Uses and misuses of evaluation in social policy

Christopher Deeming

We can achieve a sort of control under which the controlled, though they are following a code much more scrupulously than was ever the case under the old system, nevertheless feel free. They are doing what they want to do, not what they are forced to do. That’s the source of the tremendous power of positive reinforcement – there’s no restraint and no revolt. By careful cultural design, we control not the final behaviour, but the inclination to behave – the motives, desires, the wishes.


INTRODUCTION

A basic question within the social sciences, which is rarely addressed directly or well, is to ask whether the ends of social welfare can ever justify their means. This chapter sheds new light on this issue by examining the relationship between evidence and evaluation in social policy in both the Global North and Global South, as policy-makers seek to address social issues in the design and implementation of new social policies that actively govern conduct. Behavioural regulation is the order of the day. For scholars interested in the development of social policy and the idea of a society as a whole, it is timely to begin the re-evaluation of the very notion of active social policy and society beyond the behavioural policy paradigm. Here we are particularly concerned with the ends and means of the coercive, regulatory policy instruments seeking to control behaviour and the active ethical issues arising from their ‘use’ – and ‘misuse’.

When the behavioural psychologist B.F. Skinner, first published Walden Two in 1948 it was considered a utopian novel of ‘science fiction’, since science-based ‘behavioural engineering’ methods and experiments designed to alter people’s behaviour in society en masse were virtually unknown, as B.F. Skinner later recalled (1976, p. vi). The community – called ‘Walden Two’ – had a system of governance based on behavioural experimental design. It operated by continually testing the most successful, evidence-based strategies in an effort to organize the community for the better. The book proved controversial because its characters speak of a rejection of ‘free will’, albeit for the ‘common good’, as it was supposed.

Today, many observers have commented on the spread and growing use of behavioural experiments in social policy around the globe to solve perceived social problems based upon an individualistic approach to the
This new and emerging form of ‘active social policy’ (Bonoli, 2013) often involves conditioning the receipt of welfare – in the form of cash transfers, goods, or services – according to specific individual behaviours (Jones et al., 2013). The different forms that incentive-based approaches and behaviour-changing policies can take mean that the new and emerging terrain of active social policy is functionally diverse, if not conceptually ambiguous. As I have suggested elsewhere (Deeming, 2016a), the active social policy paradigm may be said to include the interventions associated with ‘nudge’ (Thaler and Sunstein, 2009) and ‘libertarian paternalism’ (Thaler and Sunstein, 2003), and behavioural economics more generally – in the advanced (Oliver, 2013) and developing economies (Banerjee and Duflo, 2011), but it also covers the vast range of active labour-market policies (ALMPs) and welfare conditionality programmes (Bonoli, 2010). Few, however, have touched on the contextual differences in terms of the application of research and evidence governing policy-making in both developed and developing world contexts. Subsequently, the debates over the ‘use’ and ‘misuse’ of evaluation in social policy in the Global North and Global South are not as well connected as they might be.

This chapter is organized as follows: the first section below looks at how conditional cash transfer (CCT) programmes have reshaped evidence-based policy for developing countries; the second section considers how active social policy designs have transformed welfare systems in the advanced economies; the third section examines how ‘evaluations’ are conducted and the ways in which ‘evidence’ is then represented by policy-makers in the policy processes, including what this may tell us about the nature of power and rule in society, and the complex ethical and human rights issues arising from ‘active’ (coercive) social policy programmes. Conclusions are drawn together in the fourth section.

**EVALUATION IN DEVELOPING COUNTRIES**

As a result of the global research effort (WHO, 2011), it is now well established that there are certain (desirable) human behaviours and conditions of living that are beneficial for our health and well-being. This knowledge should be used to form the basis of social policy, as Hartley Dean (2010) argues. In the brave new world of the behavioural economist, promoting health, wealth and happiness (and meeting policy targets such as the Millennium Development Goals – MDGs) is largely about changing people’s behaviour, and being able to observe and assess that change. Demonstrating the effectiveness or impact of an intervention is thus the keystone for policy development (Deaton, 2010). The point about clearly demonstrating success and impact in social policy processes also underscores the need for robust evaluation. In order to consider the effectiveness of a particular policy or programme, one really needs to know what outcomes would have been achieved had the programme not been in place. This is often referred to as the
‘counterfactual outcome’ (Morgan, 2014). One way of overcoming this evaluation problem is through the use of randomized controlled trials (RCTs), considered to be the ‘gold standard’ of all the methods available to researchers (Young et al., 2002). Generally, in a randomised controlled trial, study participants are randomly assigned to one of two groups: the experimental group receiving the ‘intervention’ and a comparison or ‘control’ group (Sibbald and Roland, 1998); researchers are then able to compare the outcomes between groups (see Chapter 2).

Increasingly, behavioural economists are advocating social experiments and the RCT in particular, as the main tool for studying the effectiveness of social policy interventions in development settings. In developing economies, social experiments with conditionality have long been part of the drive for evidence-based policy (Fiszbein et al., 2009). CCT programmes usually have an a priori evaluation design built into their operation, which embraces experimental or quasi-experimental features. RCT designs are common. In a recent trial in Malawi, for example, the behavioural condition targeted children at risk of not going to school without the benefit of a CCT. The behavioural condition placed upon transfer recipients was thus regular attendance of children in school (Baird et al., 2009a, 2009b, 2015). Some villages were randomly assigned to the social assistance programme and others were not (Figure 9.1). Important, at the outset each village had the same equal chance of being picked to receive the intervention; families were not told that they were part of an experiment – they were ‘blind’ or ‘blinded’ as knowing this information may have affected their behaviour and biased the outcome of the study. The development economists – sometimes referred to as the ‘randomistas’ – argue that this type of randomized experiment is the only way to be sure of identifying impact, because it helps to eliminate bias and other confounding (hidden) factors.
A systematic review of RCT results can usually be found at the very pinnacle of the research evidence hierarchy of (see Chapter XXX), as it provides a way of pooling evidence from different studies to provide an overview of outcomes (White and Waddington, 2012).2 Other non-experimental methods found at the bottom of the evidence hierarchy are largely dismissed; they are considered ‘unscientific’ and are best avoided (see Chapter XXX). The influence of the randomistas appears to be growing: NGOs (World Bank, IMF, WTO) and philanthropic agencies, and donors are increasingly giving (explicit) preference to randomized designs and systematic reviews in evaluating policy programmes and their impacts (Hickey et al., 2009; Deaton, 2010; Hagen-Zanker et al., 2012; Hall, 2015).

According to some behavioural economists like Abhijit V. Banerjee and Esther Duflo (2011), the major debates in international development can be boiled down to disagreements about the shape of a function in development theory (Figure 9.2). The S-shaped curve on the left suggests poor people are ‘trapped’ in poverty and require a ‘conditional push’ to help them out of poverty. The L-shaped curve on the other hand, suggests that poor people are gradually able to pull themselves out of poverty because they are not really ‘trapped’. If we accept this premise, the theoretical proposition and corresponding debate about which of the two graphs best represents the real world can only be settled experimentally, that is, through the use of RCTs. Of course, not everyone agrees with the worldview depicted in these two graphs or feels the need for social experiments to solve global problems like poverty. The ‘randomistas’, however, firmly believe that it is perfectly possible to make significant progress in tackling global social problems using

Figure 1: Malawi RCT design

experimentation, through the accumulation of small experimental steps, each well thought out, carefully tested, and judiciously implemented. In effect, they claim to be saving lives with a well-placed behavioural ‘nudge’ and they offer plenty of evidence to support their position.

**Figure 9.2 Different prospects for the world’s poorest**

![Graphs showing different prospects for the world’s poorest](image)

*Source: Adapted from Banerjee and Duflo (2011, Figures 1 and 2, pp. 12–13).*

Conditional social policy programmes are now firmly established across the developing world and shape the lives of millions of people. Many CCT programmes are large in scale and usually have an evaluation design built into their operation. In 2014, Oportunidades in Mexico, for example, cover about a quarter of Mexico’s national population (1.5 million households). In Brazil, some 11 million families – 46 million people – receive regular transfers under the Bolsa Familia programme. Mexico’s Progrresa (renamed Oportunidades in 2002 and now rebranded as Prospera), often seen as the seminal and model programme, began in 1997. This programme has demonstrated a range of benefits over the years, particularly in the areas of health and education (DFID, 2006). For instance, some 70 per cent of families participating in the scheme have shown improved nutritional status and stunting has been reduced. Antenatal care increased by 8 per cent, contributing to a 25 per cent drop in the incidence of illness in newborns. Immunization rates improved as a result of preventive healthcare appointments, while school enrolment rates increased by over 20 per cent for girls and by 10 per cent for boys. These social assistance programmes have fiscal appeal too, at least according to the national governments and NGOs such as the World Bank. In terms of national budget, Mexico and Brazil commit only 0.5 per cent of GDP to their programmes. Guatemala is one of the latest Latin American countries to embark on reform with Mi
Fa\textit{mi}lia Progresa (My Family Progresses), introduced in 2008 (Gaia, 2010). As well as familiar conditions relating to school attendance and the use of health services, other conditions continue to be tested, for example, adult education, microcredit, housing and accommodation schemes, also involving ‘bed net’ schemes to help protect people against malaria-carrying mosquitoes (DFID, 2011).

The evidence from CTT evaluations suggests that these programmes can be particularly effective in promoting the health and well-being of disadvantaged populations. Improvements in the use of health services, family food and diet, and school enrolment rates, all show significant gains and improvements with conditionality in place. Table 9.1 summarizes recent findings from studies published in 2014/15.

\textit{Table 9.1 Selected CCT impact evaluations of social safety-net programmes published in 2014/15}

<table>
<thead>
<tr>
<th>Country</th>
<th>Main Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nicaragua</td>
<td>Boys exposed to the CCT education programme \textit{in utero} and during the first two years of life have better cognitive outcomes when they are ten years old than those exposed in their second year of life or later. For boys aged 9–12 in 2000 (and thus aged 19–22 in 2010), the short-term programme effect of a half-grade increase in schooling was sustained into early adulthood, seven years after the end of the programme. In addition, there were significant and substantial gains in both maths and language achievement scores. Those boys of the same cohort in the early treatment group have higher earnings in the labour market than those in the late treatment group</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>This CCT health pilot had a significant impact on the incidence of wasting (low weight for height) among children who were 10–22 months old when the programme started. The pilot was also able to improve nutrition knowledge among mothers, including an increase in the proportion of beneficiary mothers who knew about the importance of exclusively breastfeeding infants until the age of six months</td>
</tr>
</tbody>
</table>
Philippines

Pantawid Pamilya CCT (Bridging Programme for the Filipino Family) encourages the trial use of modern family planning methods. The programme promotes facility-based deliveries and access to professional postnatal care and improves children’s access to some key health care services. Among Pantawid beneficiaries, the PhilHealth health insurance programme covers about nine in ten households. The programme keeps older children in school. Children (10–14 years old) in the programme work seven fewer days a month than children not in the programme. Pantawid increases households’ investments in education and does not encourage dependency or spending more on vice goods, such as alcohol. Pantawid Pamilya CCT is reaching most of its key objectives. The impacts found through this study are comparable to the levels of impact found in other CCT programmes around the world at this stage of programme maturity, particularly in terms of the programme’s achievements in improved use of health services and school enrolment.

Tanzania

In this community-based CCT programme significant impacts are observed across a broad array of areas, including health, education, and various risk-reducing behaviours: use of health insurance, insurance expenditures, non-bank savings (for the poorest households), and the purchase of livestock such as goats and chickens. In addition, the programme has led to significant increases in spending on certain children’s goods (especially children’s shoes).


CCT programmes appear to have secured significant reductions in levels of extreme poverty (Table 9.2). In Mexico for instance, severe poverty as measured by the ‘squared poverty gap’, which attempts to take account of the depth of poverty and inequality among poor people, has fallen by about 29 per cent – a significant achievement. Nevertheless, policy-makers are continually looking to refine and improve programme effectiveness (Gaarder et al., 2010; Yoong et al., 2012). Despite the huge investments in large-scale trials, however, important questions remain unanswered, particularly concerning cost-effectiveness and cost–benefit, an issue that is discussed in more detail below. Before this
however, we turn to consider social policy evaluation in the advanced economies, where concerns about the paucity of evidence and robust evaluation continue to be raised.

**Table 9.2 Poverty measures for number of people (in millions) below $1.25 a day in 2005 PPP**

<table>
<thead>
<tr>
<th></th>
<th>Headcount</th>
<th>Poverty Gap&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Squared Poverty Gap&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-transfer</td>
<td>Post-transfer</td>
<td>Pre-transfer</td>
</tr>
<tr>
<td>Brazil</td>
<td>0.2421</td>
<td>0.2369</td>
<td>0.0980</td>
</tr>
<tr>
<td>Ecuador</td>
<td>0.2439</td>
<td>0.2242</td>
<td>0.0703</td>
</tr>
<tr>
<td>Jamaica</td>
<td>0.2439</td>
<td>0.2329</td>
<td>0.0659</td>
</tr>
<tr>
<td>Mexico</td>
<td>0.2406</td>
<td>0.2222</td>
<td>0.0847</td>
</tr>
</tbody>
</table>

**Notes:**

<sup>a</sup> Depth of poverty: this provides information regarding how far households are from the poverty line.

<sup>b</sup> Poverty severity: this takes into account not only the distance separating the poor from the poverty line (the poverty gap), but also the inequality among the poor. Thus, a higher weight is placed on those households who are further away from the poverty line.


**EVALUATION IN THE ADVANCED ECONOMIES**

The origins of active social policy in the advanced economies can be traced back to the post-war period, particularly with active labour-market policy in Sweden (social democratic variant) and the ‘workfare’ or ‘work-for-your-welfare’ neoliberal variant in the USA. The 1990s witnessed the full expression of this new and emerging form of thinking, as major welfare reforms were undertaken across the advanced economies (OECD, 1994, 2006; Nativel, 2006). Compared with developing country contexts, however, there has been less direct appeal to research evidence gained from robust evaluations in order to secure welfare reform: experimentation and evaluation with RCTs has been less of a priority in the advanced economies (OECD, 2013). Consequently, welfare states were reformed on more ideological grounds, with an appeal to free-market political philosophy as Lawrence Mead and Christopher Beem (2005) observe. Putting this in
stark terms, welfare conditionality in the Global South is, arguably, being driven by an evidence-based policy-making agenda whereas in the Global North, political philosophy is clearly driving welfare reform.

Labour-market attachment programmes introduced new duties (roles) and notions of reciprocal obligation (responsibilities) (Lødemel and Moreira, 2014). Work conditionality has been intensified and now applies to most sections of the adult working age population in the advanced world. Currently, the primary obligation on social security recipients in most welfare states is to actively look for paid employment and to improve their chances. Claimants must undergo training to develop new skills and they must engage with the employment services on offer and behave in a certain way to access allowances and benefits. The original research that underpinned ‘welfare to work’ in the UK, for example, was encouraging, but certainly not convincing (Vincent et al., 1998).\(^3\) Active social policy, however, was designed to solve a problem about the effects of passive, rights and eligibilities–based welfare that arose, it was argued, because too many citizens had been reduced to ‘welfare dependency’. Supporters of active social policy have long maintained that the new duties imposed on welfare benefit claimants are perfectly justifiable as the social contribution now owed by citizens to society (Mead, 1986). There is thus a strong ideological basis to labour-market activation, particularly the ‘workfare’ variant, that sits squarely within the free market ‘neoliberalism’ according to some critics like Mitchell Dean (1995).\(^4\) However, the research evidence supporting this great transformation of social policy has been the subject of much debate.

Observers complain that research evidence has not been driving welfare reform in the advanced economies in the way that it should (Haynes et al., 2012; Roberts et al., 2012). Aside from health, and possibly education (Oakley, 2000), there appears a real lack of substantive investment in robust policy trials and systematic reviews. Until recently, evaluations of ALMPs have been rather limited in scope and design, particularly when considered against the large-scale trials in developing economies. Research is only now attempting to piece together the fragmentary evidence relating to ALMP evaluations, including the use of formal systematic review techniques (Bambra et al., 2005; Kluve, 2010; Escudero, 2015; Filges et al., 2015). At present there is huge uncertainty surrounding individual and aggregate programme effects. How ALMPs help or benefit particular constituencies of the population is not generally known, while cross-national comparisons are often compounded by institutional variations in programme design, meaning it is often difficult to generalize from one context to another (Clasen et al., 2016).\(^5\) Targeted job subsidies may, for instance, lead to additional job creation, or they might equally crowd out unsubsidized work through substitution effects. The evidence from trials to date suggests modest returns on certain activation programmes from untold costs, notably those that include job search activities accompanied by sanctions for non-compliance.
ARGUMENTS CONCERNING THE ‘USE’ AND ‘MISUSE’ OF EVALUATION IN SOCIAL POLICY

The review, albeit brief, elaborated under the rubric of ‘active’ social policy highlights a number of competing tensions and complex interrelated issues that require further elaboration. In this section we are particularly interested in how ‘evaluations’ are conducted and the ways in which ‘evidence’ is represented by policy-makers in the policy processes and what this then says about the nature of power and rule in society, and the complex ethical and human rights issues arising from the coercive social programmes.

Evaluations and Narratives about ‘the Evidence’

‘Evidence’ in one form or another is indispensable in the policy-making process, as citizens in democratic societies value coherent explanations for policy choices. While implementing welfare reforms with little or no clear evidence of their effectiveness could be considered unethical (Davies et al., 2000), it is also clear that the whole ‘evidence-based’ policy movement has become too ideologically driven (Packwood, 2002; Stevens, 2011; Wiggan, 2012). Active social policy designs lean towards neoliberal belief and value systems, as Mitchell Dean (2014) argues, determining how people and society should now function in the twenty-first century (Deeming, 2016b). For example, ‘welfare quarantining’, as it is known, in Australia’s ‘income management’ programme addresses the inappropriate use of family income – represented by policy-makers as the ‘problem’. Thus, a range of welfare benefits (including unemployment, disability and single-parenting payments) are granted only on condition that recipients do not spend any money on goods and services such as alcohol, tobacco or gambling products. Australian policy-makers and their critics make strong claims and appeals to the evidence about ‘what works’ (Nutley et al., 2009), and whether the ‘ends’ of the programme justify their ‘means’. Of course, such positioning may be expected in political life more generally, as Peter Hall (1993) observed, with the state involved in ‘social learning’ processes and policy-makers seeking to champion political ideas. At best, however, the programme evaluations from Australia suggest a diverse set of impacts (Deeming, 2016a). Faced with a deluge of inconclusive information, however, policy-makers have been moved to create persuasive policy stories and ‘evidence-based’ narratives to help justify their interventions and social programmes, as Eva Cox (2011) argues. In doing so, they have ignored other options and arguably more ‘enlightened’ ways of helping people to ‘get on’ and learn to manage their own affairs (Rowson, 2011).

Even with all of the CCT experimentation happening in the developing world, we still know little about the strengths and weaknesses of alternative programmes – cost-benefit calculations in this field of social policy are rare
(St. Clair, 2009). This makes it extremely difficult, at least at present, to make clearly informed judgements about the value for money of interventions, or the value and benefits of CCTs verses unconditional cash transfers (UCTs) (Barrientos, 2009). The impact of conditionality is usually measured at the margins, by incremental gains. For example, secondary school enrolment rates for girls in Mexico’s Oportunidades programme increased by 9 percentage points in two years, from a base of 67 per cent. Whether such modest gains are worth the costs or justify conditioning the right to welfare is another matter entirely. Ideally, the overall (system) costs and benefits, to individuals, families and society should be reflected in any programme evaluation. However, capturing the full range of information required is methodologically and technologically challenging, which itself comes at a significant cost. Conditionalities may, for example, impose non-trivial compliance costs on families not accounted for in their benefit levels (i.e., the transfer may not cover the cost of the condition). Including cost–benefit analysis in future CCT evaluations might help to inform the current debate; there is, however, growing evidence that UCTs are cheaper, more efficient to administer, support human rights (unconditionally) and demonstrably improve family well-being (Davala et al., 2015). So why, we might ask, is it necessary for policy-makers to attach conditions to transfers?

In the advanced economies, we also find cost–benefit calculations are largely absent. Policy-makers in Australia, for example, have struggled to weigh up the ‘costs’ and ‘benefits’ of activation. Despite the Work Programme having been in place for nearly two decades, little is known for certain about its costs or effectiveness (Belchamber, 2013). Activation schemes in the UK have offered a poor return to the British taxpayer according the UK’s National Audit Office (NAO). Astonishingly, getting someone into paid employment can cost the public purse as much £200,000, it claims (NAO, 2012a, 2012b). Subsequently, scandals have made headlines in the British press amid allegations of fraud, excessive pay and bonuses for those agencies charged with helping unemployed people back into the labour market (Committee of Public Accounts, 2012). Of course, politicians and policy-makers are often keen to promote their policies and reforms. Thus, apparent and favourable impacts – improvements in employment rates or reductions in state expenditures for instance – are seized upon to build public confidence and support. However, such information is usually uncontrolled knowledge. In other words, more often than not we know very little about the counterfactual outcome that would have been achieved had the social programme not been in place (Morgan, 2014).

Behavioural effects in the field of labour market activation are notoriously complex, and difficult to observe and monitor, particularly for long periods of time (Paz-Fuchs, 2008). Often there are indirect social and economic effects, unobserved consequences and population effects. Much of the evidence that has been generated seems patchy at best, meaning it is often inconclusive on key questions. Social experiments in the field of social policy
are often complex in design (Samson et al., 2010), and difficult to administer (Lomeli, 2008). Even well-designed trials present methodological challenges, and some of the observed changes in behaviour during the course of a study may be related, at least in part, to the special conditions and monitoring systems in place during the period of investigation (known as the ‘Hawthorne effect’). Researchers have recently detected a possible ‘Hawthorne effect’ in some of the trial data from Latin America, for example (De Brauw and Hoddinott, 2011).

The evidence from developing country contexts, even for improvements on core programme objectives such as school attendance, is often unclear and context dependent, making it hard to generalize from one experiment to the next (Olk et al., 2012). Evidence is also weak or unproven in key areas of development, particularly for policies addressing child poverty (Yeates, 2009) and child labour (Tabatabai, 2009). Doubts are also expressed about the longer-term sustainability of programme benefits. Impacts tend to be evaluated in the short term and as a consequence, little is known about longer-term effects. Observed benefits may be diluted over time, for example. Evaluations tend to show the frailty of poverty reduction programmes (Lomeli, 2009) and jobs created by labour market programmes may not last (Kluve, 2010). Structural issues extend far beyond individual agency and personal responsibility, as Sharon Wright (2012) observes. In times of high unemployment, for instance, there simply may not be enough jobs to go around (Dean, 2012), or high childcare costs in some countries may deter sole parents from working (Davey and Hirsch, 2011).

With the turn to activation policies that problematize individual agency (with a particular focus on the lower classes and disadvantaged sections of society), however, policy-makers appear to have lost sight of important structural and economic constraints and social inequalities. In the process, policy-makers are failing to actively engage the public on important issues, policy options and priorities for social investment and inclusive growth to address problems in the labour market and society more generally (Deeming and Smyth, 2017). Some groups face higher unemployment rates than others. Policy-makers have clearly misunderstood the complex structural causes of youth unemployment, which still remains at record levels across the OECD (2015), but may not care much if the object is to extend state compulsion on citizens regardless of their circumstances.

Policy-makers tend to assume that ‘unemployment’ (the problem) is voluntary, resulting from poor motivation. Equally, however, the ‘problem’ of unemployment could be reconceptualized as the lack of ‘decent work’, which is not just about ‘paid employment’ in society, as Hartley Dean (2014) suggests. One of the main objections to the state stepping in to create work, as the employer of last resort, is whether it is economically sustainable to do so. The political right’s view, traditionally, is that only jobs created by the private sector can be considered real productive jobs in the advanced economies; the situation is quite different in some of the coordinated economies of Europe where job-
creation strategies have long found favour in ALMPs and firms have been more enthusiastic about implementing such programmes (Martin, 2004). Under coercive workfare programmes, however, the state is increasingly prepared to create and subsidize poor-quality work and unrewarding roles in order to discipline those without jobs (Wacquant, 2010). Thus, punitive policies are designed not only to deter citizens from making welfare claims but also to act as a regulatory labour-market push factor, forcing low-skilled workers with little choice but to accept low-waged jobs.

In developing contexts, there are more fundamental, outright objections to the CCT schemes supported by international agencies. While CCT programmes are shown to have secured significant reductions in levels of extreme poverty (seen in Table 9.2), many argue that the new behavioural policies that violate human rights are failing to make a real impact on the rate of chronic global poverty (Reddy and Pogge, 2010). Extreme world poverty has fallen, down from 1.94 billion people in 1981 to 835 million people in 2015, or around 10 per cent of the global population (Table 9.3). However, nearly all of the poverty reduction over the period took place in China. Progress elsewhere, over nearly three decades, has been stubbornly slow, casting a long shadow of doubt over the individualistic explanation of the problem (Deeming and Gubhaju, 2015). By contrast, a broad coalition of bodies under the UN recommends all countries establish social protection floors (SPFs) that guarantee income security (Deacon, 2013). A floor under the global economy would certainly help to improve the lives of millions of people around the world.

Table 9.3 Global poverty estimates, the number of people (in millions)

<table>
<thead>
<tr>
<th>Region</th>
<th>Millions of People Below US$1.25 a Day</th>
<th>Projections</th>
</tr>
</thead>
<tbody>
<tr>
<td>East Asia and Pacific</td>
<td>957.1</td>
<td>324.1</td>
</tr>
<tr>
<td>Eastern Europe and Central Asia</td>
<td>7.1</td>
<td>6.0</td>
</tr>
<tr>
<td>Latin America and the Caribbean</td>
<td>52.7</td>
<td>42.0</td>
</tr>
<tr>
<td>Middle East and North Africa</td>
<td>13.1</td>
<td>9.0</td>
</tr>
<tr>
<td>South Asia</td>
<td>603.2</td>
<td>589.0</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>287.1</td>
<td>399.1</td>
</tr>
<tr>
<td>Total</td>
<td>1920.2</td>
<td>1368.1</td>
</tr>
</tbody>
</table>

Source: World Bank Group (2015, p. 19), number of people living below $1.25 a day, in 2005 purchasing power parity, PPP.
Active social policy appeals to the positivist research tradition. Within behavioural economics, for example, conditionality in welfare is often presented as a ‘model’ to be tested. However, work by Paul Stubbs (2009) reveals how the behavioural change programmes are socially constructed. Drawing on CCT experiences in Eastern Europe, Stubbs observes how policy-makers negotiate the aims and objectives of their programme, adapting trial evidence from one context (e.g., Latin America) to another (e.g., Europe). Some of the historical, cultural and institutional connections drawn here are tenuous at best he argues, and families may suffer as a result of the policy translation process as welfare is reformed. RCTs are not necessarily the ‘gold standard’ of causal inference in social policy as trial data always have to be interpreted (Cartwright, 2007). Policy conclusions cannot simply be exported from one population or context to another.

**Ethics and Human Rights**

RCTs as a method for evaluating social programmes can raise profound ethical issues. Critics, for example, argue that it is unfair to experiment with the lives of the poor and disadvantaged (Elizabeth and Larner, 2009). Experiments with social security and assistance – by their very nature – often target the most vulnerable sections of society. Therefore, critics claim that conditionality in social welfare undermines social inclusion or, worse still, reinforces existing prejudices in the community. Unemployed people are often stigmatized in the British media for instance, as Malcolm Dean (2011) observes. For this reason, the Roman Catholic Church in Australia principally opposes ‘welfare quarantining’ that targets disadvantaged families. Since human behaviour is being ‘policed’ the programme involves elaborate systems for monitoring compliance and controlling behaviour. Ethically the scheme has been the subject of much controversy (Deeming, 2016a).

State surveillance systems are often required to monitor compliance. Therefore, the unwarranted infringement on privacy has to be considered against the public interest according to communitarian thinking (Etzioni, 2014). But whose interest is being served here and at what cost? In 2011, for example, the Australian government allocated some $117 million from the federal budget simply to administer the state surveillance systems necessary to monitor the behaviour of citizens in the ‘welfare quarantining’ programme. The government was forced to introduce a new type of EFTPOS® debit card in order to monitor all forms of consumption. This new debit card could be used only at approved and participating shops and outlets. Thus, it was necessary for policy-makers to enlist retailers across the country in government-led practices, all deemed necessary to meet programme objectives. Clearly, the means and rationalities of active social policy programmes not only incorporate the institutions, state bodies, agencies and
programme administrators (necessary for the governance of active social policy programmes; cf. Heidenreich and Graziano, 2014), they also include a wide network of experts, professionals and practitioners drawn from diverse fields of health, welfare and education and beyond, who play a key role in fulfilling the political objectives of social policy. In this context, however, it is too soon to know whether the benefits of the policy outweigh its costs but the ethical implications are profound.

Critics argue that human rights are being eroded by the active social policy programmes (e.g., Standing, 2014). The issue at stake here is less about the ‘evidence base’. Conditionality can be effective in changing behaviour in certain circumstances, as we saw above in Table 9.1, but the problem, according to some observers, is that conditionality undermines basic human rights, citizenship rights and important principles governing social solidarity. CCTs are thus divisive in that they create and sustain social divisions within society. Poor people are treated differently from their fellow citizens; basic human rights are being violated. Several international instruments affirm that everyone has the right to social security. For example, Article 22 of the 1948 UN Universal Declaration on Human Rights (reconfirmed in 1993 and 1996) states that ‘everyone, as a member of society, has the right to social security’ and Article 25 states that ‘everyone has the right to a standard of living adequate for the health and well-being’. If basic social security is only transferred on condition, then basic human rights are surely undermined. There are strong counterarguments to this position, however, whereby exponents of CCTs claim that conditionality does in fact confront the denial of basic human rights by ensuring that welfare reaches those families most in need (see Lomeli, 2008). CCTs can also be seen as a form of ‘social investment’ for the recipient and their family given the explicit links to human capital development in education and health (Nelson and Sandberg, 2016).

**Governance, Power and Rule**

Fairly constrained debates about the efficacy (i.e., ‘effectiveness’) of social policies discussed so far ignore much more basic concerns regarding power structures and the nature of the ongoing struggles between dominant elites and marginalized populations. For Desmond King (1999), the coercive CCT programmes represent the triumph of ‘illiberalism’ in social policy. Mitchell Dean (2002, 2007) offers ‘authoritarian liberalism’ as another turn of phrase to capture the dark side of these political programmes. It is not hard to see why. Political elites (Jones, 2014) and international financial organizations (Hall, 2015) now control global policy processes and wield power. A form of ‘tyranny’ then overrules the desires and aspirations of most people in society, according to John Dryzek (1989). For example, while politicians seem to have persuaded the British public that further obligations in welfare services and a
general championing of duties over rights are necessary (Deacon, 2005; Dwyer, 2008; Deeming, 2015), many families who find themselves ‘on condition’ object (Rafferty and Wiggan, 2011). In Australia, the federal government even went so far as to claim that ideas about ‘welfare quarantining’ actually arose out of some of their own community consultations. Independent investigations have subsequently cast doubt on this claim, however (Deeming, 2016a).

Critics point out that conditionality in welfare policy reinforces entrenched inequalities within society, and thus preserves existing social structures and power relations (e.g., McDonald and Marston, 2005). Policy-makers now believe that vulnerable people must be threatened with penalties and sanctions (and less money ultimately) to incentivize them to behave in a particular way. Such policy practices then constitute poor people as subjects engaged in a power struggle for resources, recognition and respect (Veit-Wilson, 2009). Subsequently, we need to be much more critical about how ‘social problems’ are being defined in global social policy, as Carol Bacchi (2009) argues, and the ways in which research ‘evidence’ is represented for policy-making. We also need to think more critically about both the production and use of ‘evidence’, particularly when policy-makers use it to help legitimize policies in the policy-making process. Guy Standing (2008), for example, sees the CCT initiative in developing countries as a political device to legitimize cash transfers with middle-class voters and international agencies. Peter Taylor-Gooby (2011) makes similar claims for the UK, after decades of political attacks directed at the British welfare state.

SUMMARY

So what lessons might we draw from this review? Stated simply, the intent here is to make us think more deeply than usual about the underlying assumptions and meanings of policies; to encourage us to think much more critically about how policy-makers have sought to address ‘social problems’ in the early part of the twenty-first century with their policy designs; and the ideological frameworks and assumptions that are invoked and brought into play during the evaluation process. As John Clarke (2004) argues, there are always alternative ‘regimes of truth’ in the public realm.

The key message then is to not allow our gaze to simply gloss over the assumptions underpinning social policy evaluations themselves, but to maintain a critical distance and allow for more critical questioning and reflection, not only of how ‘social problems’ are represented and constructed by policy-makers, but also how aims and objectives and findings are constructed and framed in the messy evaluation process itself. Not only do we need to be aware of the ends of policy, and ask whether they justify the means, but we also need to question whether the uses or acts of evaluation themselves actually serve to legitimize particular policies and existing power relations and inequalities within society. Ultimately, then, the ‘use’ (or ‘misuse’) of evidence and evaluation in social policy lies in the eye of the
beholder. More often than not evaluations and findings only serve to stir and provoke our values and principles still further, concerning the kind of society we wish to live in – just as B.F. Skinner did all those years ago with his utopian/dystopian novel (Altus and Morris, 2009).\(^7\)

**NOTES**

1. According to ‘libertarian paternalist’ thinking it is quite legitimate for private and public institutions to affect behavioural change while at the same time respecting freedom of choice.

2. To be included in a systematic review or meta-analysis, a study usually has to meet strict research quality criteria relating to research design. Often this usually means a randomization or a well-designed before-and-after control study. A clear measure of at least one outcome is also required.

3. This study published in 1998 appears to be the first RCT in the field of social security in Britain.

4. ‘Neoliberal’ capitalism is said to be based on ideas about the supremacy of free markets, and the efficiency of market forces, via the privatization of public services and removing or reducing state regulations – ‘deregulation’.

5. The International Development Coordinating Group website of the Campbell Collaboration has further information on systematic reviews completed and in progress, accessed 19 October 2015 at www.campbellcollaboration.org/international_development/index.php.

6. Electronic funds transfer at point of sale.

7. *Walden Two* was and is controversial to the point of being labelled dystopian by some, criticized for dismissing purpose, mind, and freedom; while others celebrated Skinner’s utopian vision based on empirical methods to search for and discover practices that maximize people’s health, wealth and well-being.

**REFERENCES**


