EMS) SOP generated using process flow mapping (Renger, McPherson, Kontz-Bartels & Becker, 2016). Emergency medical services is one important subsystem of the response system for time critical events (e.g. stroke, cardiac arrest). An efficient response depends on the emergency medical services' ability to quickly and accurately send and receive patient information from two other subsystems: a dispatch service that coordinates emergency medical

FIGURE 1. CASCADING FAILURE TRIGGER POINTS OF AN EMERGENCY MEDICAL SERVICES SUBSYSTEM



services response (i.e. an upstream subsystem) and the hospital receiving the patient (i.e. a downstream subsystem). Minimizing human error and/or delays in relaying critical information (e.g. expected time of arrival to hospital; patient demographics) from dispatch to emergency medical services and/or from emergency medical services to the receiving hospital is essential in improving patient survival.

Points at which subsystems interact are possible cascading failure *trigger points* because if problems occur at these key transitions, then time delays (i.e. the efficiency goal) cascade throughout the system. Figure 1 illustrates cascading failure trigger points between two emergency response subsystems: emergency medical services and emergency dispatch. The arrows show where in the SOPs patient information is passed from one system to another. For example, emergency dispatch sends an initial communication to emergency medical services notifying them of their need to mobilize. Emergency medical services acknowledges receipt of this information. From this point forward there is a continuous two-way communication to track emergency medical services location and patient status.

Evaluating cascading failures

In our evaluation of emergency response systems we used two methods to evaluate cascading failures. The first was to conduct a mock exercise. The mock exercise followed the Department of Homeland Security guidelines and involved a real time simulation of a cardiac arrest (Granillo, Renger, McPherson, Dalbey & Foltysova, 2014). The focus of the mock exercise was on evaluating two major system attributes at key trigger points and the extent to which they may be contributing to system cascading failures. The first was whether the system actors possessed the needed capability to adhere to the SOPs. The second was to determine whether the system technology met the system actors' need for timely relay of critical response information (e.g. expected time of arrival to hospital, patient vitals).

The evaluation using the mock exercise found both problems with system actor training (e.g. not understanding communication protocols in the transfer of a patient to a hospital; emergency medical services not understanding how to transmit electrocardiogram records) and system interoperability (e.g. different patient record software platforms used by emergency medical services and hospital; inability of emergency medical services and/or hospital to send/receive electrocardiogram records transmissions). Both these factors resulted in significant time delays in patient treatment; delays which significantly jeopardize positive patient outcomes.

The second method for evaluating cascading failures was to conduct a secondary analysis of time data. In a time critical event, like a cardiac arrest, time is the efficiency goal (Eisenberg, 2013). Time data is routinely collected by all of the cardiac arrest subsystems. For example, dispatch tracks the time it takes for emergency medical services to gather at the ambulance station (i.e. chute time), to arrive on scene, and to arrive at the hospital. Hospitals track the time from when the patient enters the emergency bay to the time treatment is received (i.e. door to balloon time). Many of these system processes have established time standards based on research evidence (e.g. door to balloon time should be less than 90 minutes) (Bradley et al., 2006; Eisenberg, Bergner & Hallstrom, 1979; Weaver et al., 1986).

Data analyses consisted of two types, special and common cause (Renger et al., 2016). The former examined individual events and compared the event times to their respective standard. Results from this analysis were used for system-specific quality improvement (e.g. an emergency medical services agency and hospital from the same region). The latter examined data across events comparing the mean time for multiple events to the standard. Results from this analysis helped inform possible changes to policies and/or time standards. For example, data collected from rural emergency medical services agencies showed the time standards need to be adjusted to reflect the longer transport times faced by rural responders due to distance and road conditions.

Both evaluation methods worked well at pinpointing communication breakdowns due to system actors and system technology. However, planning, conducting and evaluating a mock exercise is costly. Depending on the exercise complexity and number of participating subsystems a mock cardiac arrest exercise can cost between US$50,000 and US$75,000. There are also safety concerns as real resources are being mobilized in real time. Thus, the added safety risk needs to be weighed against the training and evaluation exercise benefits (Federal Emergency Management Institute, 2016).

The challenge with the secondary data analyses lies in the trustworthiness of the data. Our analyses found numerous database inconsistencies with respect to the time variable (Renger et al., 2016). For example, many recorded transport times were simply impossible given the distance from the scene to the hospital and the top end ambulance speed.

Depicting cascading failures

Succinctly showing how a problem in one part of a system leads to ripple effects is important for directing decision makers where to allocate resources that will maximize system efficiency and effectiveness. Doing so is also true to the utility standard (Patton, 2008; Sanders, 1994).

We likened the domino effect of cascading failures to the *if-then* logic found in root cause analysis (Chien, Wang & Chen, 2007; Coşkun, Akande & Renger, 2012; Doggett, 2005; McLaughlin & Jordan, 1999; Renger, 2011; Renger & Titcomb, 2002; Venkatasubramanian, Rengaswamy, Kavuri & Yin, 2003). We explored whether the root cause analysis method used in logic modeling and in program theory reconstruction could be used to depict system cascading failures (Foltysova, 2013; Renger and Titcomb, 2002). This approach to root cause analysis was chosen simply because it seemed to intuitively fit the problem at hand and was most familiar to the authors.

To test whether the root cause analysis methodology could be adapted we used source documentation describing system failures in national data collection registries (Renger et al., 2016). Figure 2 shows the root cause analysis mapping results. This model is read from left to right by using *if-then* logic. For example, *if* there is a lack of shared understanding of the purpose of the national cardiac care registry between the leadership at the national, state, and local levels, *then* the leadership is unable to reach an agreed upon data set. *If* the leadership is unable to reach an agreed upon data set, *then* data elements selected for inclusion for national registry are of limited interest and value to those responsible for gathering the information and so forth.

Discussion

The work here supports conclusions from other early applications of SET suggesting mixed methods used in program evaluation such as qualitative interviewing, secondary data analyses, and root cause analyses can also be effectively used in evaluating modern day systems. However, in some cases new methods need to be developed in response to idiosyncrasies of system theory underpinning SET, such as depicting cascading failures. For example, to our knowledge the use of process flow mapping to develop SOPs to then define trigger points for cascading failures is novel. Further, we adapted methods used in emergency preparedness to evaluate system efficiency. As system evaluation grows it is likely the evaluation success will depend on the evaluators' knowledge, familiarity, and comfort in adapting program evaluation methods as well as the willingness to explore and adapt methods from other disciplines.

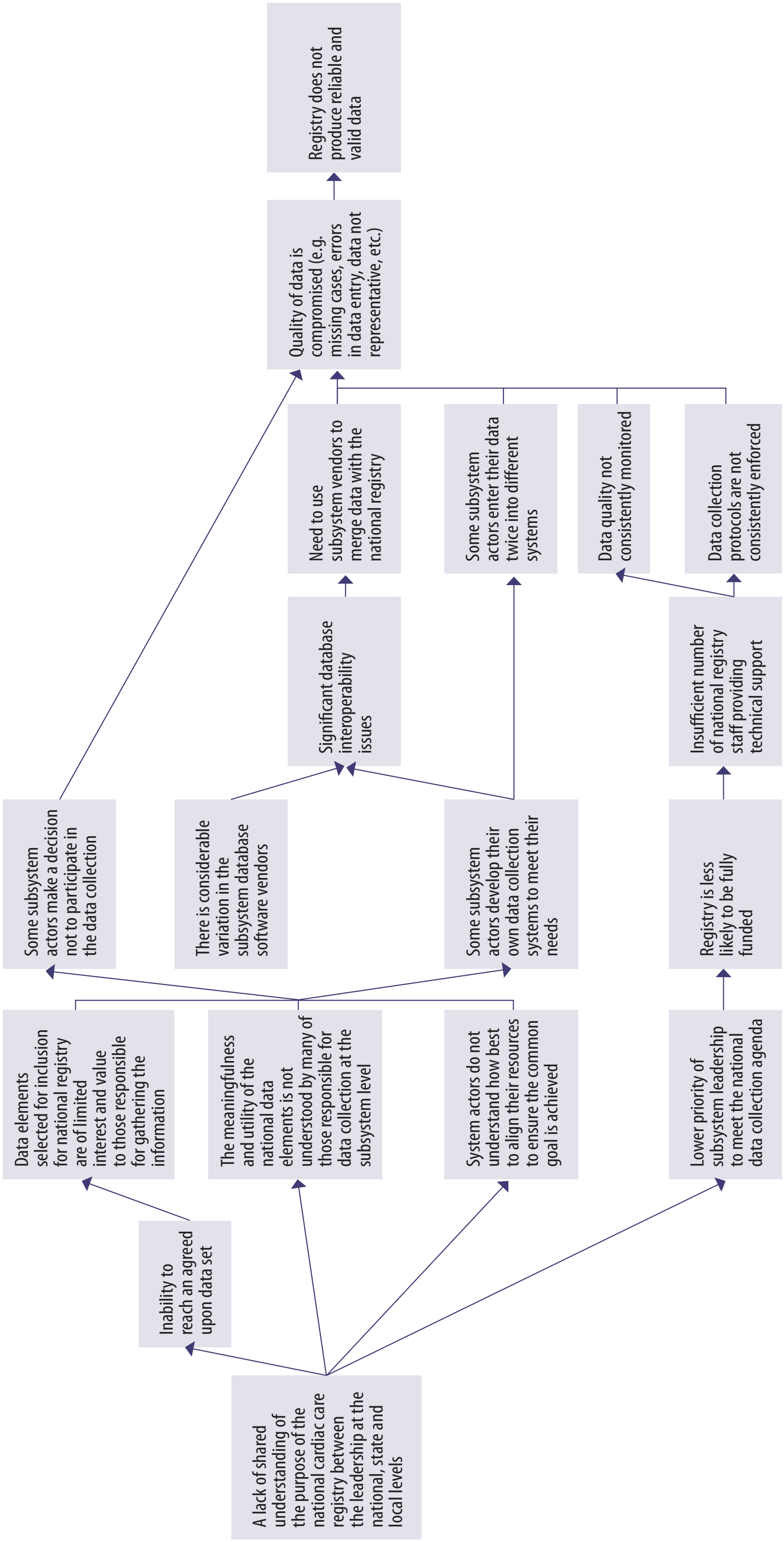
With respect to evaluating cascading failures, developing SOPs using process flow mapping proved useful, however it can be costly, requiring several interviews with subject matter experts until reaching a saturation point. On a positive note our experience is the saturation point is reached sooner than when conducting interviews for program evaluation purposes (Renger & Bourdeau, 2004; Renger, Foltysova, Becker & Souvannasacd, 2015). This seems logical since the goal of system actor training should be the standard delivery of established processes and procedures. For example, there are recommended procedures for treating cardiac arrests, trauma, and so forth. Thus, little variability should exist between system actors performing the same role. Should variability exist, it would suggest further evaluation of the effectiveness of system training is required.

It may be possible to develop SOPs cheaper and quicker using source documentation (Donaldson, 2005; Leeuw, 2003; Renger, 2011). For example, some aspects of treating cardiac arrest, like the two minute CPR cycle, are well documented (Xanthos, Bassiakou, Koudouna & Papadimitriou, 2009). However, in our experience the complete SOPs detailing necessary elements for evaluating system efficiency such as feedback loops and trigger points for cascading failures are often missing. Thus, the only option may be to engage system actors to document the SOPs. Likely the extent of available SOP documentation is system dependent. Highly regulated systems are likely to have more detailed SOPs. However, any existing documentation needs to be scrutinized to determine whether they detail the necessary elements for system evaluation (Foltysova, 2013). Regardless, evaluators should first explore the availability of existing SOP documentation before investing in primary data collection using qualitative methods.

Perhaps most encouraging was that the recommendations forthcoming from the system evaluation of cascading failures possessed high utility. Specifically, stakeholders from all subsystems worked together to address the training and system interoperability (i.e. information technology attribute) issues identified from the evaluation of the cardiac arrest and secondary data analysis. As the consequences of cascading failures were felt throughout the system it motivated all system leaders (i.e. another SET system attribute) to invest in the resources needed to resolve the issue. For example, at the time of writing the mock exercise stakeholders developed a new information technology infrastructure where patient data can now be transferred between dispatch, emergency medical services, and hospital within 17 seconds (Hardeland et al., 2014). This is in sharp contrast to the 20 minute delays discovered during the mock exercise evaluation. It is the authors' belief the motivation for action stemmed from the ability to concretely identify the issue and offer a targeted recommendation. The recommendation was derived directly from the SET principle of cascading failures. This reinforces the importance of theory and Carol Weiss' (1995) famous adage there is nothing so practical as a good theory.

Our work also suggests SET's concept of cascading failures needs to be broadened. During our evaluation of cascading failures using mock exercises we noted two types of cascading failures: within and between subsystems. This directly parallels the concepts of documenting the within and between subsystem

FIGURE 2. CASCADING FAILURES OF A NATIONAL DATA COLLECTION REGISTRY DEPICTED USING ROOT CAUSE ANALYSIS



processes. A cascading failure commencing within a subsystem can then be passed on to other subsystems.

In conclusion, evaluating modern day systems using SET continues to show great promise. SET provides simple guidance in addressing the complexity of evaluating modern day systems. It also leads to concrete recommendations for system change because it is grounded in system theory and thus has high utility. As the application of SET continues it is hoped evaluators will continue to contribute to the system evaluation toolbox by sharing their lessons learned and best practices.

Endnotes

1 Ericson (2011) defines a modern day *system* as 'an integrated composite of components that provide function and capability to satisfy a stated need or objective. A system is a holistic unit that is greater than the sum of its parts. It has structure, function, behavior, characteristics, and interconnectivity. Modern day systems are typically composed of people, products, and environments that together generate complexity and capability' (p. 402).

References

Alderman, L. (2016). Personal communication at Australasian Evaluation Society International Conference. Perth, Australia.

Bradley, E. H., Herrin, J., Wang, Y., Barton, B. A., Webster, T. R., Mattera, J. A., Roumanis, R. N., Curtis, J. P., Nallamothu, B. K., Magid, D. J., McNamara, R. L., Parkosewich, R. N., Loeb, J. M. & Krumholz, H. M. (2006). Strategies for reducing the door-to-balloon time in acute myocardial infarction. *New England Journal of Medicine*, *355*(22), pp. 2308–2320. doi:10.1056/NEJMsa063117

Chien, C., Wang, W. & Cheng, J. (2007). Data mining for yield enhancement in semiconductor manufacturing and an empirical study. *Expert Systems with Applications, 33*(1), pp. 192–198. doi:10.1016/j.eswa.2006.04.014

Coşkun, R., Akande, A. & Renger, R. (2012). Using root cause analysis for evaluating program improvement. *Evaluation Journal of Australasia, 12*(2), pp. 4–14.

Doggett, A. M. (2005). Root cause analysis: A framework for tool selection. *The Quality Management Journal, 12*(4), pp. 34–45.

Donaldson, S. I. (2005). Using program theory-driven evaluation science to crack the Da Vinci code. *New Directions for Evaluation, 106,* pp. 65–84.

Eisenberg, M. S. (2013). *Resuscitate: How your community can improve survival from sudden cardiac arrest*. Seattle: University of Washington Press.

Eisenberg, M., Bergner, L. & Hallstrom, A. (1979). Paramedic programs and out-of-hospital cardiac arrest: I. Factors associated with successful resuscitation. *American Journal of Public Health*, *69*(1), pp. 30–38.

Ericson, C. A. (2011). *Concise encyclopedia of system safety: Definition of terms and concepts*. Hoboken, NJ: John Wiley & Sons.

Federal Emergency Management Institute. (2016). *IS-130: Exercise evaluation and program planning*. Retrieved 25 April 2017 from: https://emilms.fema.gov/IS130/index.htm

Foltysova, J. (2013). *Validation of reconstructed program theory*. Dissertation. Tucson, Arizona: University of Arizona.

Gamel-McCormick, C. A. (2011). *A critical look at the creation and utilization process of logic models*. Masters Thesis. Newark, DE: University of Delaware.

Granillo, B., Renger, R., McPherson, M., Dalbey, D. & Foltysova, (2014). *Redfield-Aberdeen-Sioux Falls Cardiac Arrest Drill After-Action Report/Improvement Plan*. Unpublished technical report. Grant Forks, North Dakota: Center for Rural Health, University of North Dakota.

Hardeland, C., Olasveengen, T. M., Lawrence, R., Garrison, D., Lorem, T., Farstad, G. & Wik, L. (2014). Comparison of medical priority dispatch (MPD) and criteria based dispatch (CBD) relating to cardiac arrest calls. *Resuscitation*, *85*(5), pp. 612–616.

Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation, 24*(1), pp. 5–20.

McLaughlin, J. A. & Jordan, G. B. (1999). Logic models: a tool for telling your programs performance story. *Evaluation and Program Planning, 22*(1), pp. 65–72.

Nickols, P. (2000). Standard operating procedures. *The Quality Assurance Journal*, *4*(2), pp. 91–94.

Parsons, T. (1961). An outline of the social system. *Classical Sociological Theory,* (2), pp. 421–440.

Patton, M. Q. (2008). *Utilization-focused evaluation*. Thousand Oaks, CA: SAGE Publications.

Peters, K., Buzna, L. & Helbing, D. (2008). Modelling of cascading effects and efficient response to disaster spreading in complex networks. *International Journal of Critical Infrastructures*, *4*(1–2), pp. 46–62. doi:10.1504/IJCIS.2008.016091

Renger, R. (2011). Constructing and verifying program theory using source documentation. *The Canadian Journal of Program Evaluation, 25*(1), pp. 51–67.

Renger, R. (2015). System Evaluation Theory (SET). *Evaluation Journal of Australasia, 15*(4), pp. 16–28.

Renger, R. (2016). Illustrating the evaluation of system feedback mechanisms using system evaluation theory (SET). *Evaluation Journal of Australasia, 16*(4), pp. 14–20.

Renger, R. (2017). *Using systems evaluation theory (SET) to improve points of dispensing (POD) efficiency and effectiveness*. Manuscript submitted for publication.

Renger, R. & Bourdeau, B. (2004). Strategies for values inquiry: An exploratory case study. *American Journal of Evaluation, 25*(1), pp. 39–49.

Renger, R., Foltysova, J., Becker, K. & Souvannasacd, E. (2015). The power of the context map: Designing realistic outcome evaluation strategies and other unanticipated benefits. *Evaluation & Program Planning, 52,* pp. 118–125.

Renger, R., McPherson, M., Kontz-Bartels, T. & Becker, K. (2016). Process flow mapping for systems improvement: Lessons learned. *The Canadian Journal of Program Evaluation, 31,* pp. 109–121.

Renger, R., Qin, X., Rice, D., Foltysova, J., Souvannasacd, E., Renger, J., Markwart, B., Hart. G. & Bjerke, M. (2016). *Challenges and solutions facing emergency medical services in supporting the IOM recommendation for a national cardiac arrest registry: A system perspective*. White Paper. Retrieved 25 April 2017 from: https://ruralhealth.und.edu/pdf/ems-supporting-iom-recommendation.pdf

Renger, R. & Titcomb, A. (2002). A three-step approach to teaching logic models. *American Journal of Evaluation, 23*(4), pp. 493–503. doi:10.1016/S1098-2140(02)00230-8

Renger, R., Wood, S., Williamson, S. & Krapp, S. (2011). Systemic evaluation, impact evaluation, and logic models. *Evaluation Journal of Australasia, 11*(2), pp. 24–30.

Sanders, J. R. (1994). *The program evaluation standards: how to assess evaluations of educational programs* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N. & Yin, K. (2003). A review of process fault detection and diagnosis part III: Process history based methods. *Computers & Chemical Engineering, 27*(3), pp. 327–346.

Weaver, W. D., Cobb, L. A., Hallstrom, A. P., Fahrenbruch, C., Copass, M. K. & Ray, R. (1986). Factors influencing survival after out-of-hospital cardiac arrest. *Journal of the American College of Cardiology*, *7*(4), pp. 752–757.

Weiss, C. H. (1995). Nothing as practical as good theory: exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. P. Connell, A. C. Kubisch, L. B. Schorr, C. H. Weiss (eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts* (pp. 65–92). Washington, DC: Aspen Institute.

Williams, B. & Hummelbrunner, R. (2010). *Systems concepts in action: A practitioner's toolkit*. Stanford, CA: Stanford University Press.

Xanthos, T., Bassiakou, E., Koudouna, E. & Papadimitriou, L. (2009). Using the 30: 2 compression–ventilation ratio: Five cycles is easier to follow than 2 min of cardiopulmonary resuscitation. *European Journal of Emergency Medicine*, *16*(6), pp. 339–341.