REFEREED ARTICLE Evaluation Journal of Australasia Vol 16 | No 4 | 2016 | pp. 15–21

RALPH RENGER

Illustrating the evaluation of system feedback mechanisms using system evaluation theory (SET)

This article describes how system evaluation theory (SET) guided the evaluation of cardiac care response systems efficiency in seven rural United States. Specifically, the article focuses on the approach and methods used to evaluate system feedback mechanisms; one key factor affecting system efficiency. Mixed methods were applied to evaluate five criteria of system feedback efficiency: frequency, timeliness, credibility, specificity, and relevance. Examples from the cardiac care response system evaluation are used to illustrate each of the evaluation criteria. The discussion contrasts the role of the evaluator in system versus program evaluation, notes the post-hoc support of SET system attributes in affecting system efficiency, and offers additional consideration in evaluating system feedback mechanisms.

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

Systems thinking is gaining attention in the evaluation literature as an approach for addressing limitations associated with some of the artificialities of theory driven program evaluation (Williams & Hummelbrunner, 2010). For example, systems thinking is helpful in adding context to logic models that may unintentionally oversimplify the context of programmatic assumptions (Gamel-McCorckmick, 2011).

System thinking is also an important element in the evaluation of modern day systems. 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)

Under this definition a program may be a system component, but itself is not a system (McDavid, Huse & Hawthorn, 2013).

Coupling system thinking and system theory, Renger (2015) published the SET to guide the evaluation of complex modern day systems. SET consists of three guiding principles for bridging system theory and evaluation practice: (i) defining the system, (ii) evaluating system efficiency, and (iii) evaluating system effectiveness. Defining the system is necessary before being able to evaluate system efficiency and effectiveness. That is, it is necessary to know the details of the system components and how they are connected before being able to meaningfully evaluate whether it functions efficiently. In turn, system effectiveness is dependent on system efficiency; the greater the system efficiency the better the system effectiveness.

More specifically, SET describes three factors affecting system efficiency including a) system attributes such as leadership, organizational culture, system actor competence and capabilities, and information technology, b) inter- and intra-subsystem communication procedures and protocols, and c) inherent feedback mechanisms.

Feedback mechanisms monitor the environment for anomalies affecting system efficiency (Banathy, 1992; Burns, 2007). Presumably, detected inefficiencies leads the system to engage in corrective actions to achieve and maintain an optimal efficiency level (Banathy, 1992; Brandon, 1990; Flynn, Schroeder & Sakakibara, 1994). One purpose of SET is to determine whether inherent system feedback mechanisms exist and if so, whether they are operating efficiently. The following case example illustrates the various methods used to evaluate cardiac care system feedback mechanisms.

Case background

The cardiac care response systems in seven rural US states were evaluated[1]. A cardiac arrest is one of the most, if not the most, serious time-critical health events (Eisenberg, 2013). The key factor in improving cardiac arrest survival rates is reducing the time to definitive care[2]. At the moment of cardiac arrest the person is clinically dead (Eisenberg, 2013). The likelihood of survival falls 7–10% for every minute of delay in CPR and defibrillation (Carlbom et al., 2014). Within 10 minutes clinical death will progress to irreversible biological death (Brouwer, Walker, Koster & Chapman, 2013; Eisenberg, 2013).

There are multiple subsystems which share the common goal of expediting a patient to definitive care (see Figure 1, page 17). The response begins when the dispatch service, under the jurisdiction of law enforcement, receives an emergency call from a bystander. The goal of dispatch is to find the available emergency medical services (EMS) unit capable of the fastest response and to instruct the bystander how to conduct CPR until the arrival of EMS on scene. The goal of EMS is to stabilize and/or transport the patient to the nearest medical facility. The decision where to transport the patient lies with the attending physician at the hospital affiliated with the responding EMS unit. In rural states this is often a small hospital with limited resources. Thus, the goal of the attending physician is to decide the closest hospital for definitive care. This may be the small hospital or it may be a larger urban facility. Patients initially transported to a small hospital may require another EMS facilitated transport (air or ground) to a larger hospital with better resources, including heart imaging equipment and cardiac care specialists.

The need to reduce time to definitive care is the efficiency goal shared by all cardiac care response subsystems. As will be shown below the efficiency goal plays a crucial role in evaluating system feedback mechanisms.

Evaluating system feedback mechanisms

As per SET a two-step decision process guided the evaluation of the system feedback mechanisms (Renger, 2015). The first step was to ascertain whether inherent system feedback mechanisms were present. It is reasonable to posit in modern day systems needed feedback mechanisms may be missing. In such cases, the evaluation recommendation was to implement the needed feedback mechanisms.

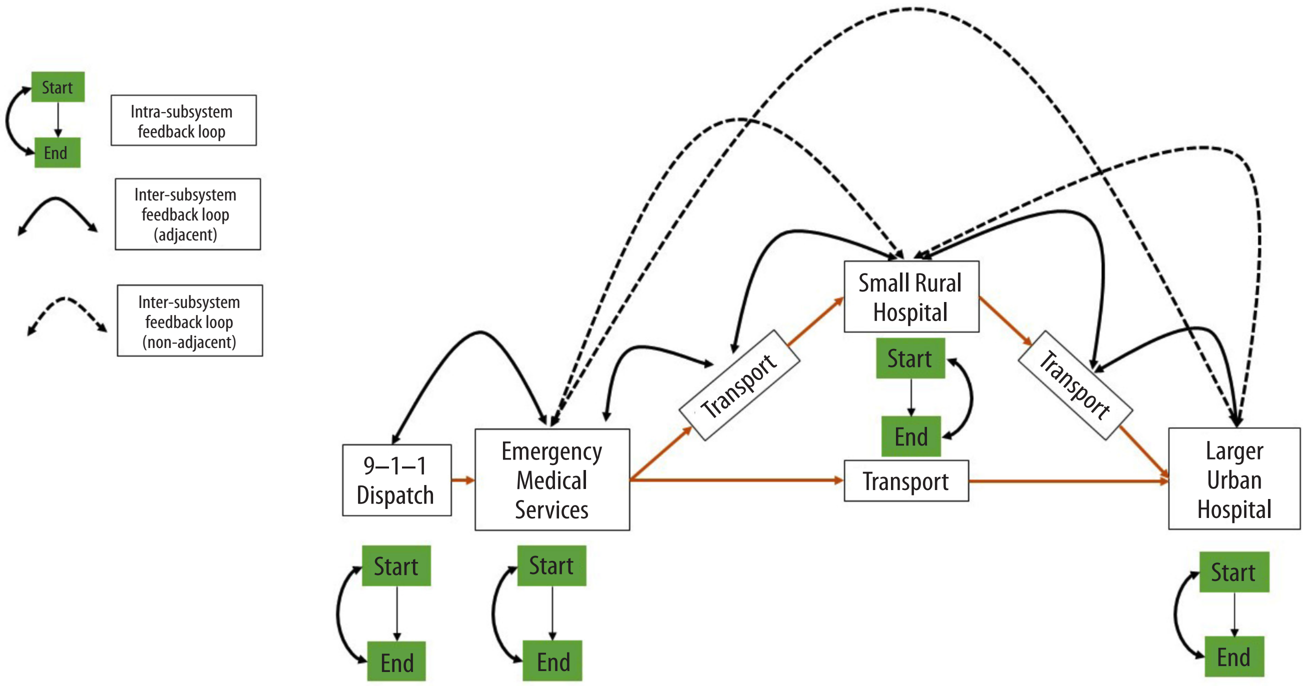
If inherent feedback mechanisms were present, then the second step was to evaluate feedback quality. Feedback was evaluated against five quality criteria: timeliness, frequency sufficiency, credibility, relevance, and specificity (Chen, Hailey, Wang & Yu, 2014). It was assumed if these five criteria are met, then the system feedback mechanism is operating efficiently. The methods used to complete both steps are now detailed.

Identifying inherent feedback mechanisms

The first challenge in evaluating the cardiac care response system feedback mechanisms was ascertaining whether and where they were operating. According to SET, systems can consist of two different feedback mechanism types. The first type is intra-subsystem. The evaluation first determined whether there was evidence a subsystem is monitoring and attempting to improve its own response components independent of other subsystems. The second type is inter-subsystem of which there are two subtypes: adjacent and non-adjacent. For these subtypes the evaluation focused on establishing whether neighboring or non-neighboring subsystems shared information to improve their respective response efficiency.

The evaluation employed mix methods to identify the presence of the different feedback mechanism types. First, the evaluation examined each subsystem database to determine whether they collected time related data (i.e. the efficiency goal). It was reasoned for cardiac care response subsystems to become more efficient (i.e. reduce time) they should be gathering data related to the efficiency goal.

FIGURE 1. CARDIAC CARE SUBSYSTEMS AND POSSIBLE LOCATIONS OF INHERENT INTER- AND INTRA-SUBSYSTEM FEEDBACK MECHANISMS



Many subsystems collected time data. For example, dispatchers gather six standard time stamp data as they track the ambulance to and from the station. Rural EMS track chute time, or the time it takes for the technicians to assemble at the ambulance station. Heart hospitals track the time from when a patient crosses the threshold of the emergency room to the time they receive necessary intervention in the catheterization lab.

Second, interviews with subsystem leaders and system actors[3] revealed they are mandated to review all time critical events as part of their continuous quality improvement (CQI) procedures. That is, CQI was the system actors' term for feedback mechanisms and served as a clue to detect other inherent feedback mechanisms. For example, the evaluation learned CQI routinely occurs for all cardiac arrest events between a medical director (i.e. hospital) and EMS personnel; an example of an adjacent, inter-subsystem feedback mechanism.

Third, using process flow mapping, a qualitative method for documenting standard operating procedures (Renger, McPherson, Kontz-Bartels & Becker, 2016), the evaluation identified inter-subsystems data sharing and communication points. This led to the discovery of health information exchanges, a mechanism for sharing patient records between hospital systems, and between hospital systems and EMS used in three states.

Figure 1 shows where the different feedback mechanism types were found operating within the cardiac care response system.

Evaluating feedback mechanism quality

Feedback specificity

Specific feedback is defined as information particular to the system. The more specific the information the more useful it is to the system actors in making necessary adjustments to optimize efficiency.

Using focus groups and available source documentation (Renger, 2011) it was discovered most cardiac care response subsystem CQI processes lack specificity. One key component to CQI is the availability of standard operating procedures, which detail the specific steps necessary in responding to a cardiac arrest. It is the 'how' of the response.

There are many important reasons for having written standard operating procedures. First, they guide training content and ensure training meets necessary standards (Nickols, 2000; World Health Organization, 2002). Second, CQI focuses on improving processes and procedures. However, if the processes and procedures are not documented, then what to change is often unclear (World Health Organization, 2002). Third, without them, ensuring changes are uniformly applied and sustained is virtually impossible (Nickols, 2000). In short, the more detailed the standard operating procedures the more specific and useful the feedback for improving response efficiency.

Cardiac arrest response standard operating procedures were sporadic, with some elements being well described (e.g. exact timing of CPR cycle), while other aspects of the response were missing. In response to this need the evaluation engaged numerous system leaders and subsystem actors in process flow mapping (Graham, 2004) resulting in detailed standard operating procedures for each subsystem.

Sufficiently frequent feedback

The frequency of feedback is critical for system efficiency. Long delays in potential life-saving information pose ethical problems. Delays also make the information less useful; over time event-specific information is decayed, forgotten, or confused with other events.

Delivering high performance CPR is critical to patient survival, but is fatiguing. In the rural context, this is made more challenging because of the long hospital transport times and the number of available EMS personnel (Grossman et al., 1997). Therefore, EMS personnel are taught to rotate CPR duties (Braithwaite et al., 2014). It is essential that compressions delivered by respective EMS personnel are of the highest quality and the pauses between compressions minimized (Braithwaite et al., 2014). To ensure EMS personnel meet the criteria for high performance CPR, compression quality and pause length statistics are automatically collected by the patient's heart monitor and are reviewable immediately after completing patient transfer. This is an excellent example of how information technology (i.e. one of the SET system attributes) plays a critical role in meeting the criteria of providing sufficiently frequent feedback (Renger, 2015).

However, many of the evaluated system feedback mechanisms had open loops (Clow, 2012). That is, critical information being collected was either not being provided to system actors or was being provided with insufficient frequency for CQI purposes. Interviews with state leadership found five of the seven states did not make the data collected by EMS after completing a call available to the ambulance services, thus leaving the feedback loop open.

Further, using focus groups the evaluation found states often operated on set reporting periods (e.g. monthly, quarterly). These long, predetermined frequency intervals posed significant problems in feedback usability; if several events occurred in these intervals then the likelihood of confusing events increased and recalling events completely decreased.

At the time of writing, several corrective actions were being implemented to improve the frequency of feedback. In one early adopting state, where the state EMS department began providing monthly reports (i.e. began closing the information loop more frequently) submission of complete patient care reports by EMS services doubled, demonstrating the motivational aspects of functional feedback mechanisms.

Feedback relevance

The feedback provided must be related to optimizing system efficiency. Irrelevant feedback creates system noise (Ash, Berg & Coiera, 2004; Slifkin & Newell, 1998). Noise will cause delays in processing feedback as system actors must filter information to determine what is critical for optimizing efficiency. Noise also increases the likelihood of drawing on incorrect information and thereby making decisions that do not optimize system efficiency.

During interviews, system leaders and actors expressed great concern with the quantity and nature of data being collected for improving cardiac care response. It is compulsory and common for many subsystem actors to collect and enter between 20–30 data elements to complete patient care reports. Many of these elements are unrelated to the efficiency metric (i.e. time) and are unrelated to subsystem level CQI. Through interviews with local and national cardiac arrest leadership the evaluation learned the problem can be traced to the competing national research agenda and local CQI agenda (Renger, Qin, Rice, Folytsova & Renger, 2016a). The impact of these competing agendas is devastating and documented in detail elsewhere (Renger, 2016b).

Feedback timeliness

Feedback about patient information must be passed accurately and quickly within and between subsystems. Delays in the passing of this information significantly impacts patient survival rates.

Using US Department of Homeland Security mock operational exercises, the evaluation discovered several bottlenecks leading to cascading failures. Problems arose with cell phone and radio coverage, internet connectivity, and hospital software platform intersystem interoperability. These issues delayed the timely delivery of patient data from subsystem to subsystem.

In response to these challenges one state is testing almost real time data flow[4] using cloud-based technology. Information from the dispatch center is used to search for patient information in the hospital record system. The medical record is then sent to the responding ambulance unit. Information gathered by EMS personnel at the scene is then sent to the receiving hospital. Finally, any documentation gathered at the hospital is shared back with the responding ambulance crew, an example of closing the feedback loop between subsystems.

Credible feedback

Credible feedback is information trusted by system leaders and actors for decision-making. Direct evaluator observation revealed poor attendance at state CQI initiatives. Following interviews with system actors the evaluation found many chose to not to participate because they did not deem the feedback credible. When asked why the feedback credibility was being questioned, some system actors admitted to fabricating data in response to state pressure to provide complete patient care records. Thus, system actors knew the data wasn't credible because they themselves were responsible for the system misinformation.

Another factor affecting feedback credibility were the data elements themselves. System actor interviews revealed many believed the data was driven by a federal research agenda and did not see the relevance of the collected data for their local CQI purposes (Renger et al., 2016b). This affected motivation to participate in data entry and helps explain why less than one third of cases entered contained complete information (Renger, et al., 2016b).

To address the credibility issue, one state was engaging in a pilot program where system actors provide direct input regarding meaningful CQI data elements. In addition, several other states were updating their data entry software to introduce higher levels of system actor configurability; allowing the tailoring of some data elements while maintaining a core statewide data set. The goal of adding configurability is to make the data more meaningful at a local level, thereby hoping to increase the motivation to enter credible information.

Discussion

This article focused on the SET principle of evaluating modern day system efficiency. Specially, it presented methods for evaluating one key component to optimizing system efficiency, feedback mechanisms.

At the system evaluation conference in Eschborn, Germany many evaluators wondered whether system thinking required any new methods or whether the available tools simply needed to be applied differently. This application of SET found many of the program evaluation tools useful in evaluating system feedback mechanisms, they just needed to be applied differently. For example, qualitative methods such as process flow mapping were essential in developing the standard operating procedures, while root cause analysis interviews were important in understanding barriers to system inefficiencies. Both methods are powerful program evaluation tools. Similarly, quantitative analyses often used in program evaluation were used to identify data credibility problems in feedback loops.

In addition, program evaluation theories and standards proved robust and useful in evaluating system feedback mechanisms. For example, utility theory lies at the heart of assessing any system feedback mechanism; if information isn't deemed useful due to a lack of credibility, timeliness, infrequency, relevance or specificity it will not be used for making system efficiency decisions.

Although the focus of this inquiry was on sharing lessons learned in evaluating feedback mechanisms it is clear the other aspects of SET system efficiency, system attributes and processes and procedures, are intertwined. With respect to system attributes, the evaluation found leadership, information technology, culture, and training significantly impact system feedback efficiency. For example, the failure to close feedback loops was directly attributed to leaderships' failure to understand the importance of providing feedback to system actors. When leadership chose to prioritize the timely and frequent provision of feedback it led to significant improvements in system actor data entry compliance. Information technology challenges significantly impacted feedback timeliness. Resolving software interoperability led to significant improvements in improving the timely and accurate transfer of patient information from one subsystem to another. Finally, providing system actor training in data entry and information technology was essential to improving feedback credibility. In summary, the evaluation supported SET's recognition of system attributes impacting system feedback efficiency.

Documented processes and procedures, that is standard operating procedures, directly impact feedback specificity. Detailed steps provide the foundation necessary to understand how to use system feedback to make corrective actions. The importance of detailing system processes parallels the importance and role of implementation protocols in program evaluation (Chen, 2005; Saunders, Evans & Joshi, 2005). An understanding of what is supposed to happen is key to being able to evaluate whether it happened and provide recommendations for making program improvements.

There were also two notable lessons learned as it relates to improving how to evaluate system feedback mechanisms. First, in addition to quality, the quantity of feedback should be evaluated. Too much feedback will result in noise; too little feedback doesn't provide the information needed for making changes to optimize system efficiency. The evaluation must assess whether there is a proper fit between the information being provided by the feedback mechanisms and the ability for it to be accurately processed.

Second, is the challenge of how to proceed when multiple system feedback mechanisms are simultaneously inefficient. In this case, the systems theory principle of cascading failure may prove useful in guiding the evaluator's recommendation. It is reasonable to posit earlier feedback mechanism inefficiencies may manifest themselves in later system feedback mechanisms as well, passing the problems down the line. If one accepts the system premise of cascading failures, then resources should be directed at first solving upstream system inefficiencies.

In conclusion, there is still much confusion in the literature as to what constitutes systems evaluation. SET guides the evaluation of modern day systems, not programs. The value of SET lies in providing evaluators a different, practical lens to guide evaluating a system and producing useful recommendations. The effect of this lens is pronounced and well-illustrated when evaluating system feedback mechanisms. From a program evaluation perspective, the role of providing feedback is often assumed by the evaluator. The evaluator gathers process and outcome data and via a report closes the program feedback loop. Hopefully, program staff then use the data to make changes, which are again monitored by the evaluator. An evaluator with an eye to sustainability may attempt to empower or build program staff capacity to assume this role. This evaluation process is very different under SET where the purpose is to evaluate the presence and functionality of system feedback mechanisms.

SET is a promising evaluation theory to systematically address system complexity. The focus of this article was on the application of SET to evaluate one specific component of system efficiency. It will be interesting to learn its utility in evaluating other aspects of system efficiency and system effectiveness.

Acknowledgments

The author would like to thank Dr. Carlos Rodriguez, Dr. Jirina Foltysova, Jessica 'Munch' Renger and Jenna 'Boo' Renger for their insights and support.

Endnotes

1. The evaluation was funded by The Leona M. and Harry B. Helmsley Charitable Trust.
2. Defined as getting the patient to a tertiary care facility with a catheterization lab.
3. System actors encompasses all system and subsystem leaders and staff.
4. The speed of data transfer depends on internet connectivity, which is sporadic in rural areas.

References

Ash, J. S., Berg, M., & Coiera, E. (2004). Some unintended consequences of information technology in health care: The nature of patient care information system-related errors. *Journal of the American Medical Informatics Association*, *11*(2), pp. 104–112. doi:10.1197/jamia.M1471

Banathy, B. H. (1992). *A systems view of education: Concepts and principles for effective practice*. Englewood Cliffs, NJ: Educational Technology.

Braithwaite, S., Friesen, J. E., Hadley, S., Kohls, D., Hinchey, P. R., Prather, M., Karonika, M., Myers, B., Holland II, W. D., Eason, C. M., & Carhart, J. (2014) A tale of three successful EMS systems: How 'pit crew' procedures improve cardiac arrest resuscitations in the field. *Journal of Emergency Medical Services Online*. Retrieved 22 October 2016 from http://www.jems.com/articles/supplements/special-topics/ems-state-science-2014/tale-three-successful-ems-systems.html

Brandon, R. N. (1990). *Adaptation and environment*. Princeton, NJ: Princeton University Press.

Burns, T. R. (2007). System theories. In Ritzer, G. *The Blackwell encyclopedia of sociology*. Malden, MA: Blackwell Publishing.

Carlbom, D., Doll, A., Eisenberg, M., Emert, J., Husain, S., Kudenchuk, P., Rea, T., Sayre, M., Sherman, L., & Stubbs, B. (2014). *Strategies to improve survival from sudden cardiac arrest: An evidence-based analysis*. Retrieved 22 October 2016 from http://www.resuscitationacademy.org/downloads/RA-35-Strategies-to-Improve-CA-Survival.pdf

Chen, H. T. (2005). *Practical program evaluation: Assessing and improving planning, implementation, and effectiveness*. Thousand Oaks, CA: Sage Publications.

Chen, H., Hailey, D., Wang, N., & Yu, P. (2014). A review of data quality assessment methods for public health information systems. *International Journal of Environmental Research and Public Health*, *11*(5), pp. 5170–5207. doi:10.3390/ijerph110505170

Clow, D. (2012). The learning analytics cycle: closing the loop effectively. In *LAK '12 Proceedings of the 2nd international conference on learning analytics and knowledge*. New York: ACM, pp. 134–138. doi:10.1145/2330601.2330636

Eisenberg, M. S. (2013). *Resuscitate! How your community can improve survival from sudden cardiac arrest*. 2nd ed. Seattle: University of Washington Press.

Flynn, B. B., Schroeder, R. G., & Sakakibara, S. (1994). A framework for quality management research and an associated measurement instrument. *Journal of Operations Management*, *11*(4), pp. 339–366. doi:10.1016/S0272-6963(97)90004-8

Gamel-McCormick, C. A. (2011). *A critical look at the creation and utilization process of logic models*. Masters Thesis. University of Delaware. Retrieved 22 October 2016 from http://dspace.udel.edu/bitstream/handle/19716/10095/Caitlin_Gamel-McCormick_thesis.pdf?sequence=1

Graham, B. B. (2004). *Detail process charting: Speaking the language of process*. Hoboken, NJ: John Wiley & Sons.

Grossman, D. C., Kim, A., Macdonald, S. C., Klein, P., Copass, M. K., & Maier, R. V. (1997). Urban-rural differences in prehospital care of major trauma. *The Journal of Trauma: Injury, Infection, and Critical Care*, *42*(4), pp. 723–729.

McDavid, J. C., Huse, I., & Hawthorn, L. R. L. (2013) Key concepts and issues in program evaluation and performance measurement. In *Program evaluation and performance measurement: An introduction to practice*, (2nd ed.). Thousand Oaks, CA: Sage.

Nickols, P. (2000). Standard operating procedures. *The Quality Assurance Journal*, *4*(2), pp. 91–94. doi:10.1002/1099-1786(200006)4:2<91::AID-QAJ78>3.0.CO;2-P

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): A practical framework for evaluators to meet the challenges of system evaluation. *Evaluation Journal of Australasia*, *15*(4), pp. 16–28.

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*(1). Retrieved 22 October 2016 from http://hdl.handle.net/10515/sy56t0hd0

Renger, R., Qin, X., Rice, D., Foltysova, J., & Renger, J. (2016a). National data collection efforts pose challenges for many EMS agencies. *EMSInsider*.

Renger, R., Qin, X., Rice, D., Foltysova, J., Souvannasacd, E., Renger, J., Markwart, B., Hart. G., & Bjerke, M. (2016b). *Challenges and solutions facing EMS in supporting the IOM recommendation for a national cardiac arrest registry: A system perspective*. White Paper. Retrieved 22 October 2016 from https://ruralhealth.und.edu/pdf/ems-supporting-iom-recommendation.pdf

Saunders, R. P., Evans, M. H., & Joshi, P. (2005). Developing a process-evaluation plan for assessing health promotion program implementation: A how-to guide. *Health Promotion Practice*, 6(2), pp. 134–147. doi:10.1177/1524839904273387

Slifkin, A. B., & Newell, K. M. (1998). Is variability in human performance a reflection of system noise? *Current Directions in Psychological Science*, 7(6), pp. 170–177. doi:10.1111/1467-8721.ep10836906

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

World Health Organization. (2002). *Handbook for good clinical research practice (GCP): Guidance for implementation*. Geneva: World Health Organization. Retrieved 22 October 2016 from http://apps.who.int/prequal/info_general/documents/gcp/gcp1.pdf