

Examining the epistemology of impact and success of educational interventions using a reflective case study of university bursaries

HARRISON, Neil and MCCAIG, Colin <<http://orcid.org/0000-0003-4364-5119>>

Available from Sheffield Hallam University Research Archive (SHURA) at:

<http://shura.shu.ac.uk/15151/>

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

Published version

HARRISON, Neil and MCCAIG, Colin (2017). Examining the epistemology of impact and success of educational interventions using a reflective case study of university bursaries. *British Educational Research Journal*, 43 (2), 290-309.

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>

Examining the epistemology of impact and success of educational interventions using a reflective case study of university bursaries

Neil Harrison (University of the West of England) and Colin McCaig (Sheffield Hallam University)

Abstract

This paper engages with the continuing emphasis given to evidence-based policy and 'what works' approaches in educational research, highlighting some of the continuing epistemological challenges from a post-positivist perspective. To illustrate these, it draws on a reflective case study provided by a research project commissioned by the Office for Fair Access in England. The aim of this project was to devise an analytical framework to allow universities to explore the impact of the bursaries that they award to students from disadvantaged backgrounds. The experiences of the project team led to a collection of insights which are discussed here alongside a more general account of the limitations of experimental and quasi-experimental designs in complex social fields. The paper concludes that we lack a strong theoretical understanding of the relationships between financial and educational disadvantage prior to and during higher education, and this undermines efforts to prove that certain interventions will 'level the playing field'.

Introduction

A key feature of the last twenty years in educational research has been the rising emphasis placed on evidence-based practice. One particularly controversial manifestation of this has been what is often referred to as the 'what works' agenda, driven in tandem by policymakers seeking accessible and actionable answers around value-for-money and researchers from the positivist tradition seeking to provide them. Its approach is strongly objectivist, privileging quantitative data collected through rigorous sampling, robust analytical techniques and formal inference to subject educational activities and innovations to a sharp test of 'impact' and 'success'.

To its supporters, 'what works' might reasonably be typified as a moral obligation. There is a responsibility to learners to ensure that the best practices are being used and that limited learning opportunities are not squandered on old-fashioned thinking, folklore or unevidenced assumptions (Slavin, 2004; Gorard, 2004, 2013). To its critics, it is reductionist, simplistic, philosophically-weak and unethical (Biesta, 2007, 2010; Oancea and Pring, 2008). It is seen within a tradition of technocratic approaches to education, where deprofessionalised teachers simply need to be programmed with the 'best' practice in order to optimise learning (Webster, 2009; Menter, 2013).

Either way it has increasingly pervaded thinking over the last ten years. The UK government, for example, has routed £125 million through the Education Endowment Foundation; a charity established to commission 'what works' style evaluations around educational interventions felt to have potential to increase outcomes for disadvantaged children (Menter, 2013; Gorard, Siddiqui and See, 2015). Other research funders increasingly require bidders to focus on innovations and impact. The 'what works' philosophy has featured in keynote addresses at the British Educational Research Association in recent years (Coles, 2012; Connolly, 2015), while public intellectuals like Ben Goldacre (2013) have voiced enthusiastic support at the government's behest.

In this paper, we critique 'what works' from a post-positivist perspective using reflections from a government-commissioned project in higher education as a case study to examine some of the challenges and traps for the unwary. We will not be rehearsing the philosophical objections to the 'what works' agenda as this has been well-covered elsewhere (e.g. Hammersley, 2013). Rather than being ideologically hostile, the paper will engage with 'what works' in its own terms; the focus here is critical and pragmatic, stressing the need for a robust epistemology to underpin this style of research, which understands the questions being asked and the ability of the methodology to answer these questions. The paper will hopefully generate some insight around the wisdom of uncritically privileging 'what works' as a form of educational research.

We begin with a brief introduction to some of the key principles of 'what works' research and some important epistemological challenges, drawing on the work of both enthusiasts and sceptics. We will then go on to introduce and discuss the case study.

Impact, causality and proof

'What works' research is premised on the need for a clear, causal and impactful relationship between an intervention and an outcome. Gorard (2013, p.65) lays out what he argues are the formal requirements for this:

1. For X (a possible cause) and Y (a possible effect) to be in a causal relationship, they must be repeatedly associated. This association must be strong, clearly observable, replicable and it must be specific to X and Y.
2. For X and Y to be in a causal relationship, they must proceed in sequence. X must always precede Y (where they both appear), and the appearance of Y must be safely predictable from the appearance of X.
3. For X and Y to be in a causal relationship, it must have been demonstrated repeatedly that an intervention to change the strength and appearance of Y then also strongly and clearly changes the strength and appearance of Y.
4. For X and Y to be in a causal relationship, there must be a coherent mechanism to explain the causal link. This mechanism must be the simplest available without which the evidence cannot be explained. Put another way, if the proposed mechanism were not true then there must be no simpler or equally simple way of explaining the evidence for it.'

The skill in devising 'what works' style projects is thus in meeting these criteria while eliminating as many potential confounding factors as possible, either through holding these factors constant or through randomisation within large samples. It is easy to see why this causal clarity is attractive to policymakers.

The methodological apogee for 'what works' researchers is the randomised controlled trial (RCT). Borrowed from medicine, it establishes an *experimental group* which undergo a planned intervention or 'treatment' (a possible cause) and a *control group* that does not. Random allocation of either individuals or clusters of individuals (e.g. schools) into the two groups statistically controls for extraneous factors that might impact on the outcome, with the objective of isolating the effect of the intervention, demonstrating causality and estimating the effect size (i.e. the degree to which the intervention influences the outcome). RCTs are typically used within medicine when exploring the efficacy of potential new drugs.

'Lesser' forms of experimental design are also popular. For example, quasi-experiments are used where the researcher is not able to dictate in advance the allocation of the intervention between groups – e.g. where some schools choose to opt-in to a new programme while others do not. Analysis may use retrospective random selection from within the experimental and control groups or on a matching process to ensure that the two groups are as similar as possible. Quasi-experiments are particularly used in medicine to determine long-term susceptibilities to disease and where purposive allocation to groups would be impossible or unethical; for example, the effects of smoking. The case study underpinning this paper is a quasi-experiment.

However, while there are some superficial parallels between medicine and education, the approaches and the underpinning principles do not transfer readily (Pawson, 2006). In the former, the researcher generally has strong control of the environment and they are able to, for example, ensure that the trial participants get the same dosages after being preselected to remove those with complicating medical histories or environmental factors that may have confounding effects. Furthermore, the causal pathway for drugs is often quite simple, where it is possible to predict the biochemical reactions that will occur with a good degree of accuracy: ‘many people, including social scientists and policy makers, who advocate the use of RCTs for shaping policy and practice in nonmedical fields seem unaware that this tool work wells only for “conceptually neat components of clinical practice” in medicine’ too (Lingenfelter, 2016, p. 137, quoting Berwick, 2008).

Social science is rarely ‘conceptually neat’ and there are therefore dangers arising from inappropriate borrowing of ‘what works’ ideas. The following list outlines some of these practical dangers in educational settings:

- *Problem complexity* – social problems worthy of policy attention tend towards complexity, or else they would be resolved without the need for it. Lingenfelter (2016, p.34) argues that ‘complexity presents formidable analytical problems to the scientific method because the number of causal factors and potential interactions among them is enormous [...]. The more numerous and complex the causal factors involved in the problem to be solved or the disease to be cured, the less likely a single intervention can yield satisfactory results.’ In particular, with respect to social problems, the likelihood is that individuals will have different ‘levers’ which are effective at shifting choices, behaviours or outcomes, rather than there being one intervention that works for all. There is no guarantee that this complexity will be eliminated by randomisation, as the match between individual and optimal intervention for them will not be known in advance, such that different statistical results could arise from different experimental/control group configurations (Bickman and Reich 2009).
- *Social complexity* – similarly, educational establishments are complex organisations, containing multiple agents with individualised values, beliefs, attitudes and emotions. They are not well understood by analogy to biochemical reactions and interventions are likely to be subject to a range of behaviours that may be resistive, performative or non-rational from teachers, learners and others (e.g. parents); Pawson (2006) argues that human interactions and intentionality should be the unit of analysis, not crude measures of ‘outcome’. ‘What works’ approaches often rely on randomisation to control for this hyper-complexity, but there is a strong challenge here to Gorard’s (2013) requirement around predictability and repeatability. Lingenfelter (2016, p. 42) conceptualises this as an issue of ‘fidelity’ which compromises both the results of the experimental study and the likelihood of later replicating these results ‘in the field’: ‘no crisp intervention, no “proven program”, can solve a complex problem at scale, in different settings, and over time as situations change.’

- *Law of unmeasured consequences* – ordinarily, there is only a single or small group of outcomes that are measured and analysed for a marked improvement in the area of interest (e.g. reading scores: Gorard, Siddiqui and See, 2015). Therefore, the evaluation may not be aware of possible negatives elsewhere for the learner. For example, they could see their stress levels rise or the attainment in a different subject fall (e.g. due to the intervention taking more of their study time). As a parallel concept, even medicines tested through RCTs are often later found to have severe side effects once implemented ‘in the field’. Indeed, practical medicine is predicated on professional judgement guiding a trade-off between primary and side effects.
- *Passage of time and overlapping interventions* – the longer that an intervention lasts or the longer the time between intervention and outcomes are measured, the less reliable the inference that can be made due to the ever-increasing number of confounding factors than can creep into the study, as well as the scope for participant attrition and contamination between the experimental and control/comparison group (Bickman and Reich, 2009; Cook et al, 2010). This is less problematic if the factors are random, but if they introduce a systematic bias (e.g. schools with poorer outcomes withdrawing) then it can significantly undermine the study. Furthermore, if there are multiple interventions during the life of the study with overlapping aims, as is often the case in educational settings, determining which are effective is particularly problematic (Pawson, 2006). Educational interventions tend to last longer than medical trials and so these difficulties are harder to overcome. However, if studies are kept short to eliminate these issues, they are then unable to provide reliable inference around the long-term (positive, negative and side) effects of an intervention, which may differ significantly from those measured within short timescales.
- *Experimentation effects* – a key challenge is removing systematic biases that derive from the very nature of experimental research (Bickman and Reich, 2009; Scriven, 2009). The so-called Hawthorne effect sensitises individuals to changes in their circumstances, while those delivering the intervention (i.e. usually teachers) will provide subconscious clues as to their standpoint on it by, for example, adding extra enthusiasm to lessons. Within medicine this is alleviated through the double-blind placebo protocol (where neither the researcher nor the participants know whether the latter are getting real drugs), but this is very much harder to achieve in an educational setting. There is a danger that results merely represent the novelty of the intervention, practitioner enthusiasm for it or a secondary feature of the intervention, while those in the control group may become demotivated by their exclusion from the intervention. For example, Gorard, Siddiqui and See (2015) reflect on how the results of their RCT on a reading intervention might be explained, at least in part, by the additional one-to-one attention given to the children rather than the intrinsic success of the intervention itself. ‘What works’ approaches may illuminate causal relationships, but they can be silent on which elements are causal and why, especially in complex situations where causality is likely to be fractured.

- *Baselining and controls* – following on from this, because of day-to-day time and resource constraints, convincing controls can be difficult to construct in educational settings. For example, a control group may be expected to continue with ‘business as usual’ while the experimental group is treated, but this ‘business as usual’ is not neutral and may differ in its quality and outcomes. Establishing a firm baseline against which interventions are measured becomes problematic, especially where trialling a new intervention means forgoing another learning activity. In clinical trials, an element of good practice is to compare the intervention with the established ‘next best’ treatment, but this is a harder judgement in social settings. There is also a danger of simplifying or altering the intervention (e.g. in terms of targeting, duration or scope) to enable it to be more readily subjected to experimental design.
- *Over-expectation* – there is a danger that policymakers, researchers and the originators of the intervention can have unrealistically high expectations of the effect sizes that mark ‘success’. This is partly due to the multiple confounding factors and interventions that are likely to impact on the outcomes, bound up in the complexity of the environment and extended exponentially if the time period is long. The risk is that effective interventions may be discarded if they do not reach a threshold for effect size which may be as arbitrary as cut-offs for statistical significance; how much impact should one expect a single intervention to have on a learner’s life? Similarly, in what proportion of educational settings does an intervention need to work for this to constitute ‘proof’? Drawing on work in the US, Lingerfelter (2016, p. 173) asserts that ‘the experience of the past several decades suggests that very few if any single interventions or programs can consistently generate sizable effects’ such that ‘it is hard to find any intervention that works dramatically well for the most difficult challenges, and only rarely do particular interventions work so much better than competing interventions that their superiority can be persuasively established through RCTs’ (ibid, p.44).
- *Crude utilitarianism* – where the objective is to show an improvement in a key outcome indicator, there is a danger that it does not matter where the change occurs. For example, the result of the intervention could be that five learners in a class do less well, while ten do significantly better. It would be simple to typify this as evidence of success, given that the overall change is positive, but it could mask negative impact for some learners. This tyranny of the majority derives from quantitative methods’ over-reliance on averages to describe distributions and, in this instance, evaluate impact (Mark, 2009). This is an extension of the side effect argument advanced earlier, but at the macro level. The question is thus the extent to which ‘collateral damage’ can be tolerated in pursuit of improved averages.

It is important to stress that none of these critiques is a ‘silver bullet’ that destroys the credibility of ‘what works’ research. They are challenges that need to be addressed and either overcome or acknowledged; poor research results when they are either not understood, ignored or obfuscated

by unsupportable claims to knowledge which are based solely on the privileged nature of RCTs and other experimental methodologies in the eyes of policymakers and others.

Lingenfelter (2016) makes two particularly salient points based on the US experience. Firstly, he asserts that 'what works' has failed to seed significant improvements in educational outcomes, arguing that 'experimental studies with random assignment to treatment and control groups have generally been inconclusive and much less productive than hoped for 50 years ago' (p. 212). Furthermore, 'the weak track record of experimental research in identifying successful social interventions means that admitting only "gold standard" evidence in evaluating social problems has negative practical effects' (p. 44) by squeezing out alternative forms of evidence and inappropriately dismissing successful interventions. Secondly, he reflects on how even the medical profession has more recently tempered its enthusiasm for a pure 'what works' philosophy, increasingly making use of a reciprocal knowledge exchange between researchers and practitioners, particularly with respect to complex medical problems with a social dimension that require a more holistic approach (Berwick, 2008).

Context: widening participation in higher education

The context for the remainder of this paper is higher education in England and specifically the longstanding efforts to widen undergraduate participation to groups that have historically been under-represented – particularly those from low income households, whose chances of attending are around a half that of their more affluent peers (Department for Business, Innovation and Skills, 2015 [BIS]; Higher Education Funding Council for England [HEFCE], 2012). Widening participation has been a policy objective for successive governments since the late 1990s (Department for Education and Employment, 2000; Department for Education and Skills [DfES], 2003), initially as part of an agenda to support the notion of social justice and latterly to support social mobility (Milburn, 2009, 2012).

Over the last ten years, pressure has grown on institutions and other organisations to demonstrate that their efforts are working (e.g. HEFCE, 2007). This is partly in the context of the resource being expended during a period of austerity (although much of the expenditure is now derived from tuition fee income and so not directly from the public purse) and partly because progress in closing the 'social gap' has been slower than hoped for, especially with respect to elite universities (Harrison, 2011, 2012; Croxford and Raffe, 2015). This new 'what works' emphasis is encapsulated in the section on 'evidence-based practice' in the recent national strategy document:

'It is essential to understand which approaches and activities have the greatest impact, and why. An improved evidence base, and a robust approach to evaluation, are critical in helping the sector and partners to understand which of their activities are most effective and have the greatest impact on access, student success and progression, so enabling effort to be focused on these areas' (BIS, 2014, p. 9).

Part of this pressure is being exerted by the Office for Fair Access (OFFA); the quasi-governmental regulatory body established to monitor universities' progress, initially in protecting access for the poorest applicants and latterly in diversifying their social mix. This is largely achieved through the negotiation of 'Access Agreements' (McCaig and Adnett, 2009; McCaig, 2015) which lay out the provisions made, particular around outreach activities and financial support for students, that are funded from income derived from tuition fees levied above a certain threshold. Of particular note, the most recent OFFA strategic plan challenges universities to use a more evidence-led approach to their activities (OFFA, 2015a), although the rhetoric is substantially softer than in BIS (2014). OFFA also commission research to inform practice across the sector, and it is one such project that forms the case study in this paper.

To date, there have been relatively few quantitative efforts to assess the impact of particular interventions, partly due to the epistemological challenges around the construction of control/comparison groups. Hoare and Mann (2012) explore this in detail in their study of 'summer schools', where young people have a residential stay on campus during which they typically undertake curriculum enhancement and motivation-building sessions.

A recent study (Harrison and Waller, in press) of past and current practitioner-managers working within widening participation revealed a high degree of scepticism about the application of 'what works' approaches to their field. While around half of responding universities did report using RCTs or quasi-experiments on an infrequent basis, fewer than 10 percent felt that these approaches provided the most reliable evidence about the success of their programmes. Most confidence was placed in longitudinal tracking studies that collate data about young people's activities, decisions and outcomes over several years. The participants identified a number of specific challenges to demonstrating 'what works', including the time periods involved, the time lag between an intervention and a shift in attitudes, the existence of multiple overlapping interventions, the nature of close partnership working, the confounding influences of teachers and parents and the reliability of data gleaned from young people about their future intentions. More generally, the participants also problematised the very idea of success in this context, with this having different meanings between institutions and individuals.

Case study: university bursaries

Policy background

This case study will focus on the means-tested bursaries introduced through the Higher Education Act 2004 and implemented from 2006 onwards (DfES, 2003). As noted above, this coincided with a significant increase in the tuition fees that universities were permitted to charge (from £1,125 to £3,000 a year), where they were only allowed to implement the rise if they agreed to spend a self-identified proportion of the additional fee income on access, including a bursary scheme for low income students. As nearly all did increase their fees, this effectively became a national

programme. The statutory requirement was for a minimum bursary of £300 a year, but universities were actively encouraged to allocate a larger proportion of their additional fee income, with the intention to attract applicants from low income backgrounds who may otherwise have been deterred by the new higher fees; there was also an intent to create a competitive quasi-market in bursaries, with universities seeking to distinguish themselves through their bursary offer (Harrison, Baxter and Hatt, 2007; Mitton 2007). It is important to note at this point that these bursaries were distinct from the small number of merit-based scholarships historically offered by many institutions, although the boundaries were sometimes blurred (Callender 2010). They were also distinct from the means-tested grants and loan provided by the government, although they generally used the same means-testing figure ('residual household income'¹) calculated when students apply for government support.

Universities' responses to this policy initiative were highly varied as the situation was novel and their freedom over implementation almost complete. A few offered the bare minimum, while others ringfenced significantly larger amounts and were able to offer bursaries of £3,000 or more. There was also a wide diversity in practice in the criteria used to award bursaries. Some used a basic 'threshold' means-test (i.e. constant amount for incomes under certain level), but others developed complex 'tapered' means-tests (i.e. variable amount depending on income) while others still devised additional criteria to target the bursaries to groups that were deemed to be disadvantaged or under-represented (e.g. by ethnicity or locality). Indeed, the situation was exceedingly complex, with over 300 different schemes in operation (McCaig and Adnett, 2009; Callender 2010; Harrison and Hatt 2012). In terms of amounts offered, there was a rapid ossification by institutional status, with elite universities offering substantially larger bursaries to fewer students than lower status universities and colleges, who tended to offer smaller bursaries to large proportions of each cohort, reflecting the differences in social mixes (Callendar, 2010; Harrison and Hatt, 2012). The system was briefly complicated further by the short-lived National Scholarship Programme that operated between 2012 and 2014 (McCaig, 2016; Bowes et al, 2016).

This system has effectively persisted to the current day, though the average bursary size has fallen; indeed, it had begun to fall within two years of their introduction (McCaig and Adnett, 2009). It accounts for an annual national spend of £430m, with around one-third of all students receiving a bursary, although the figure varies widely by institution (OFFA, 2015b). Universities have altered their expenditure, bursary amounts and criteria year-on-year, but still within the broad remit laid down in the Higher Education Act 2004. The statutory requirement on universities has been lifted, but, to date, none has withdrawn from the bursary 'market', although some are considering doing so.

Interestingly, evidence from a large-scale survey commissioned by OFFA found that relatively few students considered bursaries when making choices about entering higher education or which university to choose (Callender, Wilkinson and Hopkins 2009), which is consistent with smaller studies (Davies et al, 2009; Harrison et al, in press) and wider ones finding that students afford financial (dis)incentives a low status in their decision-making (Purcell et al, 2008).

In recent years, therefore, bursaries have been conceptualised less as means of encouraging applications (either in general or to specific institutions) and more in terms of providing funds to enable students from low income and other disadvantaged backgrounds to participate fully in the higher education experience; in other words, less as a financial inducement and more a tool to support student retention and achievement. Indeed, there is evidence from qualitative studies that students find that the bursary reduces stress and anxiety, lower the need to undertake part-time work, improves academic engagement and creates a motivational reciprocity with their university (Harrison, Baxter and Hatt, 2007; Crockford, Hordósy and Simms 2015; Byrne and Cushing 2015; O'Brien 2015; Hoare and Lightfoot 2015; Harrison et al, in press). Both students and practitioners believe that bursaries are an effective tool (Nursaw Associates 2015; Bowes et al 2016), but given the current situation where every university has one or more schemes with unique eligibility criteria, amounts, means-test algorithm and procedures, it is reasonable to hypothesise that some will be more effective than others.

However, a quantitative study of national official data (OFFA, 2014) cast doubt on whether this positivity was being translated into measurable positive changes in behaviour, albeit using bursary data at the institutional level (i.e. not actual amounts awarded) and with a single narrow definition of a positive outcome (retention into a second year of study). University managers want to know if bursaries work given the large sums invested in them, while OFFA is keen to evaluate the success against national targets for improvements in student participation, retention and success. It is in the context of resolving the epistemological conflict between the large quantitative study and the various localised qualitative studies that the case study project described below arose (OFFA, 2015c).

OFFA project

In March 2015, OFFA issued an invitation to tender for a research team to develop a methodology for universities to assess the success of their own bursary programme (OFFA, 2015c). The purpose of the project was not to directly research bursaries' effectiveness within the project, but to lay out a standardised framework for analysis which would enable institutions to evaluate their own bursary scheme in a robust and comparable way. This was based on the hypothesis that some bursary schemes might be more successful than others and that forms of good practice might emerge, as well as to help universities to better target support at those in most need.

A team drawn from five universities (Bedfordshire, Kings College London, Oxford, Sheffield Hallam and West of England) was awarded the contract in June 2015. The project was divided into two phases, with the first focusing on the statistical analysis of historical administrative data from universities' student records systems and the second on gathering data from students with bursaries through an online questionnaire around their experiences and how they had used their bursary. This case study relates solely to the first phase, the report of which was published in February 2016 (OFFA, 2016).

As outlined above, the overarching context of the project was to determine whether bursaries were effective in terms of improving the outcomes for students from low income and other disadvantaged backgrounds. As a stepping stone, the team undertook analysis of data from the five institutions represented, using a pilot version of the analytical framework. The analysis itself is considered commercially-sensitive by the participating universities and so is not reported here. Instead, this article provides reflections from the project based on grappling with the underlying epistemological challenges and contradictions of attempting to quantitatively evaluate the success of a major governmental policy initiative and a significant expenditure for individual universities: in short, do they 'work' and how can we know?

To this end, the research question used to guide the project on behalf of the future users of the analytical framework was: *Do financial bursaries for financially disadvantaged students ameliorate their educational disadvantage relative to other students?*

Problematizing the question

There were a range of epistemological issues to be grappled with within this question. Firstly, there could be no true counterfactual analysis. The project was not conceived to be fully experimental (even if this were possible to achieve within an ethical research framework) and it was predicated on the use of historic administrative data – i.e. it was quasi-experimental. There was no means of examining what would happen to students if their bursary had not existed or if it had been doubled – a putative project that was able to compare similar students *across* institutions with different bursary schemes could potentially address this issue, but this was out of scope for the project in hand.

The project therefore had to be based on a comparison between bursary holders and students who were not bursary holders. The latter could not be a control group, not least as the nature of bursaries is that, in general, all students passing the means-test receive a bursary – i.e. there should be no low income students without bursaries. Instead, the project was based around the idea of a comparison group with marginally less financial disadvantage than the bursary group (and therefore ineligible for a bursary). We will return to this mid-income comparison group shortly.

The research question itself makes the not-unreasonable assumption that students who are financially disadvantaged (i.e. come from a low income household) suffer educational disadvantage as a result. Much of this disadvantage will have accumulated prior to university and be manifest in the qualifications that the student holds (Gorard and See, 2013; Crawford, 2014) and the extent to which these represent intellectual preparedness for higher study. The acknowledgment of this underpins the 'contextualised admissions' movement, which seeks to give disadvantaged students preferential treatment within the admissions system on the basis that their qualifications are not a reliable representation of their ability and/or potential (Hoare and Johnston, 2011; Moore, Mountford-Zimdars and Wiggans, 2013).

However, other educational disadvantage is likely to accumulate during their time as a student. They are, for example, more likely to need to take on part-time work to support their living costs, limiting time for study and potentially causing anxiety (Jessop, Herberts and Solomon 2005; Humphrey, 2006; Robotham 2008; Callender 2008). They are less likely to be able to afford the books, equipment and materials to support their study. There may be other indirect forms of disadvantage too – e.g. due to affordable accommodation options (including remaining in the family home), which increase the time required for travelling or limit integration to the university community.

The subsequent assumption is that these forms of financially-related disadvantage would usually be reflected in commonly-used measures of ‘formal’ outcome that are often used as performance indicators within the sector (e.g. retention, degree result and employment). This, however, is more problematic. For example, it may be that these students develop greater resilience by overcoming barriers earlier in their education, have a higher motivation to succeed, possess good coping or study strategies or benefit from stronger family or community support (e.g. Marshall and Case, 2010; Clegg, 2011; Modood, 2012; McKay and Devlin, in press); these are exceptionally difficult to measure with reliability. Students may also simply adopt more frugal lifestyles to allow them to focus their limited finances and time towards the expenditure with the most academic value, eschewing elements of the stereotypical student experience (Harrison et al, 2015).

Furthermore, students from low income backgrounds are likely to have been those targeted by universities as part of their outreach programmes, especially if living in deprived areas (Harrison and McCaig, 2015). Many will have had access to significant long-term activities intended to develop their preparedness for higher education and to help them to achieve on arrival – e.g. summer schools, revision classes or mentoring programmes. Having a bursary may therefore act as a proxy for a lengthy positive engagement with their university (and/or others), with associated gains in terms of skills, motivation and/or commitment; indeed, many universities use participation in their outreach programme as one criterion for awarding bursaries.

All in all, it is not entirely clear how financial disadvantage causally implies poorer ‘formal’ outcomes, because of the complexity of the determinants of these outcomes, the passage of time and overlapping interventions designed to support these students. Recent reports published by and for HEFCE have suggested that students from disadvantaged neighbourhoods are slightly less likely to obtain first or upper second class degrees once other factors are taken into account, but that the intersectionality reflected in the interaction terms within the statistical model makes interpretation and inference vexed (HEFCE, 2014; Mountford-Zimdars et al, 2015). For example, it may be that this effect can be explained, at least in part, by the increased likelihood of such students remaining in the family home while studying, which is also seen as a risk factor. Conversely, the same reports are more solid in their finding that students from state schools are likely to outperform their private school peers, all else (including entry qualifications) being equal. The field is therefore clearly complex, but something of a deficit discourse around disadvantaged

students remains alongside an assumption of underperformance (Clegg, 2011; Smit, 2012; McKay and Devlin, in press).

However, it is important to remember that even if financial disadvantage does not result in lower formal outcomes, it does not preclude poorer 'informal' outcomes from higher education, including stress, anxiety, problematic debt and social isolation; such conditions may also reduce opportunities for the accumulation of valuable work skills and experience. In other words, if low income students *are* achieving as well (or better) than their more-advantaged peers, then this may be at the expense of their wellbeing, their relationships or their long-term employability. There is a clear danger of systematic unfairness in this regard and, arguably, these informal outcomes are more significant to the individual and their human flourishing.

What then is the purpose of bursaries? The discussion above notwithstanding, the overarching policy aim is that they should allow low income students to compete on (more) equal terms with their more affluent peers – in effect, 'levelling the playing field'. The purpose is *not* to provide an *advantage* to the extent that students with bursaries have significantly *better* outcomes than those peers. This would likely be seen as unfair by other students – especially those marginally above the eligibility threshold. This is an important point which will be returned to shortly. What is moot is the extent to which this purpose extends to informal outcomes as well as the formal ones valorised by policymakers.

Making valid inference

Because of the quasi-experimental nature of the data available, binary logistic regression analysis was selected for this project. This technique allows the researcher to examine the individual contribution of a range of variables to a dichotomous outcome measure, while holding the others equal. A full explanation of the statistical model used is outside of the scope of this paper (see OFFA, 2016), but 15 control variables were used alongside four dichotomous measures of outcome:

1. Whether the student had been retained into their second year of study at the same university (not necessarily having progressed);
2. Whether the student completed a degree programme within five years (longer programmes such as medicine were excluded);
3. Whether the student achieved a 'good' degree – a first or upper second class degree;
4. Whether the student was in graduate-level employment six months after graduation, with various exceptions for travelling, parenthood etc.

The analytical intent was thus to see whether there was a statistically significant difference between the group of bursary holders and the comparison group of students from mid-income backgrounds who were ineligible for a bursary, whilst controlling for potentially important factors on which universities hold data, such as entry qualifications, gender, degree subject, disability and

so forth. While the data management task entailed was complex and significant, the analysis was purposefully kept relatively simple to enable a wide range of institutions to adopt it without additional software or staff training needs.

However, it was at this point that additional epistemological challenges began to emerge. As noted above, the full results of the analysis are commercially confidential to the individual universities, but three different types of result were represented – there were no specific patterns by university and most had a mixture of results across the four outcome measures:

- **Result 1:** Bursary students had significantly worse outcomes than the comparison group
- **Result 2:** Bursary students had similar outcomes to the comparison group – i.e. there was no statistically significant difference
- **Result 3:** Bursary students had significantly better outcomes than the comparison group

The problem arose in inferring impact or success from these statistical results because of the uncertain relative starting position of individual students, as discussed above. While the model took account of a wide range of demographic and situational factors, it was necessarily mute on individual dispositional factors like motivation, resilience or preparedness for which no data are available on student records systems.

For example, Result 1 initially suggests that bursaries are ineffectual; all else being equal, students with bursaries underperform relative to their comparison group. However, there is more complexity here; indeed, there is no means to infer that the bursary was not working. On the contrary, it would be equally valid to conclude that the bursaries being offered were simply insufficient to overcome the educational disadvantage resulting from students' financial disadvantage. Rather than consigning bursaries to the policy dustbin, Result 1 could, in fact, be seen as relatively strong evidence for the 'missing' causal link between financial and educational disadvantage discussed above, with the appropriate policy response potentially being to *increase* the amount of bursary available.

Similarly, Result 2 was vexed. There was a strong temptation to interpret this as being a negative result, much as had been done in OFFA (2014): bursary holders were only doing as well as the comparison group, *ergo* bursaries were not working. Given the discussion above, this inference is clearly fallacious. Firstly, it is important to remember that the comparison group is not a control group and, given their (albeit modest) relative advantage, one would expect them to show somewhat stronger outcomes, *ceteris paribus*; it does, of course, depend again on the starting point of the individual students. If financial disadvantage is causal in educational disadvantage, then Result 2 represents impact and success, with the low income students doing as well as mid-income students. However, if financial disadvantage is not causal in educational disadvantage (insofar as this is captured by the outcome measures described above), then Result 2 represents a failure of bursaries. The inference taken depends wholly on the initial assumptions.

Again, there is also a further complication. While several analyses showed no significant difference between bursary holders and the comparison group, they did show a relatively large and positive effect size (as represented by the ‘odds ratio’). This highlights a problem with an over-reliance on statistical significance (Gorard and See, 2013). In particular, the ‘bar’ for significance in logistic regression models becomes very high when the outcome is very (un)common within the population. For example, if 95 percent of students are retained into their second year, showing statistically that bursary holders are *even more likely* to be retained is difficult, especially if this subgroup is small – even a 100 percent retention rate might not be enough to demonstrate significance. The inference drawn from this type of result depends whether the reader privileges effect size or significance.

Finally, even Result 3 required consideration. *Prima facie*, this appeared to a positive endorsement of bursaries as it showed low income students having stronger outcomes than the comparison group. Indeed, this was initially felt to be the result of most interest for the project. However, on reflection, it also raised further questions. Firstly, there is the element of fairness discussed above – if bursaries are providing excessive amelioration for low income students such that they are actually outperforming mid-income students, then there is a question of proportionality. Secondly, and returning to a theme already rehearsed, it may be that the difference in outcomes is not causally related to the scale of the bursary, but rather to other relative advantages among low income students. If this were the case, then it is impossible to determine the additional contribution that the bursary might be making – they effectively become irrelevant in terms of the formal outcomes. Interestingly, Result 3 tended to be associated with bursaries given on the basis of outreach programmes and alternative entry routes into higher education where, perhaps, the effect was as much about bursaries being a proxy for preparedness and attachment as financial support; this is moot, however. The three results are summarised in Table 1.

Table 1: Summary of possible inferences to be drawn about the effectiveness of bursaries

Result 1: bursary holders have weaker outcomes than comparison group	Bursaries are either ineffective or insufficient to completely overcome the effects of financial disadvantage
Result 2: bursary holders have same outcomes as comparison group	Bursaries are effective (unless there is no impact of financial disadvantage on educational outcomes)
Result 3: bursary holders have stronger outcomes than comparison group	Bursaries are very effective (and possibly unfairly so with respect to the mid-income group)

However, the story does not end here. The wider moral questions about the participation of low income students in higher education also demand consideration. The evaluation of the impact of

bursaries is, perhaps necessarily, couched in terms of measurable formal outcomes. However, there are myriad informal outcomes which were not (and could not) be measured within the first phase of this project and which extend beyond a student's ability to see through their studies to a high standard and move into graduate-level employment. As noted above, this might include participation in the wider university community, the development of positive social relationships, the ability to undertake curricular enhancements (e.g. field trips or internships), the maintenance of mental health and wellbeing and so on. Even if the effectiveness of bursaries in terms of headline measures of formal outcomes are questioned, it is very possible that they contribute positively to students' ability to have a full and fulfilling student experience; indeed, this is what students themselves consistently report (e.g. Harrison et al, in press) and the project's second phase is designed to explore through the development of a survey tool that can be used by institutions to explore the views of bursary recipients. It would be morally questionable to remove a provision with positive human impacts due to epistemological challenges with proving impact against relatively crude policy-friendly measures. This tendency to 'privilege the measurable' always needs to be thoroughly problematised.

Concluding remarks

The purpose of this paper has been to use a reflective case study to explore some of the challenges associated with employing a 'what works' approach with respect to a high-cost policy initiative. The project has been a useful vehicle for thinking carefully about evidence and inference in a complex educational field.

A clear epistemology for evaluating success has to be predicated on a clear definition of that success. However, this has not yet been established for bursaries, partly due to the relatively recent co-option from their original purpose of bolstering demand to a generalised tool in supporting retention and success. There is currently an absence of a clear and widely-accepted causal model for the effects of financial disadvantage on educational outcomes; continuing the earlier medical analogy, this would be akin to treating a disease without having a clear idea of its pathology. Furthermore, the definition of the formal outcomes under scrutiny may not match with wider social justice objectives (prioritised by practitioners and students, if not always by policymakers) about ensuring that students remain healthy, expand their intellectual horizons and are able to enjoy a reasonable quality of life. These latter ideas will be explored in the second phase of the project by engaging directly with students.

In particular, bursaries sit in a complex space of support for disadvantaged students, who are exposed to a wide range of overlapping interventions over a lengthy time period, both before and during university, that are intended to ameliorate their disadvantage. Disentangling any unique effect of bursaries from, for example, the impact of summer schools or peer mentoring is vexed and there are limits to the number of factors that can be controlled within a statistical model, even if they can be measured. Even if it had been possible to construct a formal control group, this would still have posed a significant (and probably insurmountable) challenge. Similarly, with no readily-available measures of potential dispositional or experiential factors, it is difficult to

assess any possible advantages acquired through overcoming pre-university challenges. Perhaps most important is Lingenfelter's (2016) rejection of the idea that interventions in social fields can be assessed in isolation, especially as it is difficult to see how Gorard's (2013) criteria could be met.

Furthermore, the freedom that universities have in the implementation of bursaries means that they play a different role in each instance and so attempts to generalise about the effectiveness of bursaries in the round are likely to be misguided. Even within the five pilot universities, there was significant divergence in the administration, amounts and criteria. This was thoroughly understood by OFFA in commissioning the work, with their focus being on supporting universities to examine their own schemes in isolation, albeit with a common analytical framework. That said, there is a strong degree of consensus from practitioners and researchers that bursaries do 'work', which may exemplify the challenges of effectively evidencing practitioner 'knowledge'. While the project has met its objective of providing an analytical framework for universities to use, it maybe that quasi-experimental approaches cannot move us far beyond an agreement that Result 2 (as defined above) is good evidence for effectiveness in 'levelling the playing field', while Results 1 and 3 are causes for different forms of concern for universities.

One useful contribution of this paper is to reassert the difference between a control group and a comparison group and the ramifications of this for the interpretation of results, particularly with respect to OFFA (2014). Where the comparator group is specifically less disadvantaged than the quasi-experimental group, then it is unrealistic to expect the latter to have better outcomes than the former; statistically similar results are a marker for success, not failure, unless it can be shown that there are no educational disadvantages accruing from financial disadvantage.

Due to the ethical issues, it is rather inconceivable that the success of bursaries could be examined through an RCT (although see Angrist, Lang and Oreopolous, 2009) and even if this were possible, it is unlikely that it could untangle the web of overlapping interventions outlined above. Future work might be based around identifying universities that are at the point of changing their bursary schemes and using a natural experiment of comparing matched cohorts before and after the change (who would therefore get different amounts), provided that other components of support for disadvantaged students remained broadly constant; i.e. an interrupted time series design. This is similar to the approach used by Denny et al (2014), who were able to robustly demonstrate the success of an integrated academic, social and financial support package at an Irish university. This could potential provide more directly causal evidence about the success of bursaries, albeit that the difficulties in measuring the vital informal outcomes would remain problematic. Conversely, the data produced by robust qualitative designs may be able to provide stronger and more useful evidence for the means by which bursaries ameliorate educational disadvantage than any 'what works' approach.

Finally, this reflective case study also has ramifications for wider projects intended to investigate 'what works' in higher education, especially around the recruitment and participation of disadvantaged students. These are likely to suffer from similar issues around time, complexity of

field, unmeasurable variables, comparison groups and overlapping interventions, with dangers for invalid inference about the success of individual activities, especially where these are implemented in very different ways in different universities. This is not a manifesto for not trying, but rather a cautionary tale about the limits of what the 'what works' approach can really tell us and an *aide memoire* that evidence is not proof.

References

- Bickman, L. and S. Reich. 2009. Randomized controlled trials: a gold standard with feet of clay?, in S. Donaldson, C. Christie and M. Mark (eds.) *What counts as credible evidence in applied research and evaluation practice*, Thousand Oaks: Sage.
- Biesta, G. 2007. Why 'what works' won't work: evidence-based practice and democratic deficit in educational research. *Educational Theory* 57 (1): 1-22.
- Biesta, G. 2010. Why 'what works' still won't work: from evidence-based education to value-based education. *Studies in Philosophy and Education* 29 (5): 491-503.
- Bowes, L., R. Moreton, L. Thomas, J. Sheen, G. Birkin and S. Richards. 2016. *Evaluation of the National Scholarship Programme: year 4 report*. Leicester: CFE Research.
- Byrne, L. and S. Cushing. 2015. The impact of structured financial support on student retention case study: Buckinghamshire New University. *Widening Participation and Lifelong Learning* 17 (3): 47-59.
- Callender, C. 2008. The impact of term-time employment on higher education students' academic attainment and achievement. *Journal of Education Policy* 23 (4): 359-377.
- Callender, C. 2010. Bursaries and institutional aid in higher education in England: Do they safeguard and promote fair access? *Oxford Review of Education* 36(1): 45-62.
- Callender, C. and J. Jackson. 2005. Does the fear of debt deter students from higher education? *Journal of Social Policy* 34 (4): 509-540.
- Callender, C. and D. Wilkinson. 2013. Student perceptions of the impact of bursaries and institutional aid on their higher education choices and the implications for the National Scholarship Programme in England. *Journal of Social Policy* 42 (2): 281-308.
- Callender, C., D. Wilkinson, and R. Hopkin. 2009. *The impact of institutional financial support in England: higher education students' awareness, knowledge and take-up of bursaries and scholarships*. Bristol: OFFA.
- Clegg, S. 2011. Cultural capital and agency: connecting critique and curriculum in higher education. *British Journal of Sociology of Education* 32 (1): 93-108.
- Cook, T., M. Scriven, C. Coryn and S. Evergreen. 2010. Contemporary thinking about causation in evaluation: a dialogue with Tom Cook and Michael Scriven. *American Journal of Evaluation* 31 (1): 105-117.
- Coles, J. 2012. Keynote address to *British Educational Research Association Annual Conference*, Manchester, 6th September.
- Connolly, P. 2015. The trials of evidence-based practice in education, keynote address to *British Educational Research Association Annual Conference*, Belfast, 16th September.
- Crockford, J., R. Hordósy and K. Simms. 2015. 'I really needed a job, like, for money and stuff': Student finance, part-time work and the student experience at a northern red-brick university. *Widening Participation and Lifelong Learning* 17 (3): 89-109.
- Davies, P., K. Slack, A. Hughes, J. Mangan and K. Vigurs. 2008. *Knowing where to study: fees, bursaries and fair access*. Stoke-on-Trent: Institute for Educational Policy Research and Institute for Access Studies.

- Denny, K., O. Doyle, P. McMullin and V. O'Sullivan. 2014. Money, mentoring and making friends: the impact of a multidimensional access program on student performance. *Economics of Education Review* 40: 167-182.
- Department for Business, Innovation and Skills [BIS]. 2014. *National strategy for access and student success in higher education*. London: BIS.
- Department for Business, Innovation and Skills [BIS]. 2015. *The proportion of 15 year olds from low income backgrounds in English maintained schools progressing to HE by the age of 19: October 2015*. London: BIS.
- Department for Education and Skills [DfES]. 2003. *Widening participation in higher education*, London: DfES.
- Esson, J. and H. Ertl. In press. No point worrying? Potential undergraduates, study-related debt, and the financial allure of higher education, *Studies in Higher Education*, DOI:10.1080/03075079.2014.968542.
- Goldacre, B. 2013. *Building evidence into education*. Available online at http://dera.ioe.ac.uk/17530/7/ben%20goldacre%20paper_Redacted.pdf.
- Gorard, S. 2004. The British Educational Research Association and the future of educational research. *Educational Studies* 30 (1): 65-76.
- Gorard, S. 2013. *Creating robust approaches for the social sciences*. London: Sage.
- Gorard, S. and B.H. See. 2013. *Overcoming disadvantage in education*. Abingdon: Routledge.
- Gorard, S., N. Siddiqui and B.H. See. 2015. An evaluation of the 'Switch-on Reading' literacy catch-up programme. *British Educational Research Journal* 41 (4): 596-612.
- Hammersley, M. 2013. *The myth of research-based policy and practice*. London: Sage.
- Harding, J. 2011. Financial circumstances, financial difficulties and academic achievement among first-year undergraduates. *Journal of Further and Higher Education* 35 (4): 483-499.
- Harrison, N. In press. Exploring social trends in higher education participation: 2000 to 2015. In R. Waller, N. Ingram and M. Ward, *Higher education and social inequalities: Getting in, getting on and getting out*. Abingdon: Routledge/BSA.
- Harrison, N., A. Baxter and S. Hatt. 2007. From opportunity to OFFA: discretionary bursaries and their impact on student finance, academic success and institutional attachment. *Journal of Access Policy & Practice* 5: 3-21.
- Harrison, N. and S. Hatt. 2012. Expensive and failing? The role of student bursaries in widening participation and fair access in England. *Studies in Higher Education* 37: 695-712.
- Higher Education Funding Council for England [HEFCE]. 2007. *Higher education outreach: targeting disadvantaged learners* (Report 2007/12). Bristol: HEFCE.
- Higher Education Funding Council for England [HEFCE]. 2012. *POLAR3: Young participation in higher education* (Report 2012/26). Bristol: HEFCE.
- Higher Education Funding Council for England [HEFCE]. 2014. *Differences in degree outcomes: key findings* (Report 2014/03). Bristol: HEFCE.
- Hoare, T. and R. Johnston. 2011. Widening participation through admissions policy - a British case study of school and university performance. *Studies in Higher Education* 36 (1): 21-41.
- Hoare, T. and J. Lightfoot. 2015. Student funding, student retention and student experiences: perspectives from Bristol. *Widening Participation and Lifelong Learning* 17 (3): 110-125.

- Hoare, T. and R. Mann. 2012. *The impact of the Sutton Trust's Summer Schools*. London: Sutton Trust.
- Humphrey, R. 2006. Pulling structured inequality into higher education: the impact of part-time working on English university students. *Higher Education Quarterly* 60 (3): 270-286.
- Jessop, D., C. Herberts and L. Solomon. 2005. The impact of financial circumstances on student health. *British Journal of Health Psychology* 10(3): 421-439.
- Lingenfelter, P. 2016. *"Proof", policy and practice: understanding the role of evidence in improving education*. Sterling, VA: Stylus.
- McCaig, C. 2016. The retreat from widening participation? The National Scholarship Programme and new access agreements in English higher education. *Studies in Higher Education* 41 (2): 215-230.
- McCaig, C. and N. Adnett. 2009. English universities, additional fee income and Access Agreements: their impact on widening participation and fair access. *British Journal of Educational Studies* 57: 18-36.
- McKay, J. and M. Devlin. In press. 'Low income doesn't mean stupid and destined for failure': challenging the deficit discourse around students from low SES backgrounds in higher education. *International Journal of Inclusive Education*, DOI:10.1080/13603116.2015.1079273
- Mark, M. 2009. Credible evidence: changing the terms of the debate, in S. Donaldson, C. Christie and M. Mark (eds.) *What counts as credible evidence in applied research and evaluation practice*, Thousand Oaks: Sage.
- Marshall, D. and J. Case. 2010. Rethinking 'disadvantage' in higher education: a paradigmatic case study using narrative analysis. *Studies in Higher Education* 35 (5): 491-504.
- Menter, I. 2013. From interesting times to critical times? Teacher education and educational research in England. *Research in Teacher Education* 3 (1): 38-40.
- Milburn, A. 2009. *Unleashing aspiration: the final report of the Panel on Fair Access to the Professions*. London: PFAP.
- Milburn, A. 2012. *University challenge: how higher education can advance social mobility*. London: Cabinet Office.
- Modood, T. 2012. Capitals, ethnicity and higher education. In S. Tomlinson and T. Basit (eds.) *Social inclusion and higher education*. Bristol: Policy Press.
- Moore, J., A. Mountford-Zimdars and J. Wiggans. 2013. *Contextualised admissions: examining the evidence*. Cheltenham: Supporting Professionalism in Admissions Programme.
- Mountford-Zimdars, A., D. Sabri, J. Moore, J. Sanders, S. Jones and L. Higham. 2015. *Causes of difference in student outcomes*. Bristol: HEFCE.
- National Committee of Inquiry into Higher Education [1997] *Higher Education in the Learning Society*, (the Dearing Report). London: HMSO.
- Nursaw Associates. 2015. *What do we know about the impact of financial support on access and student success?* Bristol: OFFA.
- Oancea, A. and R. Pring. 2008. The importance of being thorough: on systematic accumulations of 'what works' in education research. *Journal of Philosophy of Education* 42 (s1): 15-39.
- O'Brien, M. 2015. Student finance, progression and 'inclusivity': indicative data from the University of Liverpool. *Widening Participation and Lifelong Learning* 17 (3): 74-88.

- Office for Fair Access [OFFA]. 2010. *Have bursaries influenced choice between universities?* Bristol: OFFA.
- Office for Fair Access [OFFA]. 2014. *Do bursaries have an effect on retention rates?* Bristol: OFFA.
- Office for Fair Access [OFFA]. 2015a. *Strategic plan: 2015-2020*. Bristol: OFFA.
- Office for Fair Access [OFFA]. 2015b. *Outcomes of access agreement monitoring for 2013-14*. Bristol: OFFA.
- Office for Fair Access [OFFA]. 2015c. *Invitation to tender – understanding the impact of financial support on student success: an institutional approach*. Bristol: OFFA.
- Office for Fair Access [OFFA]. 2016. *Understanding the impact of institutional financial support on student success: phase one report*. Bristol: OFFA/Sheffield Hallam University.
- Pawson, R. 2006. *Evidence-based policy: a realist perspective*. London: Sage.
- Purcell, K., P. Elias, R. Ellison, G. Atfield, D. Adam and I. Livanos. 2008. *Applying for higher education: the diversity of career choices, plans and expectations*. Coventry: Warwick Institute for Employment Research.
- Reed, R. and B. Hurd. In press. A value beyond money? Assessing the impact of equity scholarships: from access to success, *Studies in Higher Education*, DOI: 10.1080/03075079.2014.968541.
- Robotham, D. 2008. Stress among higher education students: towards a research agenda. *Higher Education* 56 (6): 735-745.
- Scriven, M. 2009. Demythologising causation and evidence, in S. Donaldson, C. Christie and M. Mark (eds.) *What counts as credible evidence in applied research and evaluation practice*, Thousand Oaks: Sage.
- Slavin, R. 2004. Education research can and must address 'what works' questions. *Educational Researcher* 33 (1): 27-28.
- Smit, R. 2012. Towards a clearer understanding of student disadvantage in higher education: problematising deficit thinking. *Higher Education Research & Development* 31 (3): 369-380.
- Webster, R.S., 2009. How evidence-based teaching practices are challenged by a Deweyan approach to education. *Asia-Pacific Journal of Teacher Education* 37 (2): 215-227.
- Whitehead, J., J. Raffan and R. Deaney. 2006. University choice: What influences the decisions of academically successful post-16 students? *Higher Education Quarterly*, 60 (1): 4-26.

¹ This works from the income of the student's parent(s) if they are under 25 or their own household if they are 25 or over or otherwise deemed to be independent of their parents. A number of algorithmic deductions are then made (e.g. for multiple siblings in higher education) before arriving at the 'residual household income' figure which is used to determine government financial support and passed to universities (if the student gives permission) to determine bursary eligibility.