

Four suppositions about emergent properties of complex interventions that operate and function as systems: A call for evaluation research

Evaluation Journal of Australasia

2024, Vol. 24(1) 56–69

© The Author(s) 2023

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1035719X231217335

journals.sagepub.com/home/evj



Ralph Renger 

Just Evaluation Services, LLC, USA

Jessica Renger 

Kaiser Permanente, USA

Melanie T Pescud

University of Western Australia, Australia

Richard N Van Eck

University of North Dakota, USA

Marc D Basson

Northeast Ohio Medical University College of Medicine, USA

Jirina Renger

Walden University, USA

Abstract

This article focuses on evaluating the emergent properties of complex interventions that operate and function as systems. After defining key terms, four suppositions of emergence are presented. The suppositions are based on the authors' reflections on

Corresponding author:

Ralph Renger, Just Evaluation Services, LLC, 14777 E. Circle M Ranch Place, Vail, AZ 85641, USA.

Email: ralph@justevaluation.com

evaluating several complex interventions using a systems approach. Supposition 1 posits that there is an *operational* emergent property, paralleling the theory-driven evaluation concept of implementation theory noted by Carol Weiss. Supposition 2 suggests that every complex intervention also has a *functional* emergent property that speaks to the effectiveness of the intervention. Several important distinctions between a functional emergent property and the long-term outcomes derived based on a reductionist worldview are noted. Supposition 3 builds on the work of Chalmers to posit that it is possible for a complex intervention to simultaneously have more than one operational emergent property and more than one functional emergent property. Supposition 4 is grounded in the system definition to propose that operational emergence is a prerequisite for functional emergence. Suppositions by definition require proof to become theories. The article calls for evaluators to share whether their experiences support these suppositions and encourages investments in evaluation research to test them.

Keywords

complexity, evaluation, system, interdependence, emergence

What we already know

- Attention is being drawn to the need to evaluate interdependencies and emergent properties of complex interventions.
- An emergent property is qualitatively different than a long-term outcome and cannot be achieved independently by any one intervention component.
- One type of emergent property is related to the function/purpose/impact of the complex intervention (functional emergence).

The original contribution this makes to theory and/or practice

- There may be more than one type of emergent property.
- One type of emergent property may be related to the operation/implementation/process of a complex intervention (i.e., operational emergence).
- A complex intervention may be characterized by both operational and functional emergence.
- Operational emergence may be a prerequisite for functional emergence.

This article focuses on the challenges in defining and evaluating emergent properties of complex interventions that meet the system test (Renger, 2022). A complex intervention “involves at minimum multiple components and a complex pathway ...” (Guise et al., 2017, p. 7). In contrast, simple interventions are defined by a singular component and a linear theory of change (Petticrew, 2011). The authors have evaluated both types of interventions. An example of a simple intervention is using an electronic health system to remind a clinician to discuss breast cancer screening with a patient. An example of a complex intervention is the integration of community outreach, clinician reminders, patient reminders, diagnostic service providers, and treatment support services to improve breast cancer survival rates.

Researchers have suggested that evaluation approaches grounded in systems theory are a good fit for evaluating complex interventions because systems are by their nature complex (Holland, 1992; Renger, 2022). However, a system evaluation approach is fit for such purposes only if the complex intervention is designed to operate and function as a *system* (Renger, 2022). Therefore, an important first step for evaluators confronted with evaluating complex interventions is to engage in a system test to determine whether the intervention is indeed operating and functioning as a system (Renger, 2022) or is “just a bunch of stuff” (Meadows, 2008).

The system test asks evaluators to assess whether the complex intervention they plan to evaluate conforms to the system definition. Ackoff (1994) defines a system as an integrated whole whose essential property *emerges* from the *interdependencies* among system components. Many system definitions use the more general word “relationship” to describe the interaction among intervention components rather than the more specific term “interdependence” (e.g., AEA, 2018; Ison, 2008). In our experience evaluators using the term “relationship” either do so because it is a more benign introduction to the topic (Renger, 2022), or because they do not yet appreciate the nuances of how the term relationship differs from the system property of interdependence. Ackoff (1994) adds that the essential property that emerges is a *product* of the interdependencies of system parts, challenging the idiom that a system is the sum of its parts. Hereinafter we refer to the essential system that property that emerges as a product of the interactions as the emergent property.

Interdependence conveys a qualitatively deeper meaning about the nature of relationships among the components of an intervention. It highlights the holistic nature of a system; the operation of all parts is needed to achieve the higher order function (i.e., emergent property), and the system suffers or collapses if one-part malfunctions or is missing.

If the complex intervention meets the system test, (i.e., meets the system definition), then the intervention can then be evaluated using a systems approach. To be truly fit for purpose, however, the evaluation *approach* should also be *aligned* with the system definition. At a minimum, this requires that the evaluation approach must account for interdependencies and emergence. The evaluation of interdependence necessarily

precedes the evaluation of emergence because interdependence is a prerequisite for emergence (Ackoff, 1994). For example, system evaluation theory is an evaluation framework intentionally aligned to the system properties of interdependence and emergence (Renger, 2015). System evaluation theory answers the call of Walton et al. (2021) to move "...beyond a conceptual application of STCS [systems thinking and complexity science] to more deeply applying theory, methodologies, and approaches" (p. 163). System evaluation theory requires three steps: (1) defining the complex intervention operating and functioning as a system, (2) evaluating interdependencies, and (3) evaluating emergence.

Regardless of the evaluation approach used, evaluating emergence requires a complete understanding and appreciation of the emergent properties. The literature dedicated to the topic of evaluating emergence is scant and seems chiefly focused on emergence as an indicator of an intervention's effectiveness (Chalmers, 2006; Renger, 2022; Renger et al., 2023). However, based on our own work and experience evaluating several complex interventions using a systems approach, including the Housing and Urban Development funded Housing Opportunities for People Everywhere; the Centers for Disease Control funded Well Women Health Check Program; and the National Institutes of Health funded Center for Translational Research, we have identified four suppositions of emergent properties that we feel can help guide evaluators. We want to be clear that our suppositions represent our belief as to the nature and role of emergent system properties in relation to evaluating complex interventions. We acknowledge that they are based on our finite experiences evaluating complex interventions that operate and function primarily in the health sector and that we have "no proof" that they are robust, stable, and generalizable. However, Walton et al. (2021) noted that advancing systems evaluation thinking "requires a team" (p. 168). We concur and hope that by sharing these suppositions we will motivate others to test them in their own evaluations and to share their own findings so that together we can build a repository of knowledge to better understand and evaluate emergence in complex interventions.

We focus heavily on our Center for Translational Research evaluation to illustrate our suppositions, for two reasons. First, the basic structure is replicated in many other interventions. Thus, it is hoped evaluators reading this work will more easily find connections between the Center for Translational Research structure and the structure of the complex interventions they are evaluating. Second, our focus on emergence is a natural extension of our previously published work on the Center for Translational Research. Specifically, we applied system evaluation theory in defining the Center for Translational Research (Renger et al., 2020) and evaluating its interdependencies (Souvannasac et al., 2022). Using the Center for Translational Research to present and illustrate our suppositions related to emergence creates the contextual continuity needed to better understand the foundational work upon which discussions of emergence depend.

The Dakota Community Collaborative on Translational Activity Center for Translational Research is a complex, grant-funded organization that spans multiple

universities, hospital systems, and clinics across North and South Dakota and which is dedicated to changing the regional culture to foster more clinical and translational research focused on key health issues important to Dakotan communities. Funded by the National Institutes of Health, the Dakota Community Collaborative on Translational Activity includes seven cores and a practice-based research network which interact in a complex fashion to recruit investigators, support their efforts both financially and intellectually, and engage with the communities which the Dakota Community Collaborative on Translational Activity serves (Renger et al., 2020).

As described in our previous publications, the evaluation of Center for Translational Research interdependencies was straightforward and mechanistic in that we systematically applied system principles like feedback loops, cascading events, and reflex arcs to evaluate Center for Translational Research core standard operating procedures. However, evaluating the Center for Translational Research emergent property proved to be more challenging. We believe that at least one reason for this challenge can be traced back to Ackoff's qualifier that an emergent property is the *product* of the component interdependencies. The term "product" suggests that the whole is not *greater* than the sum of its parts as some authors suggest (Bryne & Callaghan, 2014), but is rather defined by a more complex algebraic expression that results in a whole that is qualitatively as well as quantitatively *different* from the sum of its parts (Law et al., 1998).

Supposition 1: Every complex intervention that meets the system test has an operational emergent property.

The operation of a complex system refers to the execution of processes, workflows, or standard operating procedures (Renger, 2022). Therefore, supposition 1 focuses on the *implementation* of complex interventions. In theory-driven evaluation, the implementation evaluation focuses on monitoring whether the steps of an intervention adhere to its prescribed protocol (Scriven, 1991; Weiss, 1997). However, with some exceptions (e.g., Smith et al., 2020), theory-driven evaluation methods for evaluating implementation are largely inadequate for capturing the interdependencies of complex interventions that operate as systems because they were designed to evaluate *simple* interventions that operate *independently* (Renger, 2022). Thus, for simple interventions, the evaluation of interdependencies using a systems approach is often moot, as there are no multiple components that are dependent on each other.

Weiss (1997) noted that the depth of implementation evaluations must go beyond merely checking whether key elements of the intervention protocol were executed with fidelity; they must include an evaluation of the *implementation change theory*. "Implementation theory refers to something deeper than the simple steps of an activity; it captures the essence of how the activities should be conducted to affect the mechanisms of change identified in the program theory" (Renger et al., 2013, p. 28). For example, Renger et al. (2013) noted that a "fun experience" (i.e., in German "ein Erlebnis") was a necessary component in the successful implementation of a life skills intervention in a

German trade school. The implementation theory (i.e., the educational activity must be fun) was viewed as a prerequisite for engaging students in learning life skills (i.e., the intervention outcomes).

Consistent with [Chen and Rossi's \(1983\)](#) observation four decades ago, it remains our experience that evaluations of the implementation theory underlying an intervention are rare, in no small part because in most cases, there *is* in fact no implementation theory underpinning the intervention. While intervention design flaws are not the fault of the evaluators, we do have a responsibility to inquire about whether there is an implementation theory that underpins an intervention when we develop an evaluation plan. One reason evaluators fall short in meeting Chen and Rossi's call is that evaluators may be unaware of the need for interventions to be grounded in implementation theory.

We contend that an evaluation of a complex intervention's implementation should, as is the case for simple interventions, include an evaluation of its implementation theory. Since the interdependencies detail how a complex intervention operates, an understanding of what emerges from the interdependencies is central to defining its implementation theory. We refer to the property emerging from the product of a complex intervention's interdependencies as the operational emergent property.

A good example of an operational emergent property can be found in team sports. When constructing a team, managers attempt to get the correct mix and balance ([Renger, 2022](#)) of players, coaches, and support staff. The goal is to win. However, the likelihood of winning is greater if the interdependencies between the team components (i.e., players, coaches, and managers) can be optimized. The evidence of success of optimizing these interdependencies emerges in the form of "team chemistry." Sport psychologists and sociologists have devoted their careers to evaluating team chemistry, or as it is known in Australia the "vibe of the playing group," although quantifying it remains elusive ([Gerschgoren et al., 2016](#)).

In complex social interventions, the "chemistry" emerging from the interdependencies between components has been referred to as the "secret sauce" (M. Fry, personal communication, 10 February 2023) or "where the magic happens" (L. Atkinson, personal communication, 1 March 2023). In our experience, evaluations often operationalize the "secret sauce" of complex social interventions as the level of collaboration between intervention components. This default position is understandable because the success of many complex social interventions depends on human collaboration between and among system actors representing different component parts. The dependence on human interaction for intervention success is evidenced by the finding that the implementation failure of many interventions can be traced back to poor communication ([McCulloch et al., 2011](#)).

In the spirit of full transparency, in our Center for Translational Research evaluation, we also defaulted to using collaboration as an indicator of implementation success. Specifically, we used the levels of collaboration scale to evaluate the success of the Center for Translational Research core interdependencies ([Frey et al., 2006](#)). The levels of collaboration scale asks system actors representing different intervention

components to assess the level of collaboration with other intervention components along a continuum spanning networking, cooperation, coordination, coalition, and collaboration. Our attraction to the levels of collaboration scale stemmed from the fact that the anchors used to evaluate collaboration are couched in systems language. For instance, the highest level on the levels of collaboration scale (collaboration) is anchored by the statements “members belong to one system” and “there is frequent communication characterized by mutual trust.”

When reflecting on our use of the levels of collaboration scale to evaluate Center for Translational Research success, we identified two primary concerns. First, tools like the levels of collaboration scale do not necessarily get to the heart of what it means to work interdependently. For example, in the levels of collaboration scale, the understanding of what it means to “belong to one system” is left to the individual respondent to decide. This idiosyncratic interpretation does not necessarily include an appreciation for what it means to work together interdependently. The understanding that “I need to weigh how a change to my processes will influence other processes,” gets to the heart of what it means to “operate as a system” and is qualitatively richer than simply having an understanding or appreciation of what other intervention components do. Second, the idea of collaboration in the context of evaluating complex interventions must extend beyond human-to-human interactions and include technology-to-technology interoperability and human-to-technology interactions.

Chen and Rossi (1983) wrote: “we also make a special plea for more intensive attention to developing knowledge and theory concerning how human services organizations work, so that our general understanding of *implementation systems* [emphasis added] will be advanced” (p. 300). We believe that one key to developing this knowledge is to define and evaluate the operational emergent properties of complex interventions. Therefore, we call upon our colleagues to share how they define and evaluate the operational emergent properties of the complex interventions they evaluate. We understand this will be challenging because many of the operational emergent properties may be seemingly abstract. How do we measure team chemistry, fun, and collaboration? However, this should not deter us from working with experts in scale development in operationalizing the operational emergent property. Sports psychologists once thought concepts like team chemistry to be abstract, while social scientists grappled with measuring quality of life and collaboration. With the necessary time and resources, all produced relatively reliable and valid quantitative and qualitative measures of these constructs and for different settings (e.g., Burckhardt & Anderson, 2003; DeLong, 2013; Son et al., 2022). Perhaps through sharing, we can identify common operational emergent properties for which we can then develop the tools necessary to capture the “chemistry” or “secret sauce” emerging more accurately.

Supposition 2: Every complex intervention that meets the system test has a functional emergent property.

Function is defined as the purpose or superordinate goal of the complex intervention. Therefore, supposition 2 focuses on evaluating a complex intervention’s impact

(Scriven, 1991). We refer to the impact that emerges from the interdependence of component parts as the functional emergent property.

The functional emergent property is qualitatively different than the theory-driven evaluation outcomes used to evaluate an intervention's impact. Theory-driven evaluation outcomes possess an inherent chronological property that is a direct result of the reductionist cause-and-effect thinking that underpins the theory of change. This chronology is evidenced in the labeling of outcomes as immediate, intermediate, and long term in the logic model (McLaughlin & Jordan, 2015). Emergence, on the other hand, by its very definition, "appears" when all the intervention components are working together, interdependently. By the same token, an emergent property can "disappear" if one of the dependent component parts fails. By this logic, we suppose that an emergent property is not time-bound, but rather occurs at the "tipping point" when the intervention becomes effective.

In theory-driven evaluation, the linear construction of the logic model is well suited for evaluating the impact of simple interventions. However, many theory-driven evaluation evaluators will develop *independent* logic models for each intervention component when confronted with the challenge of evaluating the impact of a complex intervention (Renger, 2022; Rogers, 2008). Theory-driven evaluation evaluators then attempt to demonstrate what we coin here as "complexity coverage" by scaffolding or coupling logic models (Walton et al., 2021). We argue that this is an ineffective approach when evaluating complex interventions that meet the system test: evaluating the effectiveness of each intervention component *independently* (logic models) is inconsistent with evaluating the system property of *interdependence*.

In reflecting on our practice, we found quality of life to be a recurring example of a functional emergent property common to many complex social interventions. One such complex intervention that uses quality of life as a success indicator is the Housing and Urban Development Housing Opportunities for People Everywhere VI "program." Although labeled a program, the Housing Opportunities for People Everywhere VI intervention consists of multiple components (e.g., life skills, education, housing, and childcare) intended to work together to address the complexity of underlying conditions facing those in public housing. While each component has a logic model and outcomes for which it is held responsible, for example, building new homes, relocating residents in mixed-income neighborhoods, and offering childcare, the "hope" is that a better quality of life emerges from the interdependencies between these components (Popkin et al., 2004). Evaluating the program with independent logic models would tell you little, if anything, about whether the system is producing better quality of life. However, with an understanding of the supposition of a functional emergent property, evaluators' attention can be directed to defining and evaluating what the interventions components are jointly trying to achieve, that is, quality of life, that they cannot do independently.

Quality of life is also used as a success measure for complex interventions targeting a wide range of health and social issues, including cancer treatment (Sanchez et al., 2011), urban studies (Rezvani et al., 2013), and aging (Leung et al., 2004). However, there are surely also many different functional emergent properties in other sectors like

transportation safety, economic security, climate change, etc. In our evaluation of the Center for Translational Research, we defined the functional emergent property as Researcher Self-efficacy (Phillips & Russell, 1994). We reasoned that no support core by itself could create the confidence and skills necessary for clinicians to be successful in pursuit of an independent research program. Further, the failure of any single core to provide its support could keep researcher self-efficacy from emerging.

We call upon our colleagues to share how they define and evaluate functional emergent properties. We recognize that other sectors are not necessarily oriented toward a system perspective, so uncovering potential functional emergent properties may require some detective work. One suggestion is to search for interventions in which multidimensional measures of success are being used. If complex interventions consist of multiple components, then it is reasonable to posit that there is a functional emergent property and that it may also be multidimensional.

Supposition 3: Complex interventions can consist of more than one operational emergent property and more than one functional emergent property.

As stated above, suppositions 1 and 2 imply that at a minimum, each complex intervention must have one operational and one functional emergent property. However, we also believe that complex interventions meeting the system test can produce more than one operational emergent property and more than one functional emergent property.

To support this supposition, we cite the work of Chalmers (2006) who forwards the notion of weak and strong emergence. If a complex intervention can have a weak and strong emergent property, then by definition multiple functional emergent properties can exist. Our understanding of the distinction between these two emergence types is that weak emergence is that which might be expected from interactions between systems parts, while strong emergence is unexpected. To illustrate the difference between weak and strong emergence, Renger (2022) draws on a hypothetical example in designing a public transportation system. One can imagine that the functional emergent property guiding the transportation system design might be to develop users' trust by ensuring that the system reliably moves a user from point A to point B. The reliability might be evaluated by tracking system delays, while the trust might be evaluated by rider loyalty. However, how might one explain high reliability in a situation in which rider loyalty remains low? Perhaps loyalty might depend upon something that had not been expected, such as users' feelings of safety that might be threatened at some transportation hubs. In such a case, the system would need to be redesigned to also consider the location and conditions of transportation stops and connections.

If there is weak and strong emergence, then by definition it is possible for a complex intervention acting as a system to have more than one functional emergent property. However, it is also our supposition that a complex intervention can also have more than one operational emergent property. For example, one might predict that a culture of collaboration is an operational emergent property that is a prerequisite for a more nuanced emergent property like team "chemistry" to emerge. Similarly, our reflection also leaves open the possibility that an emergent property can be both operational and

functional. For example, one might argue, depending on how the intervention is framed, that an emerging culture of collaboration of a complex intervention might be considered both evidence of operational (i.e., implementation) and functional (i.e., impact) success. Therefore, we call upon our colleagues to consider whether there are additional and/or qualitatively different emergent properties, beyond those which was might be reasonably predicted.

Supposition 4: Both operational and functional emergent properties are prerequisites for achieving intermediate and long-term outcomes.

This supposition hinges on the system definition that interdependence is a prerequisite for emergence. We believe that an operational emergent property may be a prerequisite for a functional emergent property. For example, in our Center for Translational Research evaluation core “chemistry,” defined as collaborating as a system, is the operational emergent property and is a prerequisite for researcher self-efficacy, the functional emergent property.

However, our reflection also led us to the supposition that the functional emergent property may not only be an important “outcome” in its own right, but may also in some circumstances be a prerequisite for meeting immediate, intermediate, and long-term outcomes typically found in theory-driven evaluation. As discussed under supposition 2, we view immediate, intermediate, and long-term outcomes as qualitatively different than emergent properties. However, this does not mean that theory-driven evaluation outcomes and the functional emergent property are mutually exclusive; both might exist in harmony. Indeed, the functional emergent property may sometimes even be a prerequisite for fully achieving theory-driven evaluation outcomes.

For example, in our Center for Translational Research evaluation, we predicted that researcher self-efficacy would emerge if the support cores worked together interdependently. Researcher self-efficacy is the functional emergent property. We also reasoned that researcher self-efficacy is a prerequisite for clinicians to be successful in meeting outcomes of interest required by the funder (National Institutes of Health) such as developing an independent research program, publishing, and achieving independent extramural grant funding. Since funder evaluation requirements are non-negotiable and the outcomes they require are derived from traditional theory-driven evaluation, such as those found in logic models, it would be beneficial to demonstrate how the addition of a system-based evaluation, specifically the measurement of the functional emergent property (researcher self-efficacy), is necessary to ensure the conditions needed to achieve these outcomes are met.

Discussion

We shared four suppositions about emergent properties of complex interventions, based on our own evaluation work with complex intervention systems. We acknowledge that these are suppositions, not established theory, yet they seem reasonable based on reflection of our evaluation practice. Our goal in sharing these suppositions was to

stimulate thought and to encourage other evaluators to share the extent to which their evaluations of complex interventions operating and functioning as systems support (or refute) these suppositions.

We believe that it is reasonable to posit that if multi-component interventions working interdependently are necessary for emergence, then what emerges must also be a multi-dimensional construct (Renger, 2022). We believe there is a tremendous opportunity to adapt knowledge from other professions in trying to define the multi-dimensionality of emergent properties beyond the operational and functional definitions posited in the current work. For example, in exploring how to create a culture where teamwork is the “special sauce,” we might consider measuring human experiences and interactions on a physical, emotional, and mental level and how they can come into alignment. We might also borrow lessons from biology and neuroscience to gain a deeper understanding of emergent properties. For example, while one neuron does not produce consciousness or self-awareness, the interdependence of billions of neurons does. This is a stark representation of the whole being different than the sum of its parts. Cognitive science and the studies of distributed cognition may also shed light on defining emergent properties (Hutchins, 2000). We suggest that these may all be fruitful areas for further contemplation and research.

There might also be value in following the recommendations for developing multidimensional constructs (Johnson et al., 2012) and taxonomies (Law et al., 1998) developed in the management sector as we apply knowledge from other professions to define emergent properties. For example, Law et al.’s (1998) conceptualization of latent models (those that exist at deeper levels than their dimensions) seems a promising avenue to guide how measures can be developed to capture the essence of operational and functional emergent properties. In addition, the differences that Law et al. (1998) propose between latent and algebraic models seem to align directly with the notion of Chalmer’s weak and strong emergence and thus may represent a path forward for predicting whether a complex intervention is producing one or more emergent properties.

It should be emphasized that systems-driven evaluation approaches can exist in harmony with other program-based evaluation approaches. Evaluating the operational emergent property need not preclude engaging in traditionally mechanistic evaluation such as quality assurance. In fact, intervention component standard operating procedures/workflows must be evaluated for efficiencies before an evaluation of the operational emergent property. This is because, as Chen and Rossi (1983) noted, the “neglect of understanding implementation has made it ambiguous in many cases Whether the program or the implementation system or both were at fault in a demonstrated failure to achieve outcomes” (p. 299).

Finally, thought should be given as to who is responsible for evaluating emergent properties. Since no single intervention component can independently produce an emergent property, it stands to reason that no single intervention component should be held responsible for its evaluation. Using our previous examples, the responsibility for evaluating quality of life and researcher self-efficacy, rests with the external evaluator

overseeing the intervention. However, in other complex interventions, the responsibility for collecting data on emergent outcomes may fall on the agency overseeing the implementation of multiple interventions. We find Friedman's (2005) conceptualization of community-based indicators to be consistent with the understanding that emergent properties must be collected by higher system levels.

In summary, while there is progress in the methods used to evaluate the interdependencies within complex interventions that operate as systems, the evaluation of emergent properties is in its infancy. We believe that evaluating whether the "secret sauce" is emerging is key to knowing the success at an intervention's operational level and is a prerequisite for complex interventions to meet their functional purpose. However, to move our discipline forward we need the help of our evaluation community, a team, to form exemplars and research to test our suppositions.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by National Institutes of Health Award Number U54GM128729.

ORCID iDs

Ralph Renger  <https://orcid.org/0000-0002-0224-1444>

Jessica Renger  <https://orcid.org/0000-0002-5120-9504>

References

- Ackoff, R. (1994). If Rusbs Ackoff had given a TED talk. [Video]. YouTube. <https://www.youtube.com/watch?v=OqEeIG8aPPk>
- AEA. (2018). Principles for effective use of systems thinking in evaluation. *Systems in evaluation TIG*. AEA. <https://www.aes.asn.au/images/stories/regions/QLD/SETIG-Principles-FINAL-DRAFT-2018-9-9.pdf>
- Bryne, D., & Callaghan, G. (2014). *Complexity theory and social sciences: An introduction*. Routledge.
- Burekhardt, C. S., & Anderson, K. L. (2003). The quality of life scale (QOLS): Reliability, validity, and utilization. *Health and Quality of Life Outcomes*, 1(60), 60. <https://doi.org/10.1186/1477-7525-1-60>
- Chalmers, D. J. (2006). *Strong and weak emergence*. Australian National University. www.consc.net/papers/emergence.pdf
- Chen, H. T., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review*, 7(3), 283–302. <https://doi.org/10.1177/0193841X8300700301>

- DeLong, C. (2013). *Team chemistry: The missing link in skill assessment for teams*. [Doctoral Dissertation, University of Minnesota. <https://hdl.handle.net/11299/161624>
- Frey, B. B., Lohmeier, J. H., Lee, S. W., & Tollefson, N. (2006). Measuring collaboration among grant partners. *American Journal of Evaluation*, 27(3), 383–392. <https://doi.org/10.1177/1098214006290356>
- Friedman, M. (2005). *Trying hard is not good enough: How to produce measurable improvements for customers and communities*. Trafford Publishing.
- Gershgoren, L., Basevitch, I., Filho, E., Gershgoren, A., Brill, Y. S., Schinke, R. J., & Tenenbaum, G. (2016). Expertise in soccer teams: A thematic inquiry into the role of shared mental models within team chemistry. *Psychology of Sport and Exercise*, 24, 128–139. <https://doi.org/10.1016/j.psychsport.2015.12.002>.
- Guise, J., Butler, M. E., Chang, C., Viswanathan, M., Pigott, T., & Tugwell, P., Complex Interventions Workgroup. (2017). AHRQ series on complex intervention systematic reviews—paper 6: PRISMA-CI extension statement and checklist. *Journal of Clinical Epidemiology*, 90, 43–50. <https://doi.org/10.1016/j.jclinepi.2017.06.016>.
- Holland, J. H. (1992). Complex adaptive systems. *Dædalus*, 121(1), 17–30. <http://www.jstor.org/stable/20025416>
- Hutchins, E. (2000). Distributed cognition. *International Encyclopedia of the Social and Behavioral Sciences*. Elsevier Science, 138, 1–10.
- Ison, R. L. (2008). Systems thinking and practice for action research. In P. W. Reason, & H. Bradbury (Eds.), *The Sage handbook of action research* (2nd ed., pp. 139–158). Sage Publications.
- Johnson, R. E., Rosen, C. C., Chang, C. H., Djurdjevic, E., & Taing, M. U. (2012). Recommendations for improving the construct clarity of higher-order multidimensional constructs. *Human Resource Management Review*, 22(2), 62–72. <https://doi.org/10.1016/j.hrmr.2011.11.006>
- Law, K. S., Wong, C. S., & Mobley, W. H. (1998). Toward a taxonomy of multidimensional constructs. *Academy of Management Review*, 23(4), 741–755. <https://doi.org/10.2307/259060>
- McCulloch, P., Rathbone, J., & Catchpole, K. (2011). Interventions to improve teamwork and communications among healthcare staff. *British Journal of Surgery*, 98(4), 469–479. <https://doi.org/10.1002/bjs.7434>
- McLaughlin, J. A., & Jordan, G. B. (2015). Using logic models. In E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of practical evaluation* (4th ed., pp. 62–87). <https://doi.org/10.1002/9781119171386.ch3>
- Meadows, D. (2008). *Thinking in systems: A primer*. Chelsea Green Publishing.
- Petticrew, M. (2011). When are complex interventions “complex”? When are simple interventions “simple”. *The European Journal of Public Health*, 21(4), 397–398. <https://doi.org/10.1093/eurpub/ckr084>
- Phillips, J. C., & Russell, R. K. (1994). Research self-efficacy, the research training environment, and research productivity among graduate students in counseling psychology. *The Counseling Psychologist*, 22(4), 628–641. <https://doi.org/10.1177/0011000094224008>

- Popkin, S. J., Katz, B., Cunningham, M. K., Brown, K., Gustafson, J., & Turner, M. A. (2004). *A decade of HOPE VI: Research findings and policy challenges*. The Urban Institute. <https://www.urban.org/sites/default/files/alfresco/publication-pdfs/411002-A-Decade-of-HOPE-VI.PDF>
- 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), 16–28. <https://doi.org/10.1177/1035719X1501500403>
- Renger, R. (2022). *System evaluation theory (SET): A blueprint for practitioners evaluating complex interventions acting and functioning as systems*. Information Age Publishing.
- Renger, R., Bartel, G., & Foltysova, J. (2013). The reciprocal relationship between implementation theory and program theory in assisting program design and decision-making. *Canadian Journal of Program Evaluation*, 28(1), 27–41. <https://doi.org/10.3138/cjpe.028.002>
- Renger, R., Basson, M. D., Hart, G., Van Eck, R., Souvannasacd, E., Renger, J., & Foltysova, J. (2020). Lessons learned in evaluating the infrastructure of a centre for translational research (CTR). *Evaluation Journal of Australasia*, 20(1), 6–22. <https://doi.org/10.1177/1035719x20909910>
- Renger, R., Renger, J., Van Eck, R. N., Basson, M. D., & Renger, J. (2023). A comparative case study illustrating the influence of a system perspective on the outcome evaluation of simple versus complex interventions. *Evaluation Journal of Australasia*, 23(2), 101–110. <https://doi.org/10.1177/1035719X231160584>
- Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), 29–48. <https://doi.org/10.1177/1356389007084674>
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Sage Publications.
- Smith, J. D., Li, D. H., & Rafferty, M. R. (2020). The implementation research logic model: A method for planning, executing, reporting, and synthesizing implementation projects. *Implementation Science: Iscus*, 15(1), 84. <https://doi.org/10.1186/s13012-020-01041-8>
- Son, G., Yun, J., & Jeong, H. (2022). Quantifying team chemistry in scientific collaboration. *Physics and Society*, 07252, arXiv:2202. <https://doi.org/10.48550/arXiv.2202.07252>
- Souvannasacd, E., Engblom, H., Van Eck, R. N., Renger, R., & Basson, M. (2022). Lessons learned in evaluating system interdependencies using qualitative methods. *Evaluation Journal of Australasia*, 22(2), 108–125. <https://doi.org/10.1177/1035719X211066961>
- Walton, M., Gates, E. F., & Vidueira, P. (2021). Insights and future directions for systems and complexity-informed evaluation. *New Directions for Evaluation*, 2021(170), 159–171. <https://doi.org/10.1002/ev.20459>
- Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review*, 21, 501–524.