Deliberating Policy: Where morals and methods mix – and not always for the best

Nancy Cartwright, UCSD and Durham University

CHESS Working Paper No. 2014-02
Durham University
July 2014

Deliberating Policy: Where morals and methods mix – and not always for the best

Nancy Cartwright

Nancy Cartwright
Department of Philosophy
UCSD and Durham University
nancy.cartwright<at>durham.ac.uk

In 2004 in the London Borough of Haringey, 17-month-old Peter Connelly was found dead in his crib. The child had suffered fractured ribs and a broken back after months of abuse at home. His mother, her partner and a lodger were jailed for his death. Peter had been seen by health and social services professionals from Haringey Council 60 times in the eight months before he died.

There were two kinds of government responses to this that I shall discuss. First, the Minister of Education Ed Balls sacked the Director of Children's Services in Haringey, Sharon Shoesmith, with immediate effect in a live press conference on television. Ms Shoesmith defended herself and her Services: 'We should not be put into blame'; it does not produce 'anything productive' and obscures 'the bigger picture'. A BBC news interviewer argued to the contrary: If nobody accepts the blame, '…how can we stop this happening again?'

A second response was that of then Prime Minister Tony Blair (Joseph Rowntree Foundation Speech, York, 2006). He argued that the government can make children and young people more safe by identifying at-risk families and intervening early on behalf of the children in those families:

> Let me summarise my argument. I am not talking about...trying to make the state raise children, or interfering with normal family life. I am saying that where it is clear, as it very often is, at young age, that children are at risk of being brought up in a dysfunctional home where there are multiple problems, say of drug abuse or offending, then instead of waiting until the child goes off the rails, we should act early enough, with the right help, support and disciplined framework for the family, to prevent it... It may be the only way to save them and the wider community from the consequences of inaction.

According to Blair, "We can predict. We can then...'intervene'."

Both these responses are morally questionable. I do not mean by this that they are wrong policies; I mean only just that: that they are morally questionable. Look first at Blair's. Blair's program is intended to identify at risk-families and offer help. But there is evidence that labeling families 'at risk' can increase the anxiety of the parent and thereby increase the likelihood of abuse or neglect. Also, as other experts report, "Parents who have been through the process of child protection investigations and registration are often bruised and stigmatized by the experience and wary of accepting the help or support which may follow". So the program may cause harm overall.

Also the question of parents' rights and family autonomy looms. As Elizabeth Brake and John Millum (Parenthood and Procreation , Stanford Encyclopedia of Philosophy, 2013) put it: "Parents have moral and legal rights regarding their children. They have the liberty to make decisions on behalf of their children regarding matters such as diet, schooling, [and] association with others". Also, as others argue, the government acts paternalistically when it "aim[s] to take over or control what is properly within the [individual's] own legitimate domain of judgment or action" (Shiffrin 2005 p216). In giving directions to parents the government substitutes its judgment of how to raise their child in place of the parent's judgment. So all this considered, even if the interventions will produce the predicted benefits, there remains a question about whether such interventions are justified.

As to blame: blame is retributive, it is often vindictive, it attacks the moral character of the culprit, not the deed, it vilifies the culprit, and as Gareth Williams explains 'there is clear evidence from social psychology that blame is frequently – and inappropriately – attributed to [the free choices of] individuals in modern Western societies', tending to 'overestimate the extent to which people's behavior is due to internal, dispositional factors, and to underestimate the role of situational factors.' Blaming a person is more than grading them negatively. It is, as Jay Wallace argues, to make them the object of negative emotions – resentment, indignation – and subject to adverse treatment – avoidance, reproach, scolding, denunciation, remonstration, and punishment.

Sharon Shoesmith was later interviewed by the BBC. She described the effects of her 'vilification' to the interviewer

"It was horrific. It was frightening. I was terrified. I had death threats. I was afraid on the streets of London. It wasn't just me and my imagination - the police were advising me that I was probably at risk. When people take photographs of you on the trains and on the buses and point you out and start to shout 'that's that woman' you're fearful of where that can go." [See http://www.bbc.co.uk/news/uk-13570959.]

Here is a case from the US. Psychologist Robert Hare developed a test to identify characteristics of psychopaths, like lack of empathy, lack of remorse, lack of guilt. One study in which the test was administered to prisoners found that those who did not have the characteristics identified by Hare were re-convicted of a crime within 5 years about 20-25% of the time. Those who did have the characteristics were re-convicted within 5 years about 80% of the time. Once this study was released (in the 1980s), parole boards across the US and Canada began using the test when considering the release of prisoners. In several states parole boards are even mandated by law to do so. Because the psychopathy test predicts a high rate of recidivism, it is unlikely prisoners with Hare's characteristics will be paroled, considering the risk to the public -- and the political risk: if a people with 'known' psychopathic characteristics re-offend, the parole board is in trouble. So there is little incentive for parole board to release those with Hare's characteristics from prison.

Yet many moral and religious points of view would think that people should not be kept in prison because of predictions about their future behavior. Prison is for people who have committed crimes (in the past-tense). Those up for parole have already served the minimum sentence for their crime. Yet for those prisoners with Hare's characteristics, other considerations favoring their release on parole are overridden by the results associating those characteristics with a high recidivism rate.

Here is yet another example reported last winter recently on the front page of the *NY Times*. In an effort to cut down on robbery, the New York City Police are aggressively intervening in the lives of young people. As they describe it, they aim to 'make them radioactive' and thus isolate them from their friends. Which young people? - Those 'destined for trouble' and 'most likely to commit' these crimes.

So we have a number of policy responses – Ed Ball's and Tony Blair's reactions to the death of Peter Connelly; a 'psychopath' test to refuse parole; and the NYPD's targeting of 'types of teenagers destined for trouble' - all of which are morally questionable; and all of which offer promise to succeed, promise based on what we can claim to know. Now, a good policy decision always requires a mix of considerations: who benefits? who suffers? who pays? how much? what are possible good side effects? bad ones? will the effects last? Etc. Central among these are issues of effectiveness and issues of legitimacy: Will the policy achieve the desired ends? And, is it morally, politically, culturally acceptable? Few policies will be all to the good for all concerned; few have only a moral upside and no moral downside. So it is inevitable that a balance be struck. In particular it can be perfectly acceptable to adopt a policy that is morally questionable or that has morally negative aspects if we can be sure it will achieve good ends, so long as the balance is reasonable and we operate within a range of what is at least morally permissible.

My concern is with cases where we get the balance wrong because we are overconfident in our predictions about what the policy results will be. This in turn generates concerns about the current drive for evidence-based policy (EBP) which is advocated, indeed often mandated, across the policy domains, from medicine to education to crime to economic development. Of course in general we will get more reliable predictions about policy outcomes if we take into account the evidence than if we ignore it. But there is the promise – or perhaps just the hope – of far more certainty than the evidence can deliver. The danger of these mistakes is that they encourage an unjustified degree of optimism about how effective our policies can be, in which case we are likely to get policy deliberation wrong in the delicate balance between considerations of effectiveness and considerations of legitimacy.

Here then is the basic problem I want to underscore: Morals and methods ought to mix when it comes to policy deliberation. That's because---leaving costs aside---whether a particular policy should be implemented depends on whether

(1) The policy will be effective, i.e. it will produce the expected effects,

and on whether

(2) The policy is morally, socially, politically, culturally acceptable.

But they don't always mix in the way they should, because we often focus on (1) and discount the importance of (2). A plausible explanation for this is that we think we have methods which will provide us with objective and fairly certain answers to the effectiveness question, e.g. randomized controlled trials -- RCTs. By contrast, we do not have methods that can give us objective and certain answers to the moral question; here, things are much muddier and open to debate.

This slide from

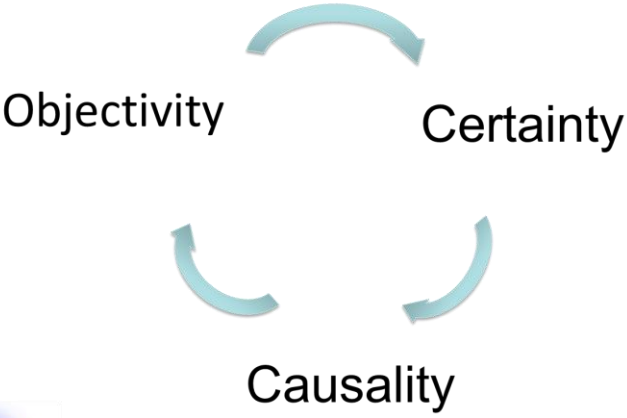
A: The methods we have for ascertaining (1) are better than the ones we have for ascertaining (2)

to

B: We should give (1) more weight than (2) in policy deliberations

is eased along by the prestige that evidence-based policy (EBP) enjoys. The point of EBP is to ensure that effectiveness predictions are based on sound evidence. But the expectations EBP creates can lead us to give this evidence and the predictions it supports greater weight than concerns of moral acceptability. Effectiveness can become the cardinal value in policy deliberations, as I think we are now seeing in UK political calls for coercive interventions into families that seriously changes the state-family relationship.

What I want to do in the rest of this talk is to hack away some mistaken philosophical stances that can make effectiveness considerations loom larger than they should. These involve a circle of mistaken ideas about objectivity, certainty, and causality:



Objectivity

Certainty

Causality

There are three mutually supporting stances:

4. We bank on certainty.

5. We suppose objectivity is the path to certainty.

   Objectivity = elimination of the subject, especially judgment; the use of methods that have manuals that fix correct procedures.

6. We assume that causality is ‘linear’ and that it is God given.

EBP champions objectivity and certainty in social policy deliberation. It insists that, for policy evaluation and prediction, we rely only on ‘objective’ methods like randomized controlled trials (RCTs) that can provide certainty: in the ideal RCTs can clinch causal claims – and they can so without the intrusion of ‘subjective judgment’. From this position we slide easily into our third problematic assumption, that causality is linear and God given.

Look at linearity first. The slide is easy here -- and easy not to notice. That’s because, looking through the lens of RCTs, complex causal webs get projected onto a line. Here is a picture of the causal process as RCTs tell it:

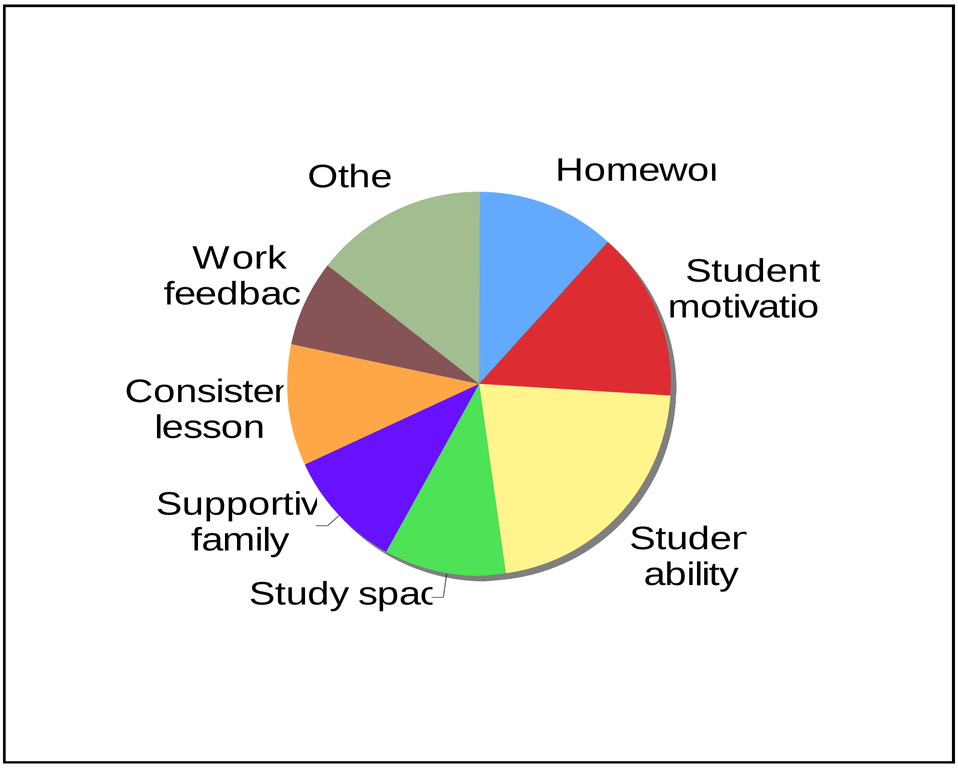
Domino causation

There are two different senses of ‘linear’ involved in this image, and we tend to suppose both.

1 – There is no single sufficient cause.

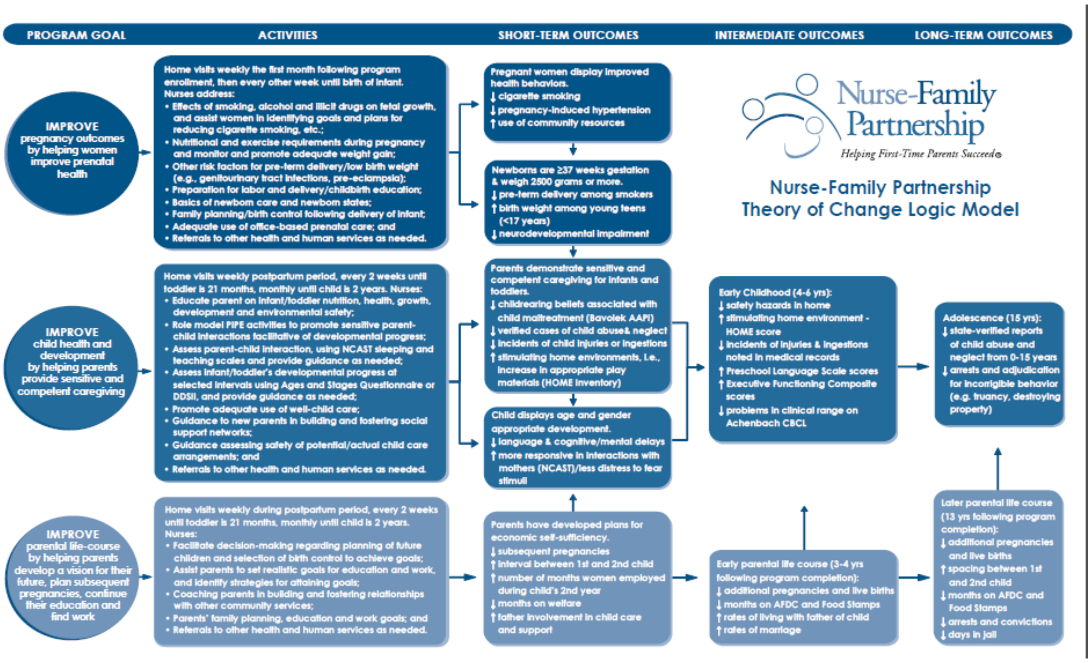
What we label ‘the cause’ – the policy whose results we aim to predict or the actions we want to blame for some disastrous outcome – is seldom enough on its own to produce the effect in question. It needs help. Generally a whole team of support factors must act together with the highlighted cause or no contribution to the effect will be produced. Epidemiologists illustrate this with what they call ‘causal pie diagrams’. Here is an example based on work on the effectiveness of homework by Harris Cooper.

Homework is one condition contributing to higher test scores. Other conditions are necessary to ensure that homework is maximally effective. These other conditions include student motivation, student ability, access to a proper study space, supportive family, getting a consistent message from teachers and parents, receiving teacher feedback on assignments.



I refer to these not as causal pies but as causal pancakes. To make pancakes you need: flour, milk, eggs, and baking powder; and you need them all if you are to get a pancake at all. With just 3 of the 4 ingredients, you don't get ¾ of a pancake – you don't get a pancake at all. Similarly for the ingredients in the causal cake diagrams: all the ingredients are required together or you not get the expected level of effect at all.

The cake diagrams make vivid a crucial point. But they also make it look too simple. For most policies the connection between cause and effect is not immediate: there is a long chain of steps in between; each one has to occur at the appropriate time to lead on the next. Consider this diagram for the Nurse Family Partnership.



PROGRAM GOAL — ACTIVITIES — SHORT-TERM OUTCOMES — INTERMEDIATE OUTCOMES — LONG-TERM OUTCOMES

Nurse-Family Partnership
Helping First-Time Parents Succeed®

Nurse-Family Partnership
Theory of Change Logic Model

IMPROVE pregnancy outcomes by helping women improve prenatal health

Home visits weekly the first month following program enrollment, then every other week until birth of infant. Nurses address:
- Effects of smoking, alcohol and illicit drugs on fetal growth, and assist women in identifying goals and plans for reducing cigarette smoking, etc.;
- Nutritional and exercise requirements during pregnancy and monitor and promote adequate weight gain;
- Other risk factors for pre-term delivery/low birth weight (e.g., genitourinary tract infections, pre-eclampsia);
- Preparation for labor and delivery/childbirth education;
- Basics of newborn care and newborn states;
- Family planning/birth control following delivery of infant;
- Adequate use of office-based prenatal care; and
- Referrals to other health and human services as needed.

Pregnant women display improved health behaviors.
↓ cigarette smoking
↓ pregnancy-induced hypertension
↑ use of community resources

Newborns are ≥37 weeks gestation & weigh 2500 grams or more.
↓ pre-term delivery among smokers
↑ birth weight among young teens (<17 years)
↓ neurodevelopmental impairment

IMPROVE child health and development by helping parents provide sensitive and competent caregiving

Home visits weekly postpartum period, every 2 weeks until toddler is 21 months, monthly until child is 2 years. Nurses:
- Educate parent on infant/toddler nutrition, health, growth, development and environmental safety;
- Role model PIPE activities to promote sensitive parent-child interactions facilitative of developmental progress;
- Assess parent-child interaction, using NCAST sleeping and teaching scales and provide guidance as needed;
- Assess infant/toddler's developmental progress at selected intervals using Ages and Stages Questionnaire or DDSII, and provide guidance as needed;
- Promote adequate use of well-child care;
- Guidance to new parents in building and fostering social support networks;
- Guidance assessing safety of potential/actual child care arrangements; and
- Referrals to other health and human services as needed.

Parents demonstrate sensitive and competent caregiving for infants and toddlers.
↓ childrearing beliefs associated with child maltreatment (Bavolek AAPI)
↓ verified cases of child abuse& neglect
↓ incidents of child injuries or ingestions
↑ stimulating home environments, i.e., increase in appropriate play materials (HOME Inventory)

Child displays age and gender appropriate development.
↓ language & cognitive/mental delays
↑ more responsive in interactions with mothers (NCAST)/less distress to fear stimuli

Early Childhood (4-6 yrs):
↓ safety hazards in home
↑ stimulating home environment - HOME score
↓ incidents of injuries & ingestions noted in medical records
↑ Preschool Language Scale scores
↑ Executive Functioning Composite scores
↓ problems in clinical range on Achenbach CBCL

Adolescence (15 yrs):
↓ state-verified reports of child abuse and neglect from 0-15 years
↓ arrests and adjudication for incorrigible behavior (e.g. truancy, destroying property)

IMPROVE parental life-course by helping parents develop a vision for their future, plan subsequent pregnancies, continue their education and find work

Home visits weekly during postpartum period, every 2 weeks until toddler is 21 months, monthly until child is 2 years. Nurses:
- Facilitate decision-making regarding planning of future children and selection of birth control to achieve goals;
- Assist parents to set realistic goals for education and work, and identify strategies for attaining goals;
- Coaching parents in building and fostering relationships with other community services;
- Parents' family planning, education and work goals; and
- Referrals to other health and human services as needed.

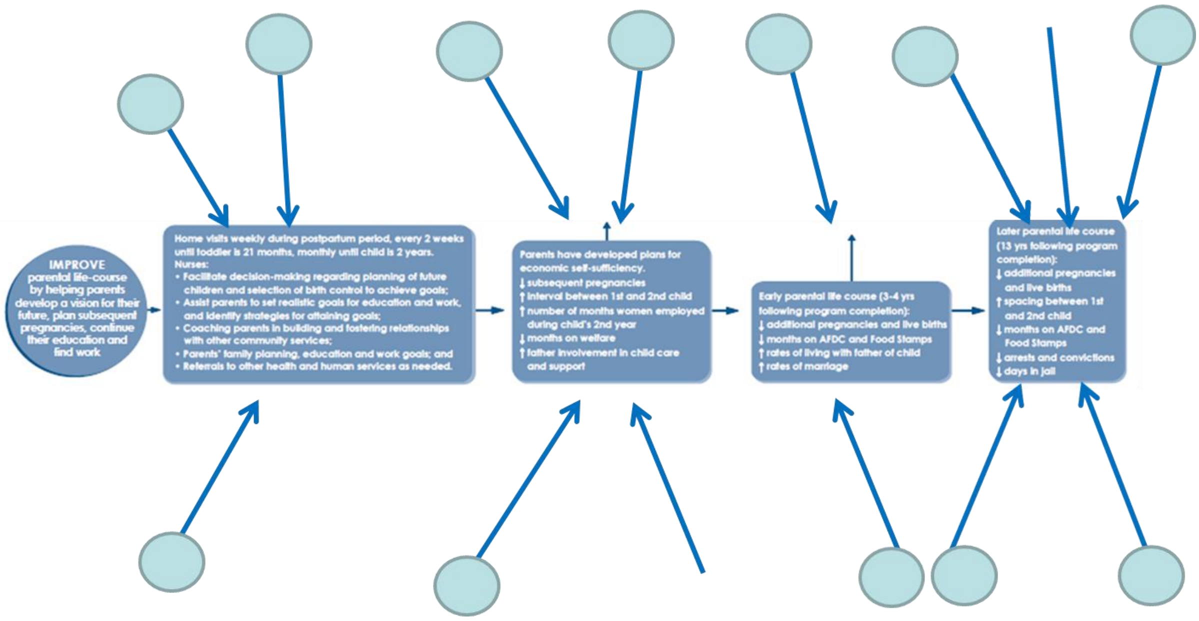
Parents have developed plans for economic self-sufficiency.
↓ subsequent pregnancies
↑ interval between 1st and 2nd child
↑ number of months women employed during child's 2nd year
↓ months on welfare
↑ father involvement in child care and support

Early parental life course (3-4 yrs following program completion):
↓ additional pregnancies and live births
↓ months on AFDC and Food Stamps
↑ rates of living with father of child
↑ rates of marriage

Later parental life course (13 yrs following program completion):
↓ additional pregnancies and live births
↑ spacing between 1st and 2nd child
↓ months on AFDC and Food Stamps
↓ arrests and convictions
↓ days in jail

Already this picture is more complicated than my simple domino image since the policy initiates not just a single causal chain but 3 different policy actions that lead by interwoven chains of intermediate effects to the targeted outcomes – less child abuse and fewer young people arrested.

Focus on the bottom line, which looks like a straightforward linear sequence. To describe it thus is to miss the point about support factors and causal cakes: what we picture as 'the cause' typically cannot produce the effect on its own but needs the help of a whole team of support factors. That's going to be true for each step in this causal sequence. There's not just one causal cake here but a different causal cake for each step.


IMPROVE parental life-course by helping parents develop a vision for their future, plan subsequent pregnancies, continue their education and find work
Home visits weekly during postpartum period, every 2 weeks until toddler is 21 months, monthly until child is 2 years.
Nurses:
• Facilitate decision-making regarding planning of future children and selection of birth control to achieve goals;
• Assist parents to set realistic goals for education and work, and identify strategies for attaining goals;
• Coaching parents in building and fostering relationships with other community services;
• Parents' family planning, education and work goals; and
• Referrals to other health and human services as needed.
Parents have developed plans for economic self-sufficiency.
↓ subsequent pregnancies
↑ interval between 1st and 2nd child
↑ number of months women employed during child's 2nd year
↓ months on welfare
↑ father involvement in child care and support
Early parental life course (3-4 yrs following program completion):
↓ additional pregnancies and live births
↓ months on AFDC and Food Stamps
↑ rates of living with father of child
↑ rates of marriage
Later parental life course (13 yrs following program completion):
↓ additional pregnancies and live births
↑ spacing between 1st and 2nd child
↓ months on AFDC and Food Stamps
↓ arrests and convictions
↓ days in jail

If we want to identify the support team necessary for the initial NFP causes to produce the targeted outcomes, we have to gather all the members of all the support teams from each stage and graph them together in one huge causal cake.
Recall: the point about causal cakes is that all their ingredients have to be in place or you don't get the effect. To the extent that any of the necessary ingredients is uncertain, so too is the final outcome. But – look at our circle of problems. We bargain for certainty. The simple linear causal model makes this look a far better bargain than it generally is. So:

1. We often expect results that can't be achieved, which leads to wasted money and effort and to heartbreak and dashed hopes;

2. we don't work to put in place the support factors that can help make our policies work because we haven't noticed the need for them;

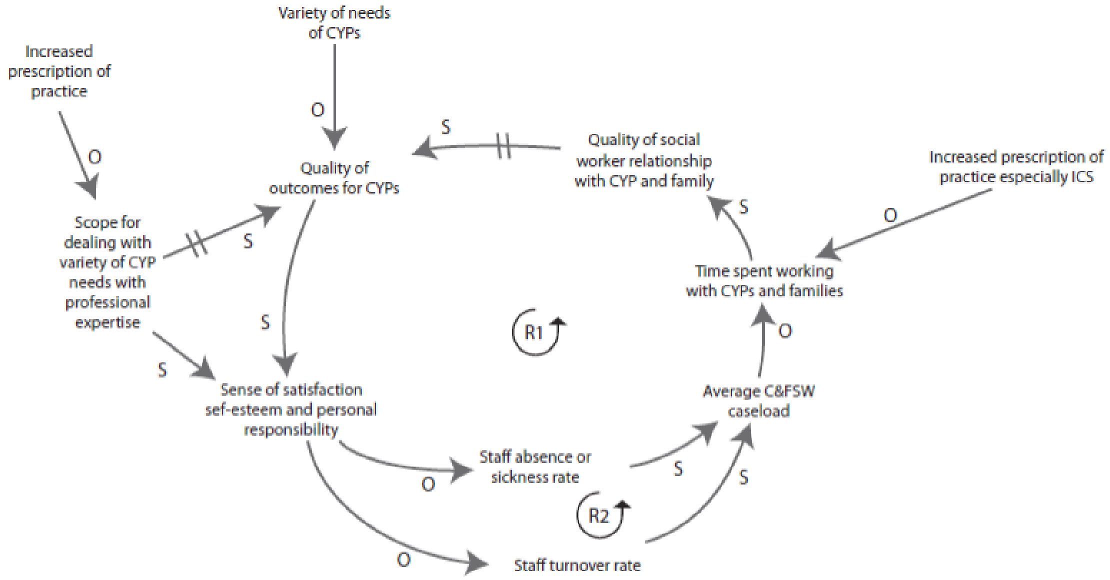
3. we blame perfectly good policies for failing that could achieve better results in better circumstances; and

4. we despair of doing anything because we cannot find the miracle cure-all.

The linear model and the omission of support factors also predisposes us to focus efforts on eliminating harmful causes at the head of a sequence, like family drug and alcohol abuse, which can be a tall order. But it can be just as effective to remove support factors anywhere along the causal path. Consider the growing body of research on resilience factors. 'Resilience' describes the product of a combination of mechanisms for coping in the face of adversity. Evidence from longitudinal studies suggests that many children can recover from short-lived childhood adversities with little detectable impact in adult life. Encouraging resilience is important because resilient children are better equipped to resist stress and adversity, to cope with change and uncertainty, and to recover faster and more completely from traumatic events or episodes.

Linear Models Ignore Cycles

Linear models don't have cycles in them. But cycles can matter. Consider the UK's recent Munro Review of Child Protection. The Review notes that policies, even good ones, can figure in negative cakes, alongside positive ones. The negative cakes diminish the good effects of the policy, and can even, if they are strong enough, outweigh the good effects. This is just the trouble that the Munro Review pinpoints for one of the big UK child welfare policies. The policy was intended to improve welfare outcomes in children and young people (that's what 'CYP' means in the graph) by providing stricter guidelines for what social workers must do in dealing with children and families and by better monitoring of what they are doing: by ensuring that specific mandated facts about the family and the child are ascertained and recorded properly and that all required meetings take place. But this policy, the Review argues, can have serious negative effects alongside the intended positive ones. How so? Through various negative feedback loops. Have a look at this diagram from the Review:


Variety of needs of CYPs
Increased prescription of practice
O
O
S
Quality of outcomes for CYPs
Quality of social worker relationship with CYP and family
Increased prescription of practice especially ICS
Scope for dealing with variety of CYP needs with professional expertise
S
S
O
Time spent working with CYPs and families
S
R1
O
S
Sense of satisfaction sef-esteem and personal responsibility
Average C&FSW caseload
Staff absence or sickness rate
O
S
S
R2
O
Staff turnover rate

An arrow linking variable A to variable B should be read as 'a change in the value of A produces a change in the value of B'. The qualitative nature of the link is indicated by a 'link polarity'. These should be read as:

- 'S': the variables move in the same direction *ceteris paribus*, so a change in variable A produces a change in variable B in the same direction: if A goes up, B goes up.
- 'O': the variables move in the opposite direction *ceteris paribus*, so a change in variable A produces a change in variable B in the opposite direction: if A goes up, B goes down.
- double bars on a link indicate a particularly long delay in the causal connection.

Note that the link polarity says nothing about the size, or quantity of the change. The indication of the effect is qualitative only. Moreover, there is no presumption of a linear relationship between the two variables.

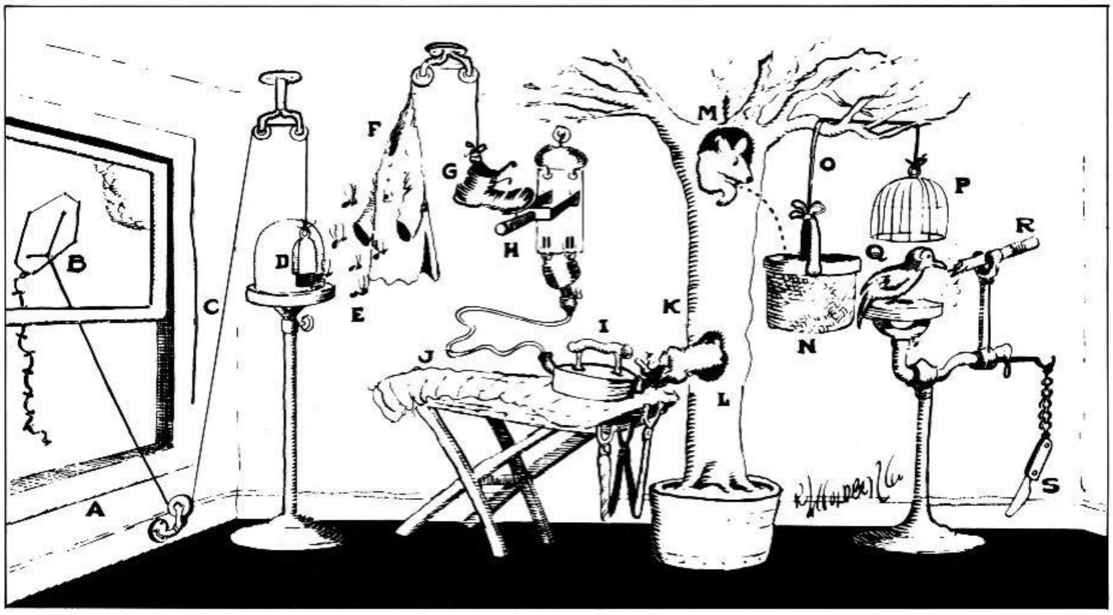
Two negative loops are pictured, R1 and R2. Both start in the same way. Increasing the amount of prescription imposed on social workers, you can reduce their sense of satisfaction and self-esteem. In R1, this increases staff sickness and absence rates; in R2, it increases staff turnover rates. Both these effects tend to result in an increase in average social worker caseload, which leads to social workers spending less time with the children and young people and their families. This in turn reduces the quality of the social workers' relationships with the children and their families, which then reduces the quality of the outcomes. So the policy may produce bad unintended consequences. Worse, these negative effects can become amplified via the feedback loops. When the outcomes are regularly too unsatisfactory, this reduces social workers' sense of self-esteem and personal responsibility, and the negative cycle is set in motion again.

God-Given Causality

Besides the habit of taking causality as linear, we take it to be God given. But the kinds of causal principles we rely on for policy prediction are not God given. They depend on intricate underlying structures. And our focus on objectivity and certainty takes or attention away from these underlying structures.

To make this point vivid I use an example not from social policy but from my own daily policies. I often write my lectures on paper and with a sharp pencil. I sharpen my pencils by putting a kite out my study window. I can do that because my study was designed by Rube Goldberg.

Pencil Sharpener



The Professor gets his think-tank working and evolves the simplified pencil sharpener.

Open window (**A**) and fly kite (**B**). String (**C**) lifts small door (**D**), allowing moths (**E**) to escape and eat red flannel shirt (**F**). As weight of shirt becomes less, shoe (**G**) steps on switch (**H**) which heats electric iron (**I**) and burns hole in pants (**J**). Smoke (**K**) enters hole in tree (**L**), smoking out opossum (**M**) which jumps into basket (**N**), pulling rope (**O**) and lifting cage (**P**), allowing woodpecker (**Q**) to chew wood from pencil (**R**), exposing lead. Emergency knife (**S**) is always handy in case opossum or the woodpecker gets sick and can't work.

Putting a kite out the window is a very effective policy for me to get nice sharp pencils. Still, I don't advise you to fly a kite to sharpen your pencils. Kite flying undoubtedly figures in the causal principles that govern pencil sharpening in my study. It would, for instance, pass any rigorous RCT. Put the kite out the window on randomly chosen days and you will certainly get more sharp pencils when you put it out than when you don't. But that principle is local. It depends on the underlying structure of my study. The causal role played by kite flying in my study is not God given: it depends on a complex pattern of interactions in an intricate underlying structure.

Of course mine is not a typical social policy case. Social policies suppose principles like 'Burnout causes turnover in child welfare service workers', or 'Age does not.' Or 'Apathetic-futile mothers are more likely to maltreat their children.' These are clearly not God-given either. And surely it is implausible to suppose that getting a good social regularity, as these are purported to be, depends less on the details of the underlying structure than getting regularities between pure physical quantities.

Note too that I am not supposing that there are no human universals, that people in Bangladesh villages are essentially different from those in New York high rises, nor across the 300 language groups and more than 50 non-indigenous communities who live in London. In fact my Rube Goldberg example of a local causal principle works in just the opposite way, by relying on other causal principles that hold widely. The pencil sharpener depends on a number of fairly universal principles – from the laws of the lever and the pulley to the familiar fact that moths eat flannel –to ensure that the arrangement and interaction of the components results in the causal principle I use to sharpen my pencils. So the fact, if it is one, that there are a large number of universal truths about human behaviours, emotions and reactions goes no way to showing that the kinds of causal principles we rely on in typical social policies will be anything other than very local.

Our aspirations for certainty divert our attention to these kinds of local causal principles since they are ones that we can nail down with objective methods, like RCTs: 'Flying kites in NC's study sharpens pencils'; or from J-PAL, the MIT-based Jameel Poverty Action Lab, after a study in certain Indian villages: 'Informing villagers of poor teaching in their villages and raising awareness of accountability mechanisms had no impact on teacher attendance.' Our efforts are taken away from the more difficult study of the underlying structures that make these casual principles possible.

When it comes not to prediction but to evaluation – looking back to see what was responsible for an outcome, the focus on linear causal principles, with their objective certifying methods, leads to skewed views about human error and individual responsibility. As Eileen Munro of the Munro Review explains elsewhere (Improving practice : child protection as a systems approach, Child and Youth Services Review 2004) , when a tragedy like the death of Peter Connelly occurs, 'The standard response is to hold an inquiry, looking in detail at the case and trying to get a picture of the causal sequence of events that ended in the child's death…We are tracing a chain of events back in time to understand how it happened.…' More: 'Unlike the police investigation, which focuses on the perpetrators of the homicide, these inquiries focus primarily on how the professionals acted, judging them against the formal procedures for working with families and principles of good practice.'

Where does this backwards tracing stop? As Munro argues, the 'events that bring the investigation to a halt usually take the form of human error. Practitioners did not comply with procedures or lapsed from accepted standards of good practice.' But as the UK Dept of Health pamphlet explains, 'There are two ways of viewing human error: the person-centred approach and the system approach. The [person-centred] … approach focuses on the psychological precursors of error, such as inattention, forgetfulness and carelessness. Its associated countermeasures are aimed at individuals rather than situations and these invariably fall within the "control" paradigm of management. Such controls include disciplinary measures, writing more procedures to guide individual behaviour, or blaming, naming and shaming' – as we saw at the start in the case of Peter Connelly with Ed Balls and the first BBC interviewer.

But as the UK Dept of Health notes:

> Aside from treating errors as moral issues, [the person-centred approach] isolates unsafe acts from their context, thus making it very hard to uncover and eliminate recurrent error traps within the system…
>
> The system approach, in contrast, takes a holistic stance on the issues of failure. It recognises that many of the problems facing organisations are complex, ill-defined and result from the interaction of a number of factors.

Just as in my Rube Goldberg pencil sharpener!

The same worry is studied in the US National Academy of Sciences' *To Err Is Human: Building a Safer Health System*:

> The title of this report encapsulates its purpose. Human beings, in all lines of work, make errors. Errors can be prevented by designing systems that make it hard for people to do the wrong thing and easy for people to do the right thing. Cars are designed so that drivers cannot start them while in reverse because that prevents accidents. Work schedules for pilots are designed so they don't fly too many consecutive hours without rest because alertness and performance are compromised.

The NAS report urges:

‘The focus must shift from blaming individuals for past errors to a focus on preventing future errors by designing safety into the system.’ Or, to put it in the terms I have been using, we should be less concerned with the easier to certify causal sequences that start with human error and end with disastrous consequences and far more with understanding – and restructuring – the underlying structures that make this kind of causal sequence likely. As Eileen Munro notes:

> When society is shocked and outraged by a child’s terrible tale of suffering, there seems a basic human desire to find a culprit, someone to bear the guilt for the disaster and to be the target of feelings of rage and frustration.

This puts us squarely in the business of finding these local linear causal principles; and, with Tony Blair, we can feel morally and epistemically safe in doing so – we are not likely to cast blame in the wrong places – because these are the kinds of claims about which with due care our objective methods can deliver reasonable certainty. But the kinds of preventative measures this leads to – recall the Dept of Health examples: disciplinary actions, writing more procedures to guide individual behaviour, or blaming, naming and shaming – these measures are often unlikely to stop these kinds of sequences occurring. As Munro urges: ‘Child protection is a systems problem’.

And so too are a good many other social problems, from poor childhood nutrition in Bangladesh and poor school attendance by teachers in Indian villages to crime, education, health and climate change adaptation almost anywhere. Our thirst for certainty and our admiration for methods that can be run by rules must not lead us to buy cheap knowledge that can’t serve our needs.

For a timely illustration of linear causal thinking in child protection consider the Serious Case Review into the death of Daniel Pelka, a 4-year-old who died at the hands of his mother and stepfather in March 2012 in Coventry. His death provoked a massive outcry across the UK in large part because, according to the *Serious Case Review* (SCR) commissioned by the Coventry Children Safeguarding Board, the social workers (as well as teachers and police officers) who had been in contact with Daniel and his family in the months leading up to his death missed 26 opportunities to help Daniel and to act in a way that would have prevented his death.

Because the SCR blames a) individuals for b) failing to behave in certain ways, the lessons it draws are about the ways those individuals should have behaved. In this case the SCR ties differences in behavior to differences in attitude towards available evidence. Lesson 15.4 for instance states: "Domestic abuse/violence is always a child protection issue and must always be approached with this as the **mind-set** of professionals."

Lesson 15.13 is: "**Professional optimism** about a family and of their potential to change or improve their parenting must be supported by objective evidence and that any contra indicators have been fully considered prior to any **optimistic stance** being taken."

These make it sound as if the social workers, teachers, doctors, and police officers involved looked at the evidence with rose-colored glasses, which I do not think the evidence reviewed in the SCR supports. This seems to be the view of Eileen Munro as well from her comment in a public interview that she can't claim she would have done better on the evidence presented in the SCR report.

In the same vein, the SCR (p. 6) castigates the professionals involved along these lines:

> In this case, professionals needed to "think the unthinkable" and to believe and act upon what they saw in front of them, rather than accept parental versions of what was happening at home without robust challenge. Much of the detail which emerged from later witness statements and the criminal trial about the level of abuse which Daniel suffered was completely unknown to the professionals who were in contact with the family at the time.

Blaming the professionals involved for failing to "think the unthinkable" in this way is odd. What "they saw in front of them" clearly wasn't that Daniel was being abused in horrific ways by his parents. For instance the pediatrician was blamed for not including abuse as a likely diagnosis when a mother brings a child to him because she is concerned that he is losing weight. But deliberate starvation as the cause is highly improbable with lots of physical illnesses being more probable. The sordid details of Daniel's abuse only emerged after a full-blown criminal investigation which involved searching the house and the parents' cell-phones, which is not, as I understand it, something the Coventry social workers or even the police could have done at an earlier stage. These social workers, by contrast, had access to concrete evidence supporting the view that Daniel was not being abused by his parents, as the SCR itself reports.

Besides the morally questionable aspects of placing heavy blame on individual social workers, the way the SCR assigns blame in the Daniel Pelka case seems to suppose that the way to prevent harmful outcomes like Daniel's death in the future is, to put it briefly, to turn the professionals involved--- who are blamed for failing to detect abuse -- into better and better detectors of abuse, where 'better' means 'fewer false negatives' (and correlatively more false positives, with more families subjected to nasty investigations and sometimes losing their children at a lower level of evidence). From the evidence in the SCR itself, the Daniel Pelka case has the earmarks of a systems issue since it is the result of many minor flaws in practice plus the basic incompleteness of information rather than one devastating individual act. Each failure of each individual was just one of many ingredients in the causal cake that led to the failure to see how much danger Daniel was in. So it is not obvious that improving the 'detectors of abuse' in that system is the best way to improve child welfare and reduce the probability of Daniel-level abuse occurring again. Perhaps the better place to devote our attentions is to redesigning the system so it is easier to do it right and harder to do it wrong. That discussion recall was all in the context of worries about linear models of causation.

My final worry about trusting that good objective methods can deliver high degrees of certainty is that these methods only deliver answers in the language in which we ask our questions, and that is often not the language in which the system under study operates. So very often we don't know what we are testing with these methods. In one sense, we do; we know all too well, and that is the problem. Our best objective methods for testing causal claims, like RCTs, require a precise characterization of both the cause and the effect. This is a crucial part of the study protocol. We must ensure that everyone in the treatment group gets the same treatment, and that there are strict criteria for deciding whether the effect has occurred. Otherwise the validity of the conclusion is impugned. This in turn means that treatment and effect descriptions are couched in concrete operational terms.

There's the rub. For policy prediction we need causal principles that hold widely – at least widely enough to cover both the study situation and the target. But these kinds of principles often relate not concrete concepts of the kind operationalized in a good study, but far more abstract ones, like the concepts of 'lever' and 'pulley' that matter in my Rube Goldberg pencil sharpener. This kind of problem looms large for social policy because the same thing doesn't mean the same in different social settings.

Consider the Integrated Nutrition Program that provides nutrition education to mothers to improve the health of their infants. This program was a success in Tamil Nadu. But not in Bangladesh. According to the World Bank post hoc analysis, that's because in many rural Bangladeshi households men do the shopping, not the mother, and the mother-in-law is in charge of food distribution.

There was a generalizable cause producing improved infant health in Tamil Nadu.

The generalizable cause was the nutritional education of a person

5.who is responsible for household food selection, and

6.who is responsible for food distribution, and

7.who holds the infant's welfare paramount in carrying out these responsibilities.

The Tamil Nadu study asked about *mothers*. And in Tamil Nadu the description 'mothers' does pick out a class of people with the requisite characteristics to improve infant health via nutritional education. But not in Bangladesh. If we wanted more generalizable results, we were asking the wrong question.

Let me begin now to tie matters together. I have been criticizing three mistaken, mutually supporting philosophical stances that contribute to bad policy decisions.

1. We take causal relations to be linear and god-given and do not tend sufficiently to the underlying systems that make these relations possible.
2. We do not trust the kind of open, mullti-method, theory-infested scientific investigations it takes to uncover the structures of these underlying systems but prefer 'objective' policeable methods, like randomized controlled trials, that deliver only surface relations.
3. We bank on certainty.

My mother maintained that a little learning is a dangerous thing. That's what we see here. It has come even to be encoded in our language, in a familiar expression now used everywhere from newspapers reports to private conversations to policy deliberations: 'Studies show...'

But studies show very little. No matter how good a study is, it can only deliver facts about the very population studied. To go beyond the boundaries of the population enrolled in the study it takes theory, and to have warrant for doing so, it takes good, well-supported theory. There's no getting round that. Sometimes we seem to act as if we believe in induction by simple enumeration: Swan 1 is white, swan 2 is white…, so all swans are white. Study population 1 does x, study population 2 does x...., so all populations do x. But with the additional drawback that we would usually be generalizing from a very small inductive base indeed, not tens of thousands of white British swans but 1 or 2 studies, or in the best of cases, a handful. And we all know what happens as soon as we move from Britain to Australia.

It is as if we have forgotten the lessons about simple induction that have been rehearsed generation after generation for eons. Recall Bertrand Russell's chicken. She infers, on very good basis, that when the farmer comes in the morning, he feeds her. That inference serves her well till Christmas morning when he chops off her head for Christmas dinner. Of course the chicken did not base her inference on a randomized controlled trial. But had we conducted one we would have obtained exactly the same results. Her problem was not her study design but rather that she was studying surface relations. She did not understand the underlying socio-economic system that gave rise to the causal relations she observed. We often act as if the methods of investigation that served the chicken so badly will do perfectly well for us.

For purposes of thinking about policy prediction and evaluation it is important to distinguish three distinct kinds of causal claims:

[1] It works somewhere.

> This is the kind of claim we can clinch with objective methods like RCTs; it is the kind of verdict that a good post hoc evaluation can deliver.

[2] It works.

> I take it this very popular expression must mean 'It works almost everywhere', or at least 'widely', or perhaps 'everywhere, other things being equal'.

[3] It will work here.

> This is what we want to know when we deliberate about whether to adopt a policy.

[1] is a good long way off from [2] or [3]; and in a very great many cases *here* is not relevant at all to *there*, either because here lacks the requisite support factors for the policy to work or because here and there have different underlying systems at work.

You may think my distinctions are obvious and ones we are not likely to lose sight of in policy deliberations. But not so. Let me illustrate with an example from a paper by Esther Duflo and Michael Kremer, who are part of a team of brilliant and dedicated development economists at the heart and core of JPAL, the MIT-centred Jamil Poverty Action Lab, and avid advocates of RCTs. Already in line 5 in one single sentence all three kinds of claims are mixed together.

> The benefits of knowing which programs work…extend far beyond any program or agency, and credible impact evaluations… can offer reliable guidance to international organizations, governments, donors, and…NGO's beyond national borders.[1]

I take it from the language and use that they mean:

1. Which programs work = it works ('in general')
2. Impact evaluation = it works somewhere
3. Reliable guidance = it will work for us.

Here Duflo and Kremer slide seamlessly between 'It works there', 'It works' and 'It will work here'. That's loose talk, and loose talk in proper academic settings where we are meant to subscribe to high scientific standards.

Exactly the same kind of slide occurs throughout evidence-based policy. To get policies that work here, we are urged to use policies 'that work'. That is okay, supposing 'It works' means 'It works widely'. If a policy does work widely, in particular widely enough to cover here, then trivially it works here. The trouble is that the standards for 'It works' are not those for establishing a solid general claim, and one using the concepts that such claims require. The standards are generally just that the policy, as described in the study protocol, has been shown to work in some study, or for highly-ranked policies, in a handful of studies.

[1] [Use of Randomization in the Evaluation of Development Effectiveness p 93]

The injunction to use ‘What works’ is endemic, from health to crime and justice systems, to child welfare, to education, to development economics. And the recommended standards for good evidence for what works are essentially the same across all these areas. And they all share the same deep problem: they are fine standards for justifying that a policy works somewhere but not for showing ‘It works’ let alone ‘It will work here’.

Before my very brief conclusion I should like to recall where we started. I began with some policy prescriptions that may be all right but that certainly call for moral scrutiny. I worry that these prescriptions get false support from a set of dodgy views about certainty, objectivity and causality. At the close we have seen that these in turn are supported by the illicit conflation of distinct kinds of causal claims.

To say, as Duflo does – and many others as well – that either randomized controlled trials or randomized field experiments ‘identify’ programs that ‘can alleviate poverty’ is loose talk, conflating different kinds of causal claims and giving a false sense of certainty and a false pride that our predictions are objective. The same thing happens in child welfare, in crime prevention, education, drug abuse,… We cannot afford to indulge in loose talk. We need to be clear exactly what it takes to provide reasonable confidence about our policy predictions. Otherwise we bank too much on the result; we can let our optimism about policy outcomes outweigh our moral concerns; and disastrous consequences can ensue.

I have then just one sentence in conclusion. Now as then we must take seriously the World War II warning, ‘Loose talk can cost lives.’