Reviewer's report

Title: Organizational interventions employing principles of complexity science have improved outcomes for patients with Type II diabetes

Version: Date: 26 February 2007

Reviewer: David C Aron

Reviewer's report:

General
This interesting paper by Leykum et al. uses a “complexity lens” in performing systematic review of organizational interventions designed to improve outcomes for patients with type 2 diabetes that as the conceptual framework. Specifically, they have examined the potential association between improved outcomes and interventions employing principles of complexity science.

1. Is the question posed by the authors for the review new and well defined?
   The use of a complexity lens is novel. However, the degree of definition depends upon the clarity and comprehensiveness of the definition. The degree of focus also depends upon the way that the definitions of characteristics of complex adaptive systems are operationalized. There are significant deficiencies in this regard (see below).

2. Are the methods appropriate and well described, and are sufficient details provided to replicate the work?

3. Are the data sound and well controlled?
   Question: Did the authors look for the appropriate sort of papers?
   The papers are limited to those in the English literature identified by standard bibliographic methods. It is not clear that all management literature would be identified by their approach. In addition, it is not clear that there was follow up from reference lists, personal contact with experts, or search for unpublished published studies. The "grey literature" is important to identify studies/projects outside the research/academic mainstream. Unpublished work may be a source for details about a study not present in the published version. This is particularly important in evaluating the characteristics of an intervention. It does not appear that the authors sought additional detail from those who conducted the actual interventions. There is also an irony in the reliance on the results of randomized trials in particular, given the shortcomings of the RCT in evaluating complex social interventions.

   Question: Was methodological quality assessed and the trials weighted accordingly?
   The study designs for the review were limited and the designs noted, but there was no weighting. Similarly, the specifics of the interventions are limited and there was no assessment of the level of detail provided. The outcomes are quite heterogeneous both in terms of their operationalization and intervention duration. The trichotomization of outcomes is an interesting approach, but, there is no weighting by number of participants or magnitude of the effect.

   Question: How sensitive are the results to the way the review has been done?
   The authors are unconvincing in persuading this reviewer that the results are not extremely sensitive to the way the review was done. First, the authors chose only some of the characteristics of complex adaptive systems. Among those omitted was “sensitivity to initial conditions.” Those initial conditions could impact upon any of the other characteristics. The degree of uniqueness of each clinical microsystem, makes randomization to eliminate the effect of differences in initial conditions difficult if not impossible. One wonders if certain types of interventions, e.g., those less dependent upon initial conditions, would be more likely to exhibit success. This could lead to biased conclusions. Second, the nature of the operational definitions and the typically low level of detail provided in a publication made it more likely that the presence of certain types of interventions would bias the results. The most obvious example is case managers. Interventions that involved components that would foster more effective interactions among existing personnel without the addition of new personnel might be less likely to be described. Third, the authors come from (essentially) a single institution and are likely to interpret the definitions similarly. I am not sure that inter-related reliability would be as high as reported if the literature were evaluated by someone from a different “school” of complexity.

   Question: Have the numerical results been interpreted with common sense and due regard to the broader aspects of the problem?
   The trichotomization of outcomes depends upon a very clear definition of what constitutes the difference between the three possibilities. I am not sure that I could reliably distinguish mixed from positive.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
   Not applicable
5. Are the discussion and conclusions well balanced and adequately supported by the data?
The limitations of the analysis are not sufficiently outlined.
6. Do the title and abstract accurately convey what has been found?
Yes
7. Is the writing acceptable?
Yes

- Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)
  1. An expanded limitations section.
  2. Inter-rater reliability involving raters from a different “school.”

- Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)
  1. I strongly recommend a more detailed, balanced, and expansive description of complexity theory(ies).

I am happy for this review to be seen by one and all.
David C. Aron, MD, MS

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.