What works in helping people and why? Lessons from evidence based practice for effective leadership in social work

Donald Forrester[1]

Abstract: *Randomized controlled trials (RCTs) are the most rigorous test of effectiveness for any intervention. This article considers RCTs as policy projects, and outlines the key elements of effective delivery of interventions within an RCT. It is argued that conceptualizing RCTs as practice delivery projects provides insights of relevance for effective leadership in social work or other helping professions. Three elements of effective RCT delivery are suggested to be crucial: (a) a theory about the nature of the issues being worked with and how people can be helped by professionals; (b) a detailed description of practice that flows from this and (c) skills development and monitoring methods for ensuring practitioners are delivering practice in this way. These are argued to be key components required for effective leadership in social work. Finally, it is suggested that while there is increasing focus on the common elements within specific interventions there has been a neglect of the common features associated with their delivery. Attention to these might help explain the 'Dodo bird effect', in which different interventions often prove to be equally effective, and thereby enrich our appreciation of what works, for whom and why.*

1. Professor of Social Work, University of Bedfordshire

Address for correspondence: *Tilda Goldberg Centre for Social Work and Social Care Research, University of Bedfordshire, Park Square, Luton LU1 3JU Donald.Forrester@beds.ac.uk*

Introduction

The question 'what works and why?' has been at the heart of attempts to develop evidence based approaches across social work, psychology and other disciplines interested in helping people for 40 years or more. The focus of this article is to review the evidence base on what works from a broader and more conceptual perspective. In doing so, the intention is not to draw-up a list of 'what works', but rather to consider what the more general lessons are from the literature. In this respect the emphasis is more on understanding the 'why' than attempting to list the 'what'. In pursuing an understanding of the underlying and sometimes unexamined reasons why some approaches 'work', it is hoped that some key lessons for the delivery of effective services are identified. In this respect the imagined audience for this piece is not only academics, students and practitioners but also one that is rarely catered for, namely managers or other leaders within social work interested in lessons from research and their implications for running effective services.

Evidence based practice and evidence based interventions

There are myriad approaches to embedding evidence within policy and practice. This article focuses primarily on the approach most strongly associated with evidence based practice, namely the attempt to develop and evaluate evidence based interventions (EBIs). EBIs are ways of working with people that have the following four characteristics:

1. A theory about what the 'problem' under study is and how people can be helped;
2. A practice that describes in some detail the way in which people can help (following from the theory);
3. A method for creating skilled practitioners able to deliver this practice;
4. Evidence that they then tend to produce better outcomes than service as usual.

One of the advantages of the existence of these elements is that as a result, EBIs can be clearly described. This may take the form of a manual or a more general description, largely depending what purpose the description is serving, but essentially the existence of the first 3 elements allows a clear specification for an EBI which can then be tested in research. Thus, for instance, cognitive behavioural therapy (CBT) starts with a theory, drawn from the behaviourist traditions of psychology, that the thoughts people have contribute to their depression (for instance, constantly feeling like a failure; thinking that one has no friends), that such 'negative cognitions' can be changed in various ways and that doing so will reduce the symptoms of depression (Hofmann, 2011; Lambert et al, 2004). CBT therefore involves working with people

to help them to change their negative cognitions. Training in CBT involves extensive clinical supervision to achieve the ability to engage and help people in this way. Research evidence shows that CBT is an effective way of helping people experiencing depression (Driessen and Hollon, 2010).

At the heart of the EBI tradition is a particular research design: the randomized controlled trial (RCT) (see Shadish et al, 2002). RCTs are a method of elegant simplicity for ruling out other explanations for findings. A before and after study does not prove that an intervention 'caused' an outcome as people may have got better (or worse) anyway. After all, people are not passive recipients of interventions but actively making decisions and (often) resolving issues in their lives; they therefore often resolve issues without professional involvement. RCTs allow us to take account of what might have happened anyway by randomising a group of participants to receive either the EBI being studied or something else (service as usual or a different intervention depending what the research question is). If people are truly randomized and there are sufficient numbers of them then any difference in outcomes is likely to be due to the intervention being studied.

RCTs are one of the greatest intellectual achievements of the 20th century (though first noted as an approach in the 19th century, and with the basic comparative method being mentioned in the book of Daniel in the Bible, they were only implemented in the 20th) (Torgersen and Torgersen, 2008). Developed in agriculture and then first used in educational and social work settings (Oakley, 2002), their potential was rapidly realised in medicine and they form the foundation for evidence about what works in health interventions (Medical Research Council (MRC), 2008). It seems certain that in various forms the RCT will always be with us as a rigorous test of whether something – from a pill to a social reform – makes a difference. While RCTs have been refined and developed to become in themselves a highly specialised research approach, at heart they retain an appealing simplicity: they make the incredibly complex real world simple and therefore allow us to look at the relationship between one intervention (for example, a pill or an EBI) and one outcome (or a small number), with everything else being the same between groups.

There is little doubt that in social work we could use RCTs more often. If we do not know whether a new way of training social workers or a different way of working with families is likely to be effective then an RCT could often be undertaken to provide an answer. There is a common misconception that RCTs are necessarily expensive or complicated. This is not the case, indeed provided there is a rigorous approach to randomization the analysis of an RCT is straightforward, and there are often natural opportunities to randomize where demand exceeds supply or a new untested way of working is being proposed. Yet we have carried out incredibly few in UK social work – and almost none in the area of child and family work (Forrester, 2012). It is perhaps little wonder that we along with other public services are being exhorted to carry out experimental trials more often by government (Haynes et al, 2011).

One of the reasons for the lack of use of RCTs within the UK is that there has been

widespread critique of their use from the social work academic community. There are valid reasons for a healthy scepticism about RCTs, however it is worth discounting one criticism of RCTs which is oft voiced. Webb's 2001 article has been identified as the most cited article in social work of the last 20 years (Webb, 2001). It articulates a widespread view (shared for instance by Adams et al, 2009 and others) that argues that helping relationships in social work are so complex that attempts to apply RCTs to them and develop general rules for evidence based practice are impossible. There are two problems with this argument. First, it suffers from a category error: EBIs and evidence based approaches more generally are considered as choices for individual practitioners. In fact, EBIs are choices for systems. Thus, Webb suggests that the complexity of each individual, their situation and the specific conversation make it impossible to specify the right response. Clearly, this is true. However, for a whole service (whether a family centre, a local authority or a country) it is both possible and appropriate for leaders to specify a way of working. Thus, for an individual woman who is depressed an approach may need to be tailored to their needs, but for a service, choosing to use CBT may be wholly appropriate, and doing so would necessitate a commitment to training, supervision and so on in order to ensure that CBT was delivered. Second, any argument that social life is too complex for the use of RCTs is a complete misunderstanding; RCTs are important *because* they help us research this complexity. The human body (in its context) is an incredibly complex system and it is not possible to be sure whether a particular medicine or treatment is responsible for a particular outcome. It is for this reason that RCTs have a particularly important role to play – they keep everything else equal and allow us to focus on one issue.

The focus of this article is not a discussion of the pros and cons of RCTs (see Forrester 2012, Bonell et al 2012 for such discussions). Rather, it considers the literature on EBIs to explore what lessons can be learnt about the effective delivery of helping services: after all, EBIs tend to out-perform 'service as usual'. How do they do this? To understand how this is achieved RCTs are considered as policy projects. At the heart of any RCT is the fact that a researcher wishes to ensure that some workers deliver a specific EBI. This article analyses how researchers attempt to do this. It turns out that the answer has potential lessons for both the effective leadership of helping services and for the research literature on 'what works'. We turn to the second of these areas first, as consideration of some unexpected findings leads to a deeper appreciation of RCTs as policy projects.

Black boxes and dodos

A challenge for EBIs is that they have not on the whole produced the type of findings that might have been expected. To illustrate this I will dwell on the substance misuse field, as it has a far better developed set of experimental studies than social

work, however the general point seems to work across a wide variety of fields. In general terms, when credible interventions are compared to service as usual they produce significantly better outcomes, but when they are compared to one another they produce similar – often almost identical - effects. In the substance misuse field there is strong evidence for this general finding. Miller and Wilbourne (2002), for instance, collate RCTs in relation to alcohol misuse into a single table (the Mesa Grande) that considers the methodological strength and research evidence for different approaches. They find that there are some approaches (including brief interventions, motivational interviewing, cognitive behavioural, community reinforcement and several others) that tend to do better than 'normal service'. In contrast, there are some – such as educational, confrontative and complementary approaches that have little or no impact (or the impact tends to be negative). Similarly, in many other fields a (rather similar) list of interventions tends to work, whether this is for parenting interventions, treatment for depression, offending or an array of other problems (Roth and Fonaghy, 2005).

So far, so good for the proponent of evidence based practice. Yet here the situation becomes considerably more complex. The obvious next step is to compare credible interventions with one another. While they all may work, there are strong theoretical grounds for thinking different people would benefit from a brief intervention compared to a longer and more intensive one; or that a more directive approach might work better with some people compared to a more client-centred approach. These types of considerations led to two of the largest ever trials of any type of talking treatment (Project MATCH in the USA and the UK Alcohol Treatment Trial (UKATT)). Project MATCH involved 1,726 people across three primary conditions, comparing Motivational Enhancement Therapy (MET – a form of motivational interviewing), with CBT (over 12 sessions) with counselling to support ongoing attendance at Alcoholics Anonymous (and therefore potentially indefinite support) (Project MATCH Research Group, 1998). UKATT compared MET with a bespoke form of help focussed on the drinker and their network (Social Behaviour Network Therapy). In both instances essentially the same result was found: all conditions worked equally well with everyone (Russell et al, 2005).

This result came as something of a shock: if one is following the medical metaphor then very different types of treatment are being compared. It seems improbable that all would be equally effective. Yet perhaps this should not have come as such a surprise. Indeed, when different credible interventions are compared they have always tended to have very similar impacts, dating back at least 80 years (Rosenzweig, 1936). There is even a name for this phenomenon: the Dodo Bird effect (named after the Dodo in Alice in Wonderland who runs a race with no start or finish and when asked who has won says 'everyone has one so all shall have prizes'). Thus, the Dodo Bird effect has been found not only in alcohol treatment but also in relation to studies looking at treatment for depression (Klein et al, 2002; Schramm et al, 2011)), parenting programmes (to at least some degree, Lindsey et al, 2011), bulimia nervosa (Agras, 2000).

Yet the Dodo effect raises serious questions for researchers and proponents of EBIs. The heart of the argument of this article is that we have been too ready to focus on the intervention and insufficiently focussed on the management lessons from RCTs. For interventions to have the same effect then the logic of RCTs is that the similarities between the groups must be far more important than the difference. The question is: which similarities are the ones creating the Dodo effect?

Three main explanations have been explored to date. The first is that it is a product of the study design, and in particular elements of data collection. There is some merit in this suggestion. For instance, in their search for comprehensiveness MATCH and UKATT provided far more time with the researcher than with the counsellor in the brief intervention – and there are very strong grounds for believing that meeting a researcher who asks about your drinking in a caring and interested way and then asks to come and talk to you again in a few weeks is in itself a strong intervention that would be likely to reduce drinking (see for instance McCambridge and Strang, 2004). Yet on its own this does not seem a credible explanation: after all, the hundreds of RCTs surveyed by Miller and Wilbourne might have been expected to produce lots more Dodo findings if the research process alone explained the findings.

The same argument can be applied to the second reason given for the Dodo Bird effect, namely that the participants in the study were so highly motivated that most of the change would have occurred anyway. The strongest evidence for this is that most of the reduction in drinking occurred before the first counselling session, and therefore perhaps people who agreed to take part in an RCT in relation to problem drinking were so motivated that it did not much matter what type of help they received (Project MATCH, 1997a).

Both of these arguments do expose the tendency of RCTs to minimise or ignore the research context and the way in which it shapes the behaviour of those taking part. Yet if either were to explain the Dodo effect they would also mean that RCTs would tend to find no effect whatever was studied – and this is far from the case. Rather RCTs tend often to find similar effects when credible interventions are studied (credible interventions being those with an evidence base of effectiveness compared to normal service) and when effort is put into delivering the interventions well.

Given this one of the most popular responses to the Dodo bird effect is to emphasise the 'common elements' or 'common factors' within effective interventions. Perhaps it is because EBIs share key common features that they work in similar ways? In other words, perhaps they are not as different as they seem, and this explains their tendency to produce similar effects. This argument has received powerful advocacy in the field of social work recently, with Barth et al suggesting that the failure of social work agencies to take up evidence based interventions might be addressed by focussing instead on identifying the key features of EBIs for specific problems and then looking to develop education and support for these key features (Barth et al, 2011).

There are strong grounds for believing that effective helping shares certain common elements. Chorpita and Daleiden (2009) have carried out exhaustive reviews of what

works for whom and why. They identify many shared characteristics in effective ways of helping. If any of us thinks about what is likely to help people it is immediately apparent that there are some commonalities across different ways of working. These would include: the helper appearing to care; demonstrating an understanding of the person's point of view (empathy); engendering hope for change; building on strengths; developing a plan for change; and so on, with important variations according to the type of presenting difficulty.

Barth et al go on to use the 'common elements' (and similar 'common factors') approach to argue for a different approach to developing better practice. They argue that EBP has failed to fundamentally change social work education or practice and that this may be because tightly defined and 'manualised' interventions are not a useful way of delivering services. Instead, it is argued that a focus on common elements would allow a broader approach to developing more effective practice, which focuses on general skills and more specific skill sets for particular problems.

There are compelling arguments for such an approach. There has been a tendency – particularly in the US – to tightly define and copyright interventions. This is not conducive to developing better practice more generally. Furthermore, it seems particularly helpful to look at foundational skills within effective interventions and ensure students are skilled in these before moving on to the more specific skills associated with different interventions.

On the other hand, there are strong reasons why many EBIs keep a very tight rein on the delivery of an intervention. There is an enormous literature that indicates that once practitioners are allowed to deliver an intervention flexibly they tend to deliver it less well (specifically the issue is one of 'implementation fidelity' – whether the intervention is being delivered as it is meant to be). An oft-quoted example is Multi-Systemic Family Therapy (MSFT), which was found to not work when delivered without very rigorous quality control by the originators (Littell et al, 2005; Littell, 2006) – leading to heated debate about whether 'it' works (Henggeler et al, 2006). The answer would appear to be that it tends to work when delivered with an extremely focussed attention on delivering it to a very, very high standard. Similar effects have been found across the field of evidence based interventions, with for instance CBT for offending 'working' when delivered by some innovators in Canada but claimed to be far less successful when rolled out in the UK (Pitts, 2010), or original versions of intensive family preservation producing impressive findings but a widely rolled-out version having no impact at all (United States Department of Health and Social Security (USDHSS), 2002). These findings provide a pretty strong case for caution about 'common elements'. It is likely that there are common elements to effective working with people – but on their own these probably do not create the type of positive change that tends to be found in the best EBIs. After all, often these 'common elements' are likely to be present in normal services – or at least one would imagine they would be.

So how *do* EBIs produce the impact they do? In particular, is there something that

makes a specified intervention greater than the sum of its parts? In the next section I argue that there is, but that this is not only about the nature of the intervention as such – but rather about the *ways* in which EBIs are delivered in RCTs. Put another way the argument is that the concept of 'common elements' should look beyond the proverbial black box of the intervention, and that there are common elements in the effective leadership of practice in RCTs that produce impressive findings.

'It ain't what you do it's the way that you do it': 'common elements' in effective delivery of evidence based interventions

If one is running an RCT it is extremely important to ensure that the EBI being studied is actually delivered. The arguments outlined in this section suggest that the processes used to achieve this are likely to create some of the impact of EBIs. As such, they contribute to the existence of the 'Dodo Bird' effect (because these leadership issues are in part creating the similarities across conditions). More importantly, they suggest that 'common elements' alone are not enough; describing good practice is only one part of the equation. Delivering it is just as important – and yet it has remained remarkably opaque. It has rarely been studied: while there are thousands of RCTs looking at the way we help people there are remarkably few studying the way we get practitioners to deliver EBIs well, the organisational context or leadership required to deliver effective practice. So how do researchers achieve the effects that they do for EBIs in RCTs?

In this discussion I am going to concentrate on pragmatic trials. There are other types of trial, and in particular trials known as 'explanatory' trials. An explanatory trial is a test of whether the intervention under study works in principle, while a pragmatic trial examines whether it works in the real world. An explanatory trial is often focussed on the how and why of effectiveness, and it therefore tends to focus on ensuring that the intervention is delivered to a high standard. There is a compelling logic to this: if I want to find out whether (for instance) Motivational Interviewing (MI) works I need to make sure workers are doing MI. In particular, in general experienced and/or enthusiastic professionals are recruited. They are then trained and supervised (as discussed below). Only those who deliver MI skilfully are used to deliver the intervention. This is an effective way of ensuring that MI is delivered, but it means that there are a number of other factors that are conflated with the MI (or whatever intervention is being studied). It may be the enthusiasm, existing skills or experience of the volunteers as much as their training in the intervention. While the explanatory trial overcomes the potential problem that people may not actually be delivering the intervention, it creates a still more important problem: even if they are delivering it, we cannot be sure that 'it' is the magic ingredient. It could be other

factors that create differences between the workers using the EBI and 'normal service'.

A pragmatic trial has a different approach: it is interested in whether an intervention works in the real world. Pragmatic trials take place in normal services, with the normal workforce delivering the intervention and (usually) a comparison between the EBI and 'normal service'. Ideally, there will be a double randomization procedure in which professionals are randomized to deliver the intervention or deliver service as usual and then clients or patients are randomized. It is obvious that this is a far better test of whether something will work in the real world. The risk is that workers will not deliver the EBI – but if they do, then the only difference between the groups should be the delivering of the EBI. Outcomes for clients should therefore be attributable to this.

Yet even in pragmatic trials there are other differences that are rarely described but that are crucially important, and which have lessons for management of effective services more generally. During a trial researchers manage projects in ways that are likely to maximise the chances of a positive outcome: after all, it remains the case that there is little point in doing an RCT if the workers are not delivering the intervention that they are intended to deliver. So what do researchers tend to do to address this?

First, and probably most importantly, staff are (usually) provided with very substantial skills development packages. These would typically include some training but would then focus on intensive supervision of practice. In UKATT for instance, experienced practitioners were given initial training in MI or SBNT and were then given weekly supervision focussed on videos of actual practice. Even then it took 6 months or longer for practitioners to become skilled (Tober et al, 2008). It was after receiving this input that practitioners were able to deliver the interventions. This may have contributed to the Dodo Effect (as may several of the other points made below) as practitioners delivering both types of intervention received this skills development, but more importantly it has enormous lessons for how we develop services to deliver more effective practice and better outcomes.

To take another example, we are currently carrying out a large RCT of MI for child protection workers. This involves staff being randomized to receive MI training now or in one year. In between clients will be randomized to groups of staff who have or have not received the training. Yet it is the quality and quantity of the skills development package that is the crucial issue here. To try to get a meaningful and measurable difference we are providing:

- a 2-day workshop introducing MI,
- a follow-up day looking at MI in child protection
- each worker will receive 14 – 18 hours of further individual and small group supervision over the next 6 months
- line managers are also being trained and supported to supervise staff in developing MI skills
- workers are also getting regular emails, books, homework exercises and other input

All of this is to maximise the possibility of creating real and measurable differences in outcomes for clients. Even with this package we are not sure this will happen: it is already obvious that creating meaningful change in practitioner behaviour is very difficult. Workers have pressing cases, entrenched habits, very different skill sets and values to begin with and so on. We genuinely do not know whether all of this will make a difference (that is why we need an RCT).

Yet stepping aside from this, what are the implications for social work more generally? What is clear is that RCTs provide intensive input to get practitioners to deliver EBIs. What do we currently do? How do we teach communication skills or social work methods on our social work courses? How do we do so within local authorities? I would suggest that it is rarely if ever anything like as intensive as this.

A second feature of our current RCT – that is a common feature in RCTs more generally - is also really important: workers are having their practice recorded and are getting feedback on their levels of MI skill. This starts with actors playing clients before and after the MI training and supervision. It then continues through the study, with tapes of early sessions with families being recorded (obviously with consent). These are then independently rated and workers are given the feedback, including things they did well and areas to work on. This type of feedback – once you have the commitment of workers and provided it is offered with ongoing support – is the most powerful aid for learning imaginable. How often do we do this in social work education or in practice?

Yet – rhetorical questions aside – there are important issues here for leadership within children's services. Whether in the university or the local authority the question is – do we have the level of commitment to really improving practice that we find hardwired into most RCTs? If not, then perhaps this is the key lesson we can learn from a close study of how and why RCTs make specified interventions 'work'.

Conclusions

If it is accepted that the processes around the specific intervention are a crucially important element of making RCTs produce positive findings then there are some important implications that flow from this. First, there are important lessons for understanding evidence based interventions themselves. We have too easily accepted these as focussed on the magical black boxes that create change – yet the way in which they are delivered, the processes of skills development, the checks on 'implementation fidelity', the meetings to get staff excited and the many other elements of delivering EBIs well are considered as if they were something separate. In fact, they are crucial elements of the intervention being studied. The mistake is to think of practice and the management of practice as separate; in reality they are integrally interlinked.

This also has implications for the emerging idea of 'common elements'. This has great potential as a way of developing our thinking about the skills of social workers and others. Yet the danger is if we think of the 'common elements' as only restricted to the practice of workers. It is likely that shared conditions supporting practice are as important in creating change as the description of practice itself. Indeed, while the focus on common elements is very helpful there may be rather less tangible elements of 'evidence based interventions' that are not written in manuals but that are just as important. For example, it may be that making workers believe in a particular way of working is itself a powerful intervention: if I believe that the method I am learning works then by definition I also believe that the clients I work with can change. This may create expectancy effects that have been shown to have powerful effects in education and counselling. These beliefs also lead workers to be motivated to try to become better: believing that I am a great practitioner is likely to lead to complacency (while a lack of belief may be even more dangerous). Believing in a particular method might be expected to lead one to work hard to become skilled, and to take negative experiences as opportunities to learn and improve rather than personal criticisms. Teaching evidence based interventions involves this type of more subtle change in attitudes and beliefs that is rarely made explicit or studied.

Third, if this more complex picture of what makes EBIs work is true it is possible that this contributes to the findings from RCTs across the field of evidence based practice. Thus, interventions that worked in one setting but not another, or 'Dodo Bird' effects may be in part because of commonalities or differences in the management of practice and skills development. At the very least these processes need to be explicit and a focus for study in their own right. If the heart of an RCT is studying the impact of one 'thing' then every element of that 'thing' needs to be understood – and that is not just the practitioner's behaviour but also the measures taken to influence and improve it.

Fourth, there are crucial lessons for the way we lead services (including university programmes for social work). EBIs in RCTs achieve the improvements that they achieve not just through a quick introduction to a way of working but through a combination of thorough-going skills development and ongoing scrutiny and feedback on practice. Providing short courses in the class room or training centre, or expecting practice educators to achieve these types of training through fairly general discussion and occasional observation of practice is extremely unlikely to be effective.

To illustrate this I sometimes carry out the following thought experiment: I imagine I wanted to get research funding to evaluate whether a social work course or a local authority programme made a real difference to the skills of workers and outcomes for clients. It would be possible to carry out such research, but it would never get funded based on what we currently do – for two reasons. The first is that the input we currently provide is not based on evidence about what changes practice - indeed it is often directly against what we know about what creates change. (Bill Miller has observed that 2-day workshops in MI are a bit like 'inoculating' people

against any chance of becoming genuinely skilled. This is because they increase our self-rated level of skill and knowledge with no actual evidence of changes in skill, practice or outcomes – thus reducing our motivation to learn without increasing our ability to deliver. If this is true for MI then it is likely to be true for much of the way we teach EBIs across social work. A little bit of everything, being a 'Jack of all trades', eclectic approaches seem likely actually to be harmful rather than helpful (Miller and Mount, 2001). A second reason it would not be funded is that the existing ways we test for the skill of practitioners would not be acceptable in any imaginable research study. The current focus tends to be on giving good accounts of practice – whether these are in portfolios, essays or verbally during supervision or interviews. Workers are – to a quite extraordinary degree - judged on their ability to talk and write about their practice (Holland, 2010). If one proposed such an approach to a research funder they would not give you money because there is no evidence that talking the talk of good practice means you can actually walk the proverbial walk. This is not to say that academic assignments are not important – many of the tasks of social work require an ability to analyse, synthesise and present complex data. Yet ultimately this is only one element of the role. It does not touch on the key issue of how we talk to people and whether we help them effectively or not.

Some of the lessons for effective leadership from this discussion will be obvious, however it is worth bringing them together here. The lesson from the development of EBIs is that delivering interventions that are better than 'service as usual' is difficult – not least because most normal services are full of professionals actively doing their best to help people. So how can a leader ensure that their service does better than this? Here it is worth recapping the essential features of an EBI. The first is having a theory about the problems being dealt with and how people can be helped. A theory does not have to be complicated or academic, indeed the most effective ways of helping people are relatively simple. However, it needs to articulate how social workers help people. I am still surprised how often workers and managers struggle when asked how their service helps people. The first job of any leader in social work is to have an answer to this question.

The next element of effective leadership involves outlining in some detail what practices are consistent with this theory of helping. Again, a feature of central and local government management of children's services has always been that there is little or no attention to this crucially important question: what do managers expect workers to do to help people. In the absence of such a clear description services can all too easily become focussed on pleasing Ofsted or meeting targets – as for these there is a clear description of what is expected and how it might be achieved. Finally, leadership involves ensuring that a service allows, supports and requires workers to deliver in the way described. This is rarely simply about training. It is likely to involve a whole organisation commitment to working in specified ways. An outstanding example of this is the Reclaiming Social Work move toward systemic practice (Cross et al, 2010, Forrester et al, 2013). Taken together having a valid theory of helping,

a clear practice and a method for ensuring the delivery of such practice is likely to create services that deliver excellent practice and improved outcomes; in essence, this is what researchers do in RCTs when they develop EBIs.

This article has therefore argued that we need a more complex and nuanced approach to evidence based practice and RCTs. In particular, we need to recognise that delivering EBIs is a management issue and that this is part of what is being evaluated. Yet it is vitally important to make one further point: none of the above arguments should suggest we should not use RCTs. Rather the opposite, most of the really interesting and important data we have or are likely to have about what works in helping people has come from RCTs and related studies. It is their detailed and rigorous approach to understanding what works that allows us to explore the types of questions considered in this paper. It is also likely that the answers to some of the questions raised here will be provided through RCTs. We therefore desperately need more RCTs in social work so that we can better understand not just how to help people but also how to create organisations that support workers to do so. This is one of the key challenges for research on what works in the 21st century.

References

Adams, K.B., Matto, H.C., and LeCroy, C.W. (2009) Limitations of evidence-based practice for social work education: Unpacking the complexity. *Journal of Social Work Education*, 45, 2, 165-186

Agras, W.S., B. Timothy Walsh, B., Fairburn, C.G., Wilson. T., and Kraemer, H.C. (2000) A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Archive of General Psychiatry*, 57, 5, 459-466. doi:10.1001/archpsyc.57.5.459

Bonell, C., Fletcher, A., Morton, M., Lorenc, T. and Moore, L (2012) Realist randomised controlled trials: A new approach to evaluating complex public health interventions. *Social Science & Medicine,* 75, 2299-2306

Brown, G.W. and Harris, T. (1978) *The Social Origins of Depression in Women : A study of psychiatric disorder in women*. London: Tavistock

Barth, R.P., Lee, B.R., Lindsey, M.A., Collins, K.S., Strieder, F., Chorpita, B.F., Becker, K.D., and Sparks, J.A. (2011) Evidence based practice at a crossroads: The emergence of common elements and factors. *Research on Social Work Practice*, published online 31st May 2011, [accessed December 28th 2011]

Chorpita, B. F., and Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: Application of the distilla- tion and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology,* 77, 566-579

Cross, S., Hubbard, A., and Munro, E. (2010) Reclaiming Social Work: London Borough of Hackney, Children and Young People's Services, An Independent Evaluation Unpacking the complexity

of frontline practice. accessed online: http://www.whatdotheyknow.com/request/51132/response/130736/attach/4/1%202816227%20RSW%20FINAL%20Report%20Sept%202010.pdf

Driessen E, Hollon SD (September 2010) Cognitive behavioral therapy for mood disorders: Efficacy, moderators and mediators. *Psychiatric Clinics of North America*, 33, 3, 537-555. doi:10.1016/j.psc.2010.04.005

Forrester, D. (2012) Evaluative research. in M. Grey Ed.) *Sage Handbook of Social Work Research*. London: Sage

Henggeler , Scott W., Schoenwald, Sonja K., Borduin, Charles M., Swenson, and Cynthia C. (2006) Letter to the Editor: Methodological critique and meta-analysis as Trojan horse. *Children and Youth Services Review*, 28, 4, 447–457

Holland, S. (2010) *Child and Family Assessment in Social Work Practice*. London: Sage

Haynes, L. , Service, O., Goldacre, B., and Torgerson, D. (2011) *Test, Learn, Adapt: Developing public policy with randomised controlled trials*. Home Office Report. [Accessed online 21.01.2013 : http://www.pacts.org.uk/docs/pdf-bank/TLA-1906126.pdf]

Hofmann SG (2011). *An Introduction to Modern CBT. Psychological solutions to mental health problems.* Oxford: Wiley-Blackwell

Klein,D.N., Schwartz J.E., Santiago,N.J., Vivian,D., Vocisano,C., Castonguay,L.G., et al. (2003). 'Therapeutic alliance in depression treatment: Controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology*, 71, 997-1006

Lambert, M.J., Bergin, A.E., and S.L. Garfield (2004) Introduction and historical overview. in M.J. Lambert, A.E. Bergin, and Garfield *Handbook of Psychotherapy and Behavior Change*. (5th ed.). New York: Wiley (pp.3-15)

Lindsey, G., Strand, S., and Davis, H. (2011) A comparison of the effectiveness of three parenting programmes in improving parenting skills, parent mental-well being and children's behaviour when implemented on a large scale in community settings in 18 English local authorities: the parenting early intervention pathfinder (PEIP).*, BMC Public Health*, 11, 962. doi:10.1186/1471-2458-11-962

Littell, J.H., Popa, M., and Forsythe, B. (2005). Multisystemic therapy for social, emotional, and behavioral problems in youth aged 10-17. (Cochrane Review). *Cochrane Database of Systematic Reviews*, *4.* Chichester, UK: Wiley

Littell, J. (2006) The case for Multisystemic Therapy: Evidence or orthodoxy. *Children and Youth Services Review*, 28, 458-472

McCambridge and Strang (2004) Deterioration over time in effect of Motivational Interviewing in reducing drug consumption and related risk among young people *Addiction,* 100, 470–478

Medical Research Council (2008) *Guide to Developing and Evaluating Complex Interventions*

Miller, W.R. and Mount, K.A. (2001) A small study of training in motivational interviewing: Does one workshop change clinician and client behaviour? *Behavioural and Cognitive Psychology,* 29, 4, 457-471

Miller, W.R., Wilbourne, P.L., and Hettema, J. (2003) What works? A summary of alcohol treatment outcome research. in R.K. Hester and W.R. Miller (Eds.) *Handbook of Alcoholism Treatment Approaches: Effective alternatives.* (3rd ed.), Boston: Allyn & Bacon (pp 13-63)

Oakley, A. (2000) *Experiments in Knowing: Gender and Method in the Social Sciences*. Bristol: Polity Press

Pickett, K. and Wilson, R. (2009) *The Spirit Level: Why equality is better for everyone*. Harmondsworth: Penguin

Project MATCH Research Group (1997a) Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol*, 58, 7-29

Project MATCH Research Group (1997b) Project MATCH secondary a priori hypotheses. *Addiction*, 92, 1671-1698

Project MATCH Research Group (1998a) Matching alcoholism treatments to client heterogeneity: Treatment main effects and matching effects on drinking during treatment. *Journal of Studies on Alcohol*, 59, 631-639

Project MATCH Research Group (1998b) Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes. *Alcoholism: Clinical and Experimental Research*, 22, 1300-1311

Roth, P. and Fonaghy, P. (2005) *What works for whom? A critical review of psychotherapy research.* (2nd ed.) New York: Guilford

Rosenzweig, S. (1936) Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry,* 6, 3, 412–415

Russell, I., Orford, J., Alwyn, T., Black, R., Copello, A., Coulton, S., Farrin, A., Godfrey, C., Morton, V., Finnegan, O., Handforth, L., Middleton, W., Raistrick, D., Thistlethwaite, G., Tober, G., Westwood, A., Fryer, K., Heather, N., Hodgson, R., John, B., Kerr, C., Parrott, S., Slegg, G., Smith, M., Smith, A., Barrett, C., Kenyon, R., Chalk, P., Champney-Smith, J., McBride, A., Crome, I., Parkes, S., Emlyn-Jones, R., Fleming, A., Kahn, A., Summers, Z., and Williams, P. (2005) Effectiveness of treatment for alcohol problems: Findings of the randomised United Kingdom Alcohol Treatment Trial (UKATT). *British Medical Journal*, 331 (7516). pp. 541-544. ISSN 0959-8138

Shadish, W.R, Cook, T.D., and Campbell, D.T. (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Stamford, CT: Brooks/Cole and Wadsworth

Schramm, E., Zobel, I. Dykierek, P., Kech,S., Brakemeier, E-L., Kűlz, A., and Berger, M. (2011) Cognitive behavioral analysis system of psychotherapy versus interpersonal psychotherapy for early-onset chronic depression: A randomized pilot study. *Journal of Affective Disorders,* 129, 1, 109-116

Torgerson, D.J. and Torgerson, C.J. (2008) Designing Randomised Trials in Health, Education and the Social Sciences, An introduction. Basingstoke: Palgrave Macmillan

USDHSS (2002) *Evaluation of Family Preservation and Reunification Programs: Final report*. Department of Health and Human Services Assistant Secretary for Planning and Evaluation . aspe.hhs.gov/hsp/evalfampres94/Final/index.htm

Webb, S. (2001) Some considerations on the validity of evidence based practice in social work. *British Journal of Social Work,* 31, 1, 57-79

Wampold, B.E., Mondin, G.W., Moody, M., Stich, F., Benson, K., and Ahn, H. (1997) A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, 'All must have prizes'. *Psychological Bulletin*, 122, 203-215