**CALCULATING THE COST EFFECTIVENESS OF HUMAN SERVICES INTERVENTIONS: IMPLICATIONS FROM SCIENTIFC REALSIM, COMPELXTIY AND BEHAVIOUR CHANGE THEORY**

**Andrew Hawkins**

**ARTD Consultants, Sydney, Australia**

**Notes**

This paper was presented at the Australasian Evaluation Society International Conference, Sydney, Australia, 29 August – 2 September 2011. The views presented are those of the author rather than ARTD Consultants.

**Abstract**

Policy makers funding human services programs usually want to know if money was well spent. Common questions include ‘was this program cost effective’, or ‘did the benefits outweigh the costs?’ In standard economic analysis these questions require the quantification of program costs and outcomes. Economics and health research have a long history of influence on evaluation, most visibly in the long running and continually re-surfacing epistemological debates about randomised controlled trials. Yet economics and health research disciplines do not handle complexity well. Evaluation, relying on mixed methods and sometimes program logic or scientific realism, is better suited to a scientific understanding of human behaviour in complex systems, but often struggles with the quantification of outcomes. Calculating the cost effectiveness and relative worth of different human services programs from either perspective alone usually feels underdone.

This paper suggests an alternative approach that draws on implications from realist evaluation, complexity theory, psychology and methods of cost effectiveness in health economics. It attempts to provide a ‘stepping stone’ to an alternative way of thinking about the economic analysis of human services programs grounded in established scientific theories. This approach should ultimately provide more accurate and useful information for policy makers.

This paper has five parts. Section 1 describes problems in approaching cost effectiveness as most economists do i.e. with an implicit assumption that programs, like bridges or drugs, are stable, replicable interventions that once deployed exert relatively predictable forces for change. It suggests taking a realist approach, which posits that programs include collections of mechanisms (M) that may help generate regular outcomes (O) for certain people in certain contexts (C), is a more useful starting point. Section 2 summarises the problem of measuring real world outcomes of programs (or mechanisms) in complex systems: the problem of measuring the impact of a butterfly flapping its wings. Section 3 describes how incremental cost effectiveness ratios (ICERs) may be calculated for programs operating in simple and possibly complicated systems *even if* good outcomes data were available. Issues to be overcome include the need to collect data on costs for individuals and finding a common metric by which outcomes of interventions can be expressed and compared. Section 4 argues that the problem of finding a common metric is overcome by focusing on measuring impact on human decision making and positive behaviour change—the real aim of most social policy. Section 5 discusses the implications for answering questions about the relative worth and cost effectiveness of human services interventions attempting to change the behaviour of people living in complex systems in terms of common factors influencing behaviour.

**Introduction**

Policy makers funding human services programs usually want to know if the money was well spent. Common questions include ‘was this program cost effective’ and ‘did the benefits outweigh the costs?’ In standard economic analysis these questions require quantification of the costs and outcomes of the program. Consideration is then given to whether the cost relative to the outcomes generated was higher or lower than for other interventions targeting the same problem (as in cost effectiveness), or whether the dollar value of outcomes generated outweighed the costs (as in cost benefit analysis). This approach assumes that programs exist as stable, replicable entities for which outcomes can be measured in the real and often complex world.

This paper challenges that view, but accepts the validity of the question, were program funds well spent?

Programs are things that get funded. They include administrative staff, program staff, polices and procedures. Yet they are only useful in the real world as specific *interventions* that provide resources, improve decision making and change behaviour. I will differentiate throughout this paper between the unit of funding (programs) and the unit of change (interventions).

Human services programs operate in the real world and the world is a complex system. A person is affected by their genes, their environment, their family, their job, their friends, the media, the weather and so on. All these factors operate at the same time and impact on each other in unpredictable ways. People learn and adapt, economies and relationships change. The term ‘complexity’ as it is used in this paper means more than just managing a large number of these factors—that’s ‘complicated’. ‘Complexity’ occurs when these things are themselves constantly changing. It’s not only very hard to isolate the impact of any of these factors on a person, you cannot measure the initial conditions of a system and the very moment you attempt to make a measurement the system is changed.

This paper rests on three premises. First, programs exist primarily as units of funding, but do not exist in the real world as stable, replicable, entities that change behaviour. Second, outcomes of programs or actual interventions in the real world cannot be measured in any but the simplest of systems or over the shortest periods of time. And third, improved decision making and behaviour change, rather than eventual results of the interactions between an uncontrollable complex world and its citizens, is the real (if not advertised) concern of policy makers. Despite these obstacles, the guiding principle for this paper is that any cost effectiveness analysis requires a common metric by which one intervention can be compared with another.

Scientific realism and complexity together suggest a focus on the mechanisms of behaviour change rather than real world outcomes as the key to improved decision making. This paper suggests that a common metric should not be based on the real world outcomes attributed to a program (which will be different for different interventions anyway), but how cost effectiveness may be undertaken by looking at the costs required to generate changes in physical circumstances or cognitive factors that drive human behaviour.

**Part 1: The false assumption that programs exist as stable entities exerting a change force**

Questions asked of evaluations are shaped by the epistemology of those commissioning evaluations. Often grounded in health or economics research, these assume that programs—like drugs or bridges—have stable, reliable and usually linear impacts. While it is sometimes accepted that outcomes are difficult to predict, the implicit assumption is of an entity to which an outcome can be attributed.

Programs are not stable entities. People working in the field know that you have to adapt a program to work in different situations. They also know that some people do this job better than others. In the policy world the unit of funding is a program, but from the individual end user’s perspective ‘the program’ may just be talking with some people about their life.

The realist approach says that programs exist in the real world as *interventions* or *mechanisms* of behaviour change implemented by one group of people using resources made available by another to provide a third group with opportunities to make better decisions about their life. This may include changing a person’s (for example, by providing physical or financial resources) or changing the way a person interacts with the world (for example, by building their self esteem, self efficacy, resilience or optimism). What they don’t do is apply some irresistible force that changes people.

A ‘program’ provides no guarantee of the effectiveness of an actual intervention because of the many different ways different people will implement the program with different recipients. The important thing is the mechanisms of behaviour change (theory of change), rather than program activities (theory of action)—the latter are just a means to ‘triggering’ the former. This is the very reason for formative evaluation before a program is ‘mature’ enough for summative evaluation.

Understanding the necessary conditions (contextual and personal) for citizens, families and communities to change their behaviour in some desired way means discarding ‘programs’ as the unit of analysis. There is just too much noise in the systems in which human services programs are implemented to allow for meaningful measures.

**The health research legacy: the assumption that you can control and measure outcomes of interventions in complicated or complex systems**

Rather than seeing interventions as one thing that impact upon humans living in the world researchers tend to want to ‘control’ for other factors to measure the impact of a program. This is the idea behind the randomised control trial, which involves testing a group of people on some desired state of being, then randomly allocating them to receive an intervention (the treatment group) or no intervention (the control group), then testing them again. If differences in state of being are found, there are attributed to the program. The first methodological problem with this approach is the assumption that programs exist as stable entities. The second is the assumption that you can average out the differences between treatment and control groups to conclude that any differences are solely due to the intervention. Life does not progress along a normal curve. We cannot take people out of their context (as in a laboratory) or control for contextual factors statistically. We can try, and this is what most proponents of these methods suggest, but trying to use a tool that is poorly suited to a task is generally a poor use of resources.

**Program logic and realist evaluation improve understanding of causation, but do not lend themselves to quantitative measurement**

The program theory approach to evaluation allowed decision makers to step into the logic of the program and tests parts of it rather than trying to test the whole thing at once as a single entity. Evaluations based on program theory often involve triangulating evidence, or inductive reasoning, to make a case about how something is or is not working, rather than by formal hypotheses testing and experimental design. Sometimes this involves comparison of competing theories as in contribution analysis. Yet we cannot often *measure* the outcomes. This approach is useful for understanding programs and interventions, but does not lend itself to quantitative answers to questions about the value of one program compared with another as required for economic analysis.

Realist evaluation provides a great opportunity for thinking about economic analysis in human services. This is because at its core it avoids the common misconception that programs either work or don’t work. It seeks to identify context- mechanism­-outcome (CMO) patterns. Mechanisms are generally the things that make interventions work. Based 0n this theory, when faced with a problem decision makers should look at the context and identify mechanisms that have been shown to be most effective at addressing the problem. Realist approaches provide a powerful way of understanding and valuing programs but tend towards an ever growing collection of CMO configurations for consideration by decision makers. In cost effectiveness we need to give up some explanatory power to increase our power to compare different programs, interventions or CMO configurations. We need to focus on what is common and measurable. Before elaborating on this we must first address complexity.

**Part 2: Facing complexity**

Dealing with interventions or mechanisms instead of programs only helps with part of the problem. We must also understand the systems into which these intervene, which can be characterised as simple, complicated and complex. (Those interested in a more detailed articulation of systems can find this in texts by Michael Quinn Patton and Patricia Rogers). Unlike experimental approaches, which can create simple systems to test interventions, most human services programs operate in complicated or complex systems. Complex systems involve interdependence, non-linearity, unpredictability, emergent outcomes and, importantly, sensitivity to initial conditions.

Social policy, like macroeconomics, deals with the chaos and complexity. Chaos theorysays that predicting outcomes of simple processes can become wholly unpredictable because of the impossibility of measuring initial conditions at sufficient detail. While you can measure air displacement under a butterfly wing and could in theory calculate the impact based on simple physics, it is not possible to predict its impact on whether systems because it is not possible to measure the initial location of the wing and the air around it precisely enough. People and the range of factors influencing their behaviour, including outcomes and learning that in turn influence behaviour, are infinitely more complex than wind dispersion, yet we often expect to be able to observe outcomes of interventions.

A complex system may appear static but actually be undergoing enormous flux. In the right context an intervention can be a catalyst but, if the system is not in a state that an intervention or catalyst will spark a reaction, it may *appear* to be doing nothing. This does not necessarily mean the intervention is not worthwhile, it may be creating the necessary conditions for change in a way that cannot be measured linearly, or on the other hand it may just not be suited to the context in which it is deployed. Too great a focus on outcomes of an intervention or program in complex system with many non-linear and uncontrollable factors will often miss what is actually of value.

Recently young people across Britain engaged in running street battles with the Police. It would appear this was a result of pent up frustration with social and economic conditions: as one journalist put it, ‘Britain is a tinderbox and on Friday someone lit a match[[1]](#footnote-1)’. Tension can build and resentment bubble away until it hits a threshold and some catalyst ignites a real world reaction. As evaluators of social policy we must be sensitive to these dynamics. We need to find mechanisms that address the underlying causes of riots. Instead we often try to draw inferences between programs and the frequency and duration of these outcomes.

In their own field, economists have come up with different ways of treating simple and complex situations by dividing them into micro and macroeconomics, but they continue to assume that complex human services interventions can be understood with tools designed for simple systems

Questions about the outcomes and worth of programs are difficult to answer using quantitative methods derived from medical science—by this I mean the experimental method and randomised controlled trials that aim to replicate simple systems. It *may* be possible to meet the requirements for experimental design in simple or complicated systems with meticulous and painstaking research, logistical effort and ethical consideration to measure the independent impact of the intervention. Yet what we are generally confronted with in human services are interventions that may or may not trigger behaviour change in intended recipients for a range of reasons. People in the real world are impacted by innumerable factors that change over time; they learn and adapt; and feedback loops and non-linear changes make it hard to measure anything when dealing with complex systems. The scientific methods that work in simple systems should not be expected to perform in complex systems in which outcomes are emergent rather predictable. Experimental designs lack ecological validity when applied to complex systems.

We cannot judge the worth of a woman’s refuge program based on whether she escapes a violent relationship or gets a job. This would ignore too many factors both within and outside of the program for getting a job to be a reliable or valid measure of the worth of the program. The experimental approach—and the random controlled experiment in particular—try to get around this problem by random assignment. They see the situation as one that is complicated, but that can be controlled and made simple. Currently, this is not possible, it a may one day be possible with better methods for measuring complexity, in time-consuming research projects with large sample sizes where efforts can isolate and measure the relative importance of aspects of an intervention. It may be possible when measures are sensitive to sequencing requirements, tipping points and feedback loops. But currently it is not useful in evaluation when we are looking at the worth of an intervention as applied to people embedded in a social situation.

**Simple, complicated and complex systems in which human services interventions operate**

Measurement of quantifiable real world outcomes from human services programs is often all but impossible. While real world outcomes of interventions are of utmost importance for people, they are unknowable by decision makers and therefore should not form the basis for decision making—even though they may be necessary for media releases.

Measuring the impact of interventions in simple systems may be possible due to the limited factors involved and the relatively brief timeframe over which outcomes are observed. Measuring outcomes of interventions in complicated systems is often practically impossible because of the problems of attribution and impossibility of controlled conditions (no matter how *rigorous* the methods). Measuring outcomes in complex systems is not only difficult, it is theoretically impossible.

An intervention in a simple system may aim to increase understanding of the guidelines for responding to the abuse of children and young people through a seminar. In this instance a randomised control trial may be appropriate to try and measure the independent impact of the seminar on people’s knowledge, attitudes and beliefs before and after attending. Here we can follow the lead of health economics or economics more broadly without much bother for the major arguments in this paper about the nature of most human service interventions.

An intervention in a complicated system may aim to improve the process of coordinated action between agencies in responding to the abuse of children and young people. If an experimental model was applied it would be necessary to include pre- and post-testing and to either randomly assign people to departments and positions, or to try and control for the other factors affecting coordinated responses, such agency culture, and political interest, through statistical controls. It would be generally accepted that so many factors shape the way agencies that it would be futile to try measure the impact of a particular initiative on the extent to which an agency coordinates. In this situation, it would be more valuable to try approaches that have a hope of working because we know coordination is a mechanism for improvement in whole-of-government responses—the independent impact of coordination does not and cannot be tested in every evaluation.

A program that seeks to increase family functioning through integrated case management is an intervention in a complex system. Here the program is often necessarily expressed in abstract terms, such as case management, because there is no simple mechanical approach that works for assisting families. These programs are not comprised of stable entities exerting an irresistible force that changes people. In this situation, rather than trying to find outcomes from programs we should be honest about the limits to measurement in complex systems and focus on how best to use resources to enhance people’s physical and mental capacity to avoid suffering and lead fulfilling lives. The important task is not to understand if these ‘programs’ work, but to understand how interventions can best influence human behaviour at the lowest cost to taxpayers. A focus on program outcomes to answer the question of cost effectiveness is unlikely to provide many reliable answers in a human services context when interventions are rarely simple, often complicated and frequently complex.

Policy evaluations must recognise that in the real world complexity reins, and isolating and measuring impacts of interventions in a complex system is as easy as measuring the impacts of a butterfly flapping its wings. If decision makers want to know the most cost effective methods of achieving some change, they should focus on aspects they can influence and promote the most cost effective mechanisms that research has shown change behaviour.

**While evaluation needs research, it is not research**

Evaluation is not research. It does not happen in a controlled system. It must instead grapple with complex systems. Evaluation can assess the worth of an intervention by looking at effects on decision making by different people in different real world situations. But evaluation should not attempt to short circuit the research process required to identify mechanisms that generate behaviour change. It takes years of careful research (with all the controls required to isolate and measure casual factors) to identify stable relationships between mechanism of behaviour change and real world outcomes.

Policy makers need evaluation, and evaluation needs research. There will always be a prominent place for experimentation in research, but it must be done when it can be done scientifically in the types of simple systems they were designed for. It is not, an appropriate approach in every evaluation designed to inform policy decisions. Evaluations cannot be single trial research projects that attempt to measure outcomes of unstable interventions in complex systems. We must design programs based on sound research, but evaluate them based on the needs of improved decision making.

Evaluation must be useful to decision makers by providing timely and accurate evidence. Unlike research, which can be broad ranging and inconclusive, evaluation is often constrained to look at the effects of one thing (a program or intervention) on people’s decisions to do something or not. Research is never ending; evaluation is bounded by time and cost. Information obtained from evaluation should not cost more than the marginal benefit of that information.

**Part 3: data requirements for calculating the cost effectiveness of simple human service interventions OR how to do cost effectiveness if you could measure the impacts of a program**

This section is about the barriers to economic analysis even if reliable data on the outcomes of human services programs were available. The problem lies to a large extent in finding a common metric by which ‘outcomes’ can be expressed and compared to the ‘outcomes’ of other projects. As already described, with simple interventions, and sometimes with complicated ones, you may be able to convince yourself (and others) that you have measured the independent impact of the intervention. In this situation, standard cost effectiveness analysis as described below is possible.

**Calculating costs of a program or intervention**

Calculating the costs of programs is relatively straightforward. There are some decisions about what costs to measure: cost to government, costs to clients, or costs to the broader community. But these decisions can be made and data collected. The single biggest limitation on cost effectiveness analysis is that costs are often collected at the program level, as though every client consumed exactly the same amount of resource—the average. Calculations then typically assume every client shared in exactly the same amount of outcome—the average. And thus we have the typical, yet unsatisfactory, average cost effectiveness ratio (ACER). In short, some people cost more and benefit less from the same intervention, and vice versa. While some costs (such as program development costs) are not incurred at the individual level there is no choice but to use aggregate costs, but these can be allocated to individual on a pro rata or some other basis. In selecting interventions for different contexts it is important to be able to measure the cost of different dosage effects and who benefits most and least from a particular intervention. Once we have collected cost data at the individual level we need a common metric that can be used across human services to measure and compare the outcomes of different interventions.

**Cost effectiveness and the need for a common outcome metric**

It is commonly (but not universally) agreed in health economics that putting a dollar value on the benefits of a health intervention is too difficult. Instead of cost benefit much analysis focuses on benefits in terms of health outcomes. To maximise the comparability of the results of studies on benefits it is important to have a common outcome metric. This has generally been in terms of life years—or quality adjusted life years (QALYs) and disability adjusted life years (DALYs)[[2]](#footnote-2)—the idea being that all health care interventions aim to increase the quality and quality of life. How does this apply to human service interventions?

The extent to which this is possible will be the extent to which the cost effectiveness of human services programs can be calculated. If a common outcome metric cannot be identified we will be restricted to making comparisons within fields, such as crime prevention (perhaps offences prevented?), child protection (incidents of abuse or neglect prevented?) housing (days of homelessness avoided). A common outcome metric allows comparisons of the worth of programs regardless of their domain, so we could compare a drink drive recidivism program with family case management program and decide which programs should be funded because they generated the greatest benefit for the least cost (ignoring the political process and concerns about equity for the moment). If no common metric can be found then the cost effectiveness of different human services interventions cannot be compared.

What could a common metric be for human services interventions? All health programs aim to increase the quantity and quality of life, all economic programs aim to increase utility or material standard of living. Ultimately, all human services programs aim to reduce human suffering and increase human potential. A common metric for human services intervention could then measure a quantum of reduced suffering and increased potential.

To advance the case for a common metric based on psychology, the final part of this paper returns to the original thesis that real world outcomes of programs cannot be measured. Instead, recognising that most human services interventions aim to change behaviour, it attempts to assess the immediate effect of interventions on the factors known to influence human behaviour. In this approach what is important is not real world outcomes, but the extent to which different interventions affect factors known to drive behaviour.

**Part 4: the real purpose of social policy is not outcomes but behaviour change**

This paper has argued that the real world outcomes for which programs are developed, for example, to reduce domestic violence or homelessness, cannot be measured due to the inherent instability of programs, the complexity of social life, and the practical difficulties of expecting evaluation to replace careful long-term research. Instead it advocates a subtle but important shift in emphasis from real world outcomes to the immediate and local changes brought about by an intervention or mechanism. The approach is to employ a psychologist to conduct a marriage between realist and complexity theory.

The approach, somewhat ironically, still tries to measure the outcomes of an intervention like the experimentalist. The first difference is to shorten the timeframe for measurement to minimise the influence of feedback loops. Second, it focuses on the situational and personal factors that determine if behaviour is likely to change in the way anticipated, rather than trying to sort out the independent impact of a program in complex real world situations. So we sacrifice the salience of real world outcomes for the more measurable immediate and local changes in mental (or physical) states known to impact human decisions and behaviour.

Whether a program aims to help women escape domestic violence or address indigenous overcrowding, what it aims to do is help people make better decisions. Sometimes this means that real world resources need to be provided, such as a house to live in. These resources are important and their costs should be considered in any deliberation among different interventions. But while resources may be necessary, they may not be sufficient for behaviour change, which requires people to make choices and decisions. These are the outcomes we should try and measure—not the eventual outcomes over which we have little influence, but the degree to which we helped people make good decisions. This is what social policy is really about. Essentially the attitude is, we can help you make better decisions, but we cannot determine what choices will present themselves to you, or make those decisions for you. We would like to be judged on what we can influence, how well we prepare you for the world, rather than how you end up being treated by the world.

Indigenous housing policy is littered with examples of programs that describe the number of houses that need to be built to meet the housing needs of Indigenous people. More useful research looks at the reasons for overcrowding and, in particular, the reasons why empty dwellings exist side by side with overcrowded households. This kind of research can uncover the perverse incentives and psychological factors that mean it is in an individual’s self interest to remain in an overcrowded house. If a program aims to reduce overcrowding, it should be sensitive to the necessary conditions in terms of numbers of houses, but it should measure the success of an intervention in terms of the behaviour change required to maximise the use of all available housing.

**Part 5: towards a method for measuring cost effectiveness that focuses behaviour change rather than real world ‘outcomes’**

The search for a common metric to meet the basic requirements of cost effectiveness will not be easy. If we want to measure whether a self help DVD, a women’s refuge or a mass media campaign is the most cost effective way to address domestic violence we need a measure of success. This will not be easy and there will be other considerations (around ethics and equality not just efficiency, for example an ad campaign doesn’t help you much in crisis). Yet as all interventions try and change human behaviour by supporting people to make better decisions there is some commonality and comparisons may be made. The task that remains is to find these common factors that drive behaviour so we can measure the relative success of interventions.

**Eight factors influencing behaviour**

In 1991 a workshop in the USA that brought together different theorists to try and synthesise a theory about the factors influencing behaviour and behaviour change. The group came up with three necessary and sufficient factors that influence whether a person performs a given behaviour.

1. The person has formed a strong positive intention (or made a commitment) to perform the behaviour.
2. There are no environmental constraints that make it impossible for the behaviour to occur.
3. The person has the skills necessary to perform the behaviour.

They also identified five further factors that influence the strength and direction of intention.

1. The person believes that the advantages (benefits and anticipated positive outcomes) of performing the behaviour outweigh the disadvantages (costs, anticipated negative outcomes); in other words that person has a positive attitude to performing the behaviour.
2. The person perceives more social (normative) pressure to perform the behaviour than to not perform the behaviour.
3. The person perceives that performance of the behaviour is more consistent than inconsistent with his or her self-image; that is, performance does not violate personal standards that activate negative self sanctions.
4. The person’s emotional reaction to performing the behaviour is more positive than negative.
5. The person perceives that he or she has the capabilities to perform the behaviour under a number of different circumstances; in other words, the person has perceived self efficacy to execute the behaviour in question.

A common outcome measure for comparing different human services interventions that aim to change behaviour could draw on these factors. Rather than asking if domestic violence programs led to women escaping domestic violence, we may wish to compare a number of programs on the extent to which they increased the woman’s’ intention to leave domestic violence, removed environmental constraints, or provided necessary skills. These things are measurable, attributable to the program and useful to know about when planning effective interventions. Whether a participant in program x or y escaped domestic violence doesn’t tell you much about how to spend money in the future to maximise the benefits that can be provided by human services interventions.

Can we compare a domestic violence program with a family management program with a homelessness intervention and a sex offender intervention? Only to the extent that a common outcome measure makes sense. If it is accepted that the common thread in most human services intervention is changing human behaviour, relieving suffering and promoting human potential, consideration of the factors influencing behaviour will provide a better starting point for calculating the cost effectiveness of human services interventions. Human service interventions can provide resources, improve decision making and change behaviour, they cannot change the world.

**Conclusion**

Evaluation is the art of posing and finding answers to questions that help policy makers make ‘evidence-based’ decisions. Program funders want to know if the costs of the program justified the benefits, and rightly so. However, as long as they construe human services programs as things with real and measurable impacts in the complex systems in which they intervene they will encounter two core problems. The first is that programs are not things that exert a stable and consistent force that changes behaviour. The second is that real world outcomes of the kind required for conventional health or economic evaluations cannot be reliably measured when intervening in complex systems. Instead, funded programs exist as interventions that provide opportunities for people to change behaviour. We have no control over eventual outcomes in the world. It is the degree to which interventions prepare people to make good decisions when confronted with the real world that should be the basis for comparing worth.

Evaluation must focus on what we need to know to make better decisions, rather than what we wish to know about the world. In short this paper says, don’t try and measure things that are impossible to adequately define and too complex to measure. Instead, do a better job of measuring what is common in attempts to improve decision making and change behaviour.

Maybe one day we will better understand complexity so as to be able to measure the impact of programs or interventions in the real world—but we can’t yet. To attempt to predict how a program or intervention will generate long-term, real world outcomes in complex systems is like asking the weather bureau to issue a long-term weather prediction.

While we wait for better measures of the impact of interventions in complex systems, a better question about the cost effectiveness of human services programs may be: what is the lowest cost way to maximise common factors that help people make better decisions, reduce suffering and increase human flourishing?

1. [Laurie Penny](http://english.aljazeera.net/indepth/opinion/profile/laurie-penny.html) 2011, *Panic on the streets of London: Speculations circle as to why the London riots have become so big, but the answer is quite obvious*. Aljazeera, Last Modified: 09 Aug 2011 16:26 [↑](#footnote-ref-1)
2. Interestingly, while cost effectiveness in health economics tends to focus on ultimate outcomes of interest—life years—traditional cost benefit analysis tends to focus on first order impacts rather than attempting to measure long-term outcomes. As already argued, cost effectiveness in human services may be more accurate if it limited its focus to first order, or immediate impacts rather than trying to measure long term real world outcomes. [↑](#footnote-ref-2)