British Journal of Community Justice

ISSN 1475-0279
Vol. xxxx

TIME TO RESET THE CLOCK ON THE DESIGN OF IMPACT EVALUATIONS IN CRIMINOLOGY: THE CASE FOR MULTI-METHODOLOGY DESIGNS

Stephen Morris, Andrew Smith and Chris Fox, Manchester Metropolitan University

Abstract
This paper highlights how qualitative research can enhance causal explanation in impact evaluations and provide additional causal leverage to findings from randomised experiments. We assess the extent to which randomised studies in criminology adopt mixed or multi-methodological approaches as seen in other fields such as health care, education and international development. We reviewed current practice in the design of experimental evaluations within criminology. Structured searched terms previously used to identify qualitative research components within randomised studies in health research, were used to search for evidence of mixed method design in 46 primary studies involving randomisation, published in four leading journals in criminology since 2013. Although such mixed-method randomised studies are increasingly seen in other fields such as health, education and international development, among the studies we identified in criminology and criminal justice our review reveals almost an entire absence of designs in which qualitative research is formally and explicitly integrated into study designs. We argue that randomised studies are significantly enhanced through incorporating explicit and planned mixed-method elements, and particularly qualitative research. We suggest reasons for this absence and what might be done to address it.

Key words
Randomised designs, criminology, mixed methods, qualitative research, evaluation

Introduction

In recent years there has been a welcome growth in intervention studies in criminology that adopt randomised designs. Evidence from these studies is playing an important role in advancing knowledge of 'what works' (Sullivan & Welsh, 2017). The US National Institute of Justice for example, through its Crime Solutions portal, provides details of nearly 100 programmes that have been deemed 'effective' as a result of evidence from randomised studies[1]. In Europe, the growth in the use of randomised experiments in criminology has

[1] https://www.crimesolutions.gov/Programs.aspx (accessed 16th April, 2018)

been slower but arguments in their favour no less compelling (Andersen & Hyatt, 2018). Between 1982 and 2004, Farrington & Welsh (2005) found that of 83 randomised trials in criminology worldwide, two were undertaken in Europe with not a single study found in the UK. By 2018, the UK centre for 'What Works in Crime Reduction' recorded some 24 randomised evaluations since 2009 of which 11 were ongoing[2]. Andersen & Hyatt (2018) in their survey of experiments published in the Journal of Experimental Criminology since 2010 (n=64) find that 19 per cent (n=12) were conducted in Europe. They also note that 8 per cent of the panel presentations at the 2014 Conference of the European Society of Criminology involved a study with a randomised design (Andersen & Hyatt, 2018, 5).

We argue that the growth in randomised intervention studies in criminology is a very welcome development (Farrington, 1983, 2003; Weisburd, Lum, & Petrosino, 2001) and that the many examples of successful randomised studies in criminology attest to the previously perceived practical and ethical barriers being readily surmountable. However, the central contention of this paper, is that usefulness and insights from randomised intervention studies are significantly enhanced through incorporating explicit and planned mixed method elements (Campbell, 1987), and particularly qualitative research. We provide evidence, however, that such mixed method intervention study designs are rare in criminology and highlight how qualitative research can promote better causal explanation. In part this paper is motivated by experience in other fields. In international development, for example, what are variously termed mixed-method randomised controlled trials or RCT+ designs have been discussed and implemented (Bamberger, Tarsilla, & Hesse-Biber, 2016; White, 2013). In health research there has been a long tradition of promoting mixed method intervention studies (Boeije, Drabble, & O'Cathain, 2015; Hansen & Jones, 2017; R. B. Johnson & Schoonenboom, 2016; Moore et al., 2015; Oakley, Strange, Bonell, Allen, & Stephenson, 2006). In education, the growing use of randomised designs has been accompanied by an increased emphasis on studies that combine randomisation with mixed method implementation process evaluation (Humphrey et al., 2016; Lendrum & Humphrey, 2012). Whilst there are calls consistent with such designs in criminology (see discussions around the EMMIE framework, Johnson, Tilley, & Bowers, (2015)), we will show that there is a dearth of such mixed method designs in the academic literature, and this in our view reveals a less well developed and possibly more polarised debate around mixed methods in the sector.

First, this paper outlines the role qualitative research can play in causal inference. This discussion is addressed to those from a quantitative background with a more sceptical position toward what qualitative might offer. In opposition to the notion of a fundamental incompatibility between qualitative and quantitative research, we outline the different ways researchers have thought about how qualitative research might be integrated with randomised designs and acknowledge the importance of philosophical pragmatism as providing an epistemological unpinning for this task. We further advance our discussion by providing three examples (among the many we could have called upon) from education, health and international development of the successful integration of qualitative research

[2] Site access on 1st February, 2018 – see http://whatworks.college.police.uk/Research/Research-Map/Pages/Research-Map.aspx

and randomisation. These examples support our claims that such mixed method interventions are achievable, insightful, and conceptually coherent.

Having established the benefits of mixed method implementation study designs, we then move on to ask how far researchers in criminology are using formal or planned mixed method randomised designs. We examine 46 recently published randomised experiments in criminology, drawn from four leading journals in criminology. We find that, in short, there is little evidence of mixed methodology or multi-method study designs in the sector despite the arguments in their favour. As a result, the final section of this paper provides and assessment of why this might be and what might be done to encourage greater use of qualitative research, as well as mixed methods more broadly, within experimental criminology. Our findings raise important questions that need to be addressed by criminologists against a back-drop in growth and importance of evidence from such study designs.

The limitations of randomised trials

Randomised experiments provide estimates of average treatment effects through comparing differences in mean outcomes between groups formed at random, where groups are exposed to different levels of an intervention (often including a control condition). If the experiment is performed correctly, estimates of average treatment effects are said to be unbiased or internally valid (Shadish et al., 2002). Furthermore, in most circumstances statistical inference is relatively straightforward (Burtless, 1995). To have an unbiased estimate of the effect of an intervention is very useful but raises further questions. These questions centre on the provision of an explanation(s) concerning how the observed effects came about and under what conditions. Where studies fail to find an effect, further questions will naturally arise as to why. The more complex the intervention, the less self-evident the answers to these types of questions. In the case of social interventions, such as those found in criminology and elsewhere in the social sciences, the means through which effects are generated can't be taken for granted and assumed to be straightforward and/or self-evident.

At least three broad types of question remain unanswered in results from standard experimental analysis: 1) questions relating to the processes or mechanisms that generate or lead to the observed effects; 2) questions addressing the factors present in the context in which the study took place that may enable or constrain the operation of the intervention; and 3) questions connected with the implementation of the intervention and how far implementation fidelity was achieved (Moore et al., 2015).

In order to address questions relating to causal processes or mechanisms, researchers have resorted to estimating structural statistical models and forms of mediator and moderator analysis (Imai, Keele, Tingley, & Yamamoto, 2011). Sample sizes in many randomised interventions, however, tend to render moderator analyses underpowered (Frazier, Tix, & Barron, 2004). Moreover, the assumptions required for the valid analysis of mediators are often highly restrictive (Keele, 2015; Suzuki & VanderWeele, 2018). These limitations have been a significant motivating factor leading many researchers to the use of mixed methods and qualitative research to explicate causal processes or mechanisms (Bamberger, 2015;

Bonell, Fletcher, Morton, Lorenc, & Moore, 2012; Jamal et al., 2015; White, 2013).

This takes us to the second type of question that remains unanswered in standard experiments: questions about context. In considering the importance of context, the varied nature of results obtained from studies of ostensibly the same intervention suggest that interventions and thus findings from randomised studies are more dependent upon context than is sometimes acknowledged (Sampson, Winship, & Knight, 2013) . For example, Berman & Fox (2016) document examples where drug courts, generally understood to be effective for adults, albeit with quite heterogenous effects (Mitchell, Wilson, Eggers, & MacKenzie, 2012), have been shown to have failed for reasons that relate to the 'political' and institutional circumstances in which they were operationalised; features of context that are particularly amenable to qualitative investigation. Context can refer to differences in 'treated' populations but also geography, historical trends, political institutions, physical environment, available resources, alternative opportunities, and so on. These are influences that Cartwright & Hardie (2012) refer to as supporting factors. Knowledge of such factors prevailing within the context in which a specific study has been conducted is important in promoting understanding of the extent to which results might hold elsewhere and indeed for meta-analysis and research synthesis. Qualitative approaches have been advanced as means of understanding and incorporating knowledge of wider contextual factors into randomised designs (Bonell et al., 2012; White, 2009).

Understanding implementation, and fidelity to intended design of an intervention, has been a key concern in evaluations using randomised designs for many years and is the third type of question that standard experiments struggle to address. In order to interpret results from randomised studies, knowledge of the nature of the intervention, its implementation, fidelity to intervention design, and nature of exposure is essential. In the field of health services research, process evaluations have been widely integrated into randomised studies (Moore et al., 2015; Oakley et al., 2006). More recently in education, formal methods of implementation process evaluation (IPE) have been proposed, drawing on the development of process evaluation approaches and techniques in health research and elsewhere (Humphrey et al., 2016). Both process evaluation and implementation process evaluation place great emphasis on qualitative methods such as depth interviews, ethnographic methods and focus groups.

An attempt has been made, when reviewing 'what works' evidence, within criminology to address these issues in the form of the EMMIE framework (S. D. Johnson et al., 2015). The EMMIE framework requires that reviews of studies identify the effect (and costs) of an intervention but also the causal mechanisms through which it works, the factors that moderate its impact, and the issues that may hinder implementation (Johnson et al., 2015). By taking into account the mechanisms and moderators (that is context) the EMMIE framework accords very much with the design of mixed method intervention studies in other fields. However, it is fundamentally a review framework and therefore relies on primary studies producing the types of information and evidence upon which the framework draws. But, as we will argue, there appears little evidence that primary studies, as published in the academic literature in criminology, are in a position to provide these forms of evidence.

Qualitative research, randomised experiments and the 'paradigm' divide

The integration of qualitative research and randomised experiments as a form of mixed methods research has been subject to criticism, particularly from those who hold to the fundamental incompatibility of quantitative and qualitative research (Guba & Lincoln, 1994). This critique stems from the notion that qualitative and quantitative research are derived from separate and fundamentally incompatible ontological positions, 'world views' or 'paradigms'. Qualitative research is often seen as inextricably linked to 'interpretivism' and the 'constructivist' paradigm, whilst quantitative research and experiments to positivism or post-positivism (Wiggins, 2011).

By way of contrast, increasing numbers of researchers argue that there is no essential link between method and paradigm (Biesta, 2010; Greene, Caracelli, & Graham, 1989; Symonds & Gorard, 2010). Furthermore, for those combining qualitative and quantitative approaches, 'pragmatism' as philosophical perspective, provides a sound footing for such endeavours (Biesta, 2010; R. B. Johnson, Onwuegbuzie, & Turner, 2007; Morgan, 2014). Separately, researchers operating in the 'realist' tradition also see no barrier to combining quantitative and qualitative approaches within single studies (Bonell et al., 2012; Pawson & Tilley, 1997; Porter, McConnell, & Reid, 2017), though the place of experimentation within 'realism' is still hotly contested (Belle et al., 2016). What these debates suggest is that the integration of qualitative approaches within randomised studies need not be seen as attempting to reconcile the fundamentally irreconcilable.

As we have noted, discussion of the role qualitative research in enhancing causal explanation, in concert with randomisation, has been prominent in methodological debates within health. This emphasis is in part a response to the growing use of randomised designs beyond evaluation of pharmacological treatments to study interventions in the form of complex human services (Moore et al., 2015; Oakley et al., 2006; Popay & Williams, 1998). A similar set of concerns is also discernible in international development, where interventions are often evaluated within complex and unstable environments (Bamberger et al., 2016; Jimenez et al., 2018; White, 2013). Given the nature of interventions that involve multiple human interactions between various law enforcement agencies, their practitioners and offenders, it would appear these discussions are as relevant in criminology as elsewhere.

Many researchers intuitively find combining quantitative and qualitative research appealing. It offers the prospect of addressing a wider set of questions than mono-method-study-designs within a single study. However, qualitative researchers have wished to avoid their research playing a subservient role to quantitative research. Popay & Williams (1998) point to a distinction between an 'enhancement' model for this relationship and a 'difference' model. In the case of the former, qualitative research addresses very specific issues relating to the improvement or enhancement of experimental findings. This is not unlike the multi-method approach proposed by Seawright (2016). According Popay & Williams (1998, page 34), however, 'qualitative research has potentially got a much greater contribution to make than the 'enhancement' model suggests'. The 'difference model' advances the notion that qualitative research provides understanding of practitioner (police

officers, probation staff, prison staff, etc.) and participant's subjective experiences and explores the meaning given to a programme or intervention by both practitioners and participants, as well as addressing unquantifiable aspects of context (e.g. institutional or organisational culture, etc.) that can determine success or otherwise of an intervention.

The distinction between 'enhancement' and 'difference' is similar to the contrast Greene, Benjamin, & Goodyear (2001) made between 'pragmatic' and 'dialectical' mixed methods, where it could be argued that the former has much in common with the 'enhancement' model, whilst the latter the 'difference' model. Similarly, Hesse-Biber, (2012) suggests that qualitative research and experimental design can be integrated either within the 'context of discovery' or 'context of justification'. Bringing qualitative perspectives to bear within the 'context of discovery' suggests an inherently multi-methodological study design, in which again there is a parity between qualitative and quantitative elements; where the focus can shift back and forth between 'objective' and 'subjective' perspectives. By way of contrast, qualitative research within the 'context of justification' refers to methods deployed during and at the end of the experimental research, with the purpose of expanding upon and augmenting experimental findings.

What this discussion suggests is that intervention study designs pay attention to the subjective perspective of both practitioners and participants, if causal explanation is to be maximized. Though not without significant challenges, those designing mixed or multi-methods intervention studies need to consider a far more fundamental mixing, or as Hesse-Biber (2012) terms weaving, of different perspectives.

Qualitative research and randomised design within intervention studies – the benefits of integration

In order to advance our argument for mixed method intervention studies in criminology, we provide examples of three trials that highlight ways in which qualitative methods specifically, and mixed methods more broadly, can be integrated with randomised designs to address questions of causal processes or mechanisms, context, and implementation. Our examples come from health, international development and education. Jamison, Karlan, & Raffler (2013) describe a mixed method randomised intervention study to evaluate an interactive text messaging platform in Uganda, that through providing access to information on sexual health aimed to improve knowledge and reduce the incidence of risky sexual behaviour in relation to HIV infection. A cluster randomised controlled involved the allocation of 60 villages to intervention or control conditions. Within intervention villages, villagers were exposure to a 'targeted high-intensity marketing campaign by a professional marketing firm' (Jamison et al., 2013, page 6) to raise the rate at which the service was accessed over a three month period – a so called 'encouragement design'. The researchers conducted eight qualitative focus groups and 39 in-depth interviews after follow-up outcome data had been collected and analysed.

Results from the experimental element of this study were disappointing and somewhat surprising. They indicated that the incidence of risky behaviour in intervention villages was actually higher than in control villages. Looking at implementation, qualitative research showed that reminders about the service were not sustained over a sufficient period of

time, reducing exposure to the service to below that expected and that the interactive topic search facility often did not work effectively. Furthermore, respondents in qualitative in-depth interviews suggested that despite accessing better information on sexual health, they did not have the resources available to act on the information. Getting treatment and access to testing was often costly and beyond the means of low-income families in the study sample.

The qualitative work also directly investigated possible explanations for increased risky behaviour in intervention villages. Researchers identified what they termed a 'sexual sorting' mechanism whereby married women gained important information about infection and 'insisted that their husbands be faithful and go for testing with them' (Jamison et al., 2013, page 15). Where husbands failed to comply with these requests, wives withheld sex and qualitative evidence suggests husbands in some cases looked elsewhere for sex, leading to an increase in promiscuity and risky behaviour. This study provides an excellent example of how qualitative evidence can not only identify difficulties in implementation and important contextual factors but also help explain experimental findings. These are vital insights that in the absence of mixed methods and particularly planned qualitative work, would have remained unknown.

Our second example comes from qualitative work incorporated into a randomised intervention study of the Accelerated Research (AR) programme in education. AR is remedial catch-up intervention delivered in schools in order to support pupils that are behind in their reading during early secondary education. The intervention consists of an online package that matches pupils by ability and sets quizzes related to chosen reading material to assess progress. Gorard & See (2016) undertook a randomised intervention study to examine the effectiveness of AR that also comprised qualitative elements, including interviews (in-depth interviews) with teachers, administrators, students (Siddiqui, Gorard, & See, 2018). Data collection took place at the start of the trial and then again toward its end.

In terms of reading attainment, those in the intervention group were found to perform better than the control group. Qualitative research revealed that the apparent success of AR was founded upon the matching up of reading material not only to ability but also interest (Gorard & See, 2016). Qualitative research also identified contextual factors whose absence or presence was central to the success of the intervention. Most notably these were around library facilities in schools and availability of reading materials to which pupils had been 'matched'. Schools also needed to maintain a subscription to the online AR service (Gorard & See, 2016). Again, without planned qualitative research the importance of these factors could have easily been over-looked.

Our third example is a complex multi-country randomised study in health care. It aimed to test two interventions, compared to control, that promoted 'facilitation' in relation to incontinence care in care homes in four European countries (Rycroft-Malone et al., 2018). Facilitation is a concept developed in implementation science from health care that aims to move beyond the adoption of evidence based interventions through simply disseminating knowledge, instead using a multifaceted approach that involves skilled individuals who

enable others, 'through a range of intervention components and approaches, to address the challenges in implementing evidence-based care guidelines within the primary care settings' (Baskerville, Liddy, & Hogg, 2012, page 63). This study is of interest because it integrates qualitative research into the experimental design within a 'realist RCT' framework, originally proposed by Bonell et al. (2012). In simple terms, realist RCTs can be considered a more developed form of mixed method intervention study. In this instance 'the purpose was to provide a theory-driven explanation of the response to facilitation interventions as they were being implemented' (Rycroft-Malone et al., 2018, page 2).

Considerable attention was paid in this study to the development of theoretical conjectures around how the intervention could bring about change and under what circumstances (Rycroft-Malone et al., 2018). The development of theory relied in part on the collection of qualitative information through workshops with stakeholders. A range of qualitative methods were also deployed in order to test theory. Overall, the experimental element of the study showed no evidence that the two facilitation approaches under investigation led to improvements (Seers et al., 2018). The qualitative research revealed that generally facilitators in response to the training they received sometimes struggled to see how they could apply what they learnt in practice. The intervention involved facilitators operating in care homes and their confidence in delivering the intervention as intended was found to be important, along with their authority and seniority within the care home setting. This in turn was reflected in the priority given to the intervention by the care home itself. Where the two factors combined positively a process of learning and adaptation took place. National policies and guidance was also found to be significant in two of the countries involved (Rycroft-Malone et al., 2018). These factors identified in the qualitative research were found to be equally important in both intervention arms of the study. Overall, in both active arms (as opposed to control) of the study the intervention itself was found to be diluted and not delivered as intended (Rycroft-Malone et al., 2018).

These three studies yield ample evidence of the potential to be gained from the integration of qualitative approaches with randomised study designs. Moreover, their existence and success acts to refute claims that integration of mixed methods is neither practical nor intellectually coherent. Furthermore, we did not encounter any difficulties in finding examples of effective and coherent integration of qualitative methods and randomised designs in international development, health and education. There were many more examples. As we will see, it has proved much harder to find such examples in criminology, particularly within the academic literature.

Mixed or multi-method intervention studies in criminology – an assessment of the academic literature

The discussion thus far, whilst acknowledging the importance of randomised studies has highlighted areas where such studies have limitations. It makes the case that such limitations can, to some extent, be addressed through the mixing or integration of experimental designs with qualitative research. In this section, we explore the extent to which such mixed method designs can be found in experimental criminology.

We reviewed studies published between 2013 and the first quarter of 2018 in four leading criminology journals (Criminology, The British Journal of Criminology, The Journal of

Experimental Criminology, The Journal of Quantitative Criminology). These journals were selected as those likely to publish high quality, exemplar randomised studies, and yield a sample representing contemporary best practice in academic criminology. Within these journals over the relevant period of time, we identified studies that described the use of randomisation. We conducted an initial hand-search of these journals for primary studies, and excluded systematic reviews and rapid evidence assessments, replication studies, and any purely methodological papers about experiments and randomisation. In addition, a number of lab experiments were identified and excluded.

Our initial search identified 46 articles that met our criteria[3], which we then analysed using full-text searches to determine whether the study made mention of qualitative methods used before, during or after the reported randomised experiment. To focus our full-text searches we adapted the search terms used by O'Cathain et al. (2013). O'Cathain et al. 2013 looked at how qualitative research was used alongside randomised study designs in health. The search terms used by O'Caithan and colleagues went through an exhaustive process of development to ensure that through their application the full breadth of qualitative approaches that might be used in conjunction with randomised designs might be identified.

Using these search terms, we find that only four of the studies we identified described the planned use of qualitative methods. Three of the 46 articles described the use of a process evaluation using methods such as feedback meetings and the analysis of qualitative data. However, from the details provided in the articles it was difficult to ascertain the exact qualitative methods used during the process evaluations, and whether process evaluations also made use of quantitative data. Four studies were not counted as examples of qualitative process evaluation as they employed quantitative data alone for this phase of the research. One study noted that preliminary qualitative fieldwork took place before the randomised experiment, although its purpose was not clear, and its findings were not reported in the article. In addition, some studies did report findings of what might be described as rather informal, piecemeal or opportunistic forms of data collection that might be construed as qualitative in the very broadest sense (for example informal discussions with law enforcement officers). These were not included in the count due to their informal, ad hoc and non-systematic nature (often referred to vaguely as 'observations' with no further elucidation of the techniques or approach involved), and due to the fact that in a number of cases such informal encounters seem not to have been conducted by researchers with any training in qualitative methods.

In addition, we hand searched the bibliographies of each of the 46 papers uncovered to determine whether any additional, supplementary qualitative studies linked to the main study were cited therein. We looked for other studies published by the named authors of the main study and studies mentioning the same or a similar intervention published within a few years of the main study could be found. This process revealed no further qualitative outputs associated with the papers included in our review.

Despite the paucity of qualitative methods used alongside randomised experiments in the

[3] A full list of identified studies is available from the authors on request.

46 articles, four of the articles suggested qualitative research as a logical next step in order to enhance understanding of the experimental findings (Groff et al., 2015), and ethnographic methods, such as those proposed by (Sherman & Strang, 2004) were cited in particular.

It might be claimed that the sample of studies we have obtained is not representative of the practice of experimental evaluation in criminology, and that there is a far wider use of planned qualitative research than our results suggest. It could be, for example, that many excellent examples of mixed method randomised intervention studies are not published in the academic literature. Or that the journals we have chosen publish studies that are unrepresentative of wider practice. Whilst we cannot categorically rule out a rogue sample, if qualitative research is widely integrated into experimental studies in criminology, then one could only have arrived at our results based on an extreme and rare sample and it seems unlikely to us that we would encounter such an unrepresentative set of studies in four leading quantitative criminology journals.

Second, an objection might be that although the studies we have found did not incorporate qualitative research within them, they may have been preceded by related studies involving qualitative research or followed by connected studies adopting qualitative research elements. If the former is the case, then there is little evidence for the existence of these studies and they do not appear to be referenced in the bibliographies of the papers we have found. It could be that such studies do still exist, even though no mention of them is made. But if true, this raises the obvious question as to why no mention of them is made; and by extension why they are deemed irrelevant within the context of the experimental findings? We cannot rule out that the experimental studies in our sample were followed by related qualitative studies. However, this also begs the question as to why plans for further qualitative work are not mentioned anywhere in the studies we have found.

Discussion: Why not qualitative research?

We have argued that studies that seek to identify the causal effects of interventions in criminology need to focus not just on causal description, but equally causal explanation. Furthermore we have stressed the role that qualitative research can play in enhancing causal explanation as part of explicit and planned mixed-method randomised intervention studies. We are not the first to argue for the integration of qualitative approaches within experimental studies (Angel et al., 2014; Sherman & Strang, 2004). However, our review of over 40 recent randomised intervention studies in criminology suggests that despite these calls the practice of integrating qualitative research formally within studies that involve randomisation is rare in criminology, at least in the academic literature. In this final section of our paper we provide some possible explanations for this finding.

There are many possible reasons for the lack of mixed methodology or multi-method studies incorporating qualitative research and randomisation. The first and probably most obvious reason for our findings is the scepticism toward qualitative research often found among advocates of experimental approaches. Concerns have been expressed over the quality of qualitative research (for example King, Keohane, & Verba, 1994) and what can seem to many as its speculative nature. The hunt for 'mechanisms' that often accompanies

qualitative inquiry has also been questioned with the implication that such endeavours are often more fraught than is sometimes acknowledged (Gerring, 2010). Nonetheless, despite these challenges, we maintain that many of the questions that arise in response to findings from randomised interventions studies that are left effectively 'hanging in the air', can be realistically addressed using qualitative methods. Moreover, without recourse to mixed or multi-methodological designs, it seems to us that mono-method randomised studies can only play a limited role in the evaluation of complex interventions.

Another possible explanation relates to the nature of criminology as a discipline. In fact some authors suggest that criminology is not a coherent discipline but a 'rendezvous' subject (Downes, 1983) where different disciplines and people converge around a common subject area (Rock, 2017). This means that criminology can be characterized by an eclectic set of methodological concerns, theoretical perspectives and normative frameworks (Liebling, Maruna, & McAra, 2017). As a field of study criminology has been stocked by 'psychiatrists, psychologists, sociologists, historians, social and physical anthropologists, lawyers, statisticians, journalists, and documentary film makers, and they have tended to talk about rather different things and share rather different traditions, audiences, styles and preoccupations. The result has been something of a cacophony in which they have not always been able to listen to, or understand, one another' (Rock, 2017, page 22). Within this cacophony, perhaps it is not a surprise that the methodological rapprochement required for the development of mixed or multi-method trials has been slow in coming. Rapid growth in the sector has, perhaps, exacerbated this tendency. Liebling et al. (2017) argue that rapid expansion of British criminology has opened up the discipline but also led to greater specialization and perhaps methodological fragmentation. In this context, criminology seems to have been more riven by the quantitative/qualitative paradigm wars than other fields of enquiry and these disputes have been slower to subside. It is surely relevant that the scientific realist approach to evaluation, with its strong critique of randomised designs emerged initially in the criminological literature in the UK (Pawson & Tilley, 1994). Although both quantitative and qualitative data have a place in realist evaluation and the approach is 'method neutral' (Marchal, van Belle, van Olmen, Hoeree, & Kegels, 2012) in practice there is more emphasis in realist evaluation on qualitative insights. The fairly recent exchange between Tilley and Sherman, in which the latter presents a strident case for randomised experiments (Sherman, 2009) and Tilley (2009) argues, to the contrary, that experiments are flawed, is a demonstration that the paradigm wars have possibly yet to wane in criminology. It doesn't, however, seem sufficient to argue that the fragmented nature of criminology and the legacy of the paradigm wars mean that those supportive of randomised experiments in criminology have less access to mixed methods expertise than their peers in other disciplines. Perhaps therefore there are other forces in play?

In the US, where many of the top-rated journals in criminology are located there is a perceived bias towards quantitative research (Andersen & Hyatt, 2018; Maruna, 2010; Sampson, 2013). Wheeldon (2012) in a review of 95 published articles in Criminology between 2010 and 2012 found that less than a tenth used qualitative or mixed-method approaches and Dooley (2010) found that only 11 percent of articles appearing in the discipline's top peer-reviewed journals between 1951 and 2008 were qualitative.

Wheeldon, Heidt, & Dooley (2014) cite these studies in support of the proposition that (in the US at least), integrating quantitative and qualitative methods is challenging if these studies are harder to publish (Wheeldon et al., 2014, page 118; and see Maruna, 2010). A related issue might be that mixed method studies do not lend themselves to easy reporting in 8,000 word academic journal articles (Maruna, 2010).

Other forces that keep quantitative and qualitative methods at arms-length may come from outside academia (Lum & Yang, 2005). Liebling et al. note that in the UK, criminology as a discipline, has expanded rapidly and now entered its third generation. As it has done so there are new pressures on academic criminologists to chase competitive contract research. Many of these funding opportunities emanate from government or from providers of public, voluntary sector and, increasingly private criminal justice service providers. Against a backdrop of tight public spending settlements, for such commissioners' evidence that can be integrated into formal cost-benefit or value for money assessments is a priority. Thus, contract budgets might tend to disproportionately favor those elements of the research that meet these needs; that is quantitative elements. Qualitative research, if it is commissioned at all as part of a wider study often tends to be relatively under-resourced and as a result poor quality. This can lead to a potential discrediting of mixed or multi methodological designs. In such an environment, quantitative evaluations and the economic evaluations are at least in some quarters, seen as dominating.

Conclusions: Encouraging mixed methods in randomised intervention studies

We tentatively suggest some steps that might be taken toward encouraging qualitative research and mixed methods in randomised intervention studies more broadly. First, if we accept the advantages of mixed method intervention studies, this may involve revisiting the way research methods are taught in universities (Wheeldon et al., 2014)but also that they are understood and appreciated by research commissioners, to say nothing of journal editors and reviewers. This is clearly a long-term project, but a useful starting point is suggested by Maruna (2010) who emphasises that the difference between quantitative and qualitative methods in criminology is overstated. The key then is greater dialogue between criminologists working within the quantitative and qualitative traditions. Secondly, we suggest that while criminology has historically drawn on many disciplines (Rock, 2017) as the discipline has grown it has become more inward looking and criminologists have possibly taken less notice of the development of multi-method, randomised intervention studies in other fields. There is now a need for the learning in other sectors to be considered and incorporated into the next generation of randomised studies. Thirdly, many of the studies discussed in the review provided above appear to be have been staffed by researchers with little understanding of qualitative research, the benefits it might bring and the types of methods that might be deployed. The make-up of research teams appears, at least with regard to skills and expertise, to be relatively homogenous. We would argue that a conscious effort should be made by principal investigators to build research teams that reflect a diversity of methodological perspectives and specifically seek to recruit qualitative researchers into their teams.

Finally, experience from other fields of study, particularly health, international development

and more recently education, suggests that the use of programme theories (theories of change, etc.) (Funnell & Rogers, 2011) can be helpful in terms of framing the integration of different methodological contributions to identifying and explaining causal effects. To some extent, developments in health intervention studies where 'realist RCTs' have been advanced reflect this (Bonell, Warren, Fletcher, & Viner, 2016). Furthermore, such approaches drawing heavily on qualitative methods, can provide a helpful interpretive framework for experimental findings (Bamberger, 2015). Our final suggestion is therefore that experiments in criminology consider making greater use of programme theories as a way of bringing to bear qualitative perspectives but also for the benefits they can bring in their own right.

References

Andersen, S. N., & Hyatt, J. (2018). Randomised experiments in Scandinavian criminal justice: Reviewing the past and looking to the future. *European Journal of Criminology*, 1477370818788015. https://doi.org/10.1177/1477370818788015

Angel, C. M., Sherman, L. W., Strang, H., Ariel, B., Bennett, S., Inkpen, N., … Richmond, T. S. (2014). Short-term effects of restorative justice conferences on post-traumatic stress symptoms among robbery and burglary victims: a randomised controlled trial. *Journal of Experimental Criminology*, 10(3), 291–307. https://doi.org/10.1007/s11292-014-9200-0

Bamberger, M. (2015). Innovations in the use of mixed methods in real-world evaluation. *Journal of Development Effectiveness*, 7(3), 317–326.

Bamberger, M., Tarsilla, M., & Hesse-Biber, S. (2016). Why so many "rigorous" evaluations fail to identify unintended consequences of development programs: How mixed methods can contribute. E*valuation and Program Planning*, 55, 155–162.

Baskerville, N. B., Liddy, C., & Hogg, W. (2012). Systematic review and meta-analysis of practice facilitation within primary care settings. *The Annals of Family Medicine*, 10(1), 63–74.

Belle, S., Wong, G., Westhorp, G., Pearson, M., Emmel, N., Manzano, A., & Marchal, B. (2016). Can ``realist'' randomised controlled trials be genuinely realist? *Trials*, 17. https://doi.org/10.1186/s13063-016-1407-0

Berman, G., & Fox, A. (2016). *Trial and error in criminal justice reform: Learning from failure* (Revised). Washington DC: Urban Institute Press.

Biesta, G. (2010). Pragmatism and the Philosophical Foundations of Mixed Methods Research. In A. Tashakkori & C. Teddlie (Eds.), SAGE Handbook of Mixed Methods in Social & Behavioral Research. Thousand Oaks, California: SAGE Publications, Inc. https://doi.org/10.4135/9781506335193

Boeije, H. R., Drabble, S. J., & O'Cathain, A. (2015). Methodological challenges of mixed methods intervention evaluations. *Methodology*, 11(4), 119–125.

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

Bonell, C., Warren, E., Fletcher, A., & Viner, R. (2016). Realist trials and the testing of context-mechanism-outcome configurations: a response to Van Belle et al. *Trials*, 17(1), 1–5. https://doi.org/10.1186/s13063-016-1613-9

Burtless, G. (1995). The case for randomised field trials in economic and policy research. *Journal of Economic Perspectives*, 9(2), 63–84.

Campbell, D. T. (1987). Guidelines for monitoring the scientific competence of preventive intervention research centers: An exercise in the sociology of scientific validity. Knowledge, *Science Communication*, 8(3), 389–430.

Cartwright, N., & Hardie, J. (2012). Evidence-based policy: A practical guide to doing it better. Oxford: Oxford University Press.

Dooley, B. (2010). Whither criminology? On the state of criminology's paradigm. PhD Thesis: University of Missouri, St Louis.

Downes, D. (1983). *Law and order: theft of an issue*. London: Fabian Society.

Farrington, D. P. (1983). Randomised Experiments on Crime and Justice. *Crime and Justice*, 4, 257–308. Retrieved from http://www.jstor.org/stable/1147511

Farrington, D. P. (2003). Methodological quality standards for evaluation research. The *Annals of the American Academy of Political and Social Science*, 587(1), 49–68.

Farrington, D. P., & Welsh, B. C. (2005). Randomised experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology*, 1(1), 9–38. https://doi.org/10.1007/s11292-004-6460-0

Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, 51(1), 115.

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. San Francisco, CA: Jossey-Bass.

Gerring, J. (2010). Causal mechanisms: Yes, but…. *Comparative Political Studies*, 43(11), 1499–1526. https://doi.org/10.1177/0010414010376911

Gorard, S., & See, B. H. (2016). Accelerated Reader as a literacy catch-up intervention during primary to secondary school transition phase AU - Siddiqui, Nadia. *Educational Review*, 68(2), 139–154. https://doi.org/10.1080/00131911.2015.1067883

Greene, J. C., Benjamin, L., & Goodyear, L. (2001). The merits of mixing methods in evaluation. *Evaluation*, 7(1), 25–44.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3), 255–274.

Groff, E. R., Ratcliffe, J. H., Haberman, C. P., Sorg, E. T., Joyce, N. M., & Taylor, R. B. (2015). Does what police do at hot spots matter? The Philadelphia policing tactics experiment. *Criminology*, 53(1), 23–53.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. Handbook of Qualitative Research, 2(163–194), 105.

Hansen, A. B. G., & Jones, A. (2017). Advancing 'real-world'trials that take account of social context and human volition. *Trials*, 18(1), 531.

Hesse-Biber, S. (2012). Weaving a multimethodology and mixed methods praxis into randomised control trials to enhance credibility. *Qualitative Inquiry*, 18(10), 876–889.

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R., & Kerr, K. (2016). *Implementation and process evaluation (IPE) for interventions in educational settings: An introductory handbook*. London: Education Endowment Foundation.

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, 105(4), 765–789.

Jamal, F., Fletcher, A., Shackleton, A., Elbourne, D., Viner, R., & Bonell, C. (2015). The three stages of building and testing mid-level theories in a realist RCT: a case-example. Trials, 16. https://doi.org/10.1186/s13063-015-0980-y

Jamison, J. C., Karlan, D., & Raffler, P. (2013). Mixed method evaluation of a passive mHealth sexual information texting service in Uganda. National Bureau of Economic Research.

Jimenez, E., Waddington, H., Goel, N., Prost, A., Pullin, H., White, H., Lahiri, S. & Narain, A. (2018). Mixing and Matching: Using Qualitative Methods to Improve Quantitative Impact Evaluations (IEs) and Systematic Reviews (SRs) of Development Outcomes. CEDIL Inception Paper 5: London.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, 1(2), 112–133. https://doi.org/10.1177/1558689806298224

Johnson, R. B., & Schoonenboom, J. (2016). Adding qualitative and mixed methods research to health intervention studies: Interacting with differences. *Qualitative Health Research*, 26(5), 587–602.

Johnson, S. D., Tilley, N., & Bowers, K. J. (2015). Introducing EMMIE: an evidence rating scale to encourage mixed-method crime prevention synthesis reviews. *Journal of Experimental Criminology,* 11(3), 459–473. https://doi.org/10.1007/s11292-015-9238-7

Keele, L. (2015). Causal Mediation Analysis: Warning! Assumptions Ahead. *American Journal of Evaluation*, 36(4), 500–513.

King, G., Keohane, R. O., & Verba, S. (1994). Designing social inquiry: Scientific inference in qualitative research. Princeton, New Jersey: Princeton University Press.

Lendrum, A., & Humphrey, N. (2012). The importance of studying the implementation of interventions in school settings. *Oxford Review of Education*, 38(5), 635–652.

Liebling, A., Maruna, S., & McAra, L. (2017). Introduction: The new vision. In A. Liebling & S. Maruna (Eds.), Oxford Handbook of Criminology, 21 (6th ed.). Oxford University Press.

Lum, C., & Yang, S.-M. (2005). Why do evaluation researchers in crime and justice choose non-experimental methods? *Journal of Experimental Criminology*, 1(2), 191–213. https://doi.org/10.1007/s11292-005-1619-x

Marchal, B., van Belle, S., van Olmen, J., Hoeree, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*. https://doi.org/10.1177/1356389012442444

Maruna, S. (2010). Mixed methods research in criminology: why not go both ways? In A. R. Piquero & D. Weisburd (Eds.), Handbook of quantitative criminology. New York: Springer-Verlag.

Mitchell, O., Wilson, D. B., Eggers, A., & MacKenzie, D. L. (2012). Assessing the effectiveness of drug courts on recidivism: A meta-analytic review of traditional and non-traditional drug courts. *Journal of Criminal Justice*, 40(1), 60–71.

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., & Hardeman, W. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, 350. https://doi.org/10.1136/bmj.h1258

Morgan, D. L. (2014). Pragmatism as a Paradigm for Social Research. *Qualitative Inquiry,* 20(8), 1045–1053. https://doi.org/10.1177/1077800413513733

O'Cathain, A., Thomas, K. J., Drabble, S. J., Rudolph, A., & Hewison, J. (2013). What can qualitative research do for randomised controlled trials? A systematic mapping review. *British Medical Journal*, 3(6), e002889.

Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *British Medical Journal,* 332, 413–415.

Pawson, R., & Tilley, N. (1994). What works in evaluation research? *British Journal of Criminology,* 34(3), 291–306.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: Sage Publications.

Popay, J., & Williams, G. (1998). Qualitative research and evidence-based healthcare. *Journal of the Royal Society of Medicine,* 91(35_suppl), 32–37.

Porter, S., McConnell, T., & Reid, J. (2017). The possibility of critical realist randomised controlled trials. *Trials,* 18(1), 133.

Rock, P. (2017). The Foundations of Sociological Theories Of Crime. In A. Liebling & S. Maruna (Eds.), The Oxford Handbook of Criminology, 21 (6th ed.). Oxford University Press.

Rycroft-Malone, J., Seers, K., Eldh, A. C., Cox, K., Crichton, N., Harvey, G., … Wallin, L. (2018). A realist process evaluation within the Facilitating Implementation of Research Evidence (FIRE) cluster randomised controlled international trial: an exemplar. *Implementation Science*, 13(1), 138. https://doi.org/10.1186/s13012-018-0811-0

Sampson, R. J. (2013). Common Concerns, Unique Cultures? European and American Criminology. *Criminology in Europe*, 12(1), 12–14.

Sampson, R. J., Winship, C., & Knight, C. (2013). Translating Causal Claims: Principles and Strategies for Policy-Relevant Criminology. *Criminology & Public Policy*, 12(4), 585–586.

Seawright, J. (2016). Multi-method social science: Combining qualitative and quantitative tools. Strategies for social inquiry, Cambridge: Cambridge University Press.

Seers, K., Rycroft-Malone, J., Cox, K., Crichton, N., Edwards, R. T., Eldh, A. C., … Wallin, L. (2018). Facilitating Implementation of Research Evidence (FIRE): an international cluster randomised controlled trial to evaluate two models of facilitation informed by the Promoting Action on Research Implementation in Health Services (PARIHS) framework. *Implementation Science,* 13(1), 137. https://doi.org/10.1186/s13012-018-0831-9

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. New York, NY: Houghton Mifflin and Company.

Sherman, L. W. (2009). Evidence and liberty: The promise of experimental criminology. *Criminology & Criminal Justice,* 9(1), 5–28.

Sherman, L. W., & Strang, H. (2004). Experimental Ethnography: The Marriage of Qualitative and Quantitative Research. *The Annals of the American Academy of Political and Social Science*, 595, 204–222. Retrieved from http://www.jstor.org/stable/4127621

Siddiqui, N., Gorard, S., & See, B. H. (2018). The importance of process evaluation for randomised control trials in education. *Educational Research*, 1–14.

Sullivan, C. J., & Welsh, B. C. (2017). Methodological advances in crime prevention research. In B. Teasdale & M. S. Bradley (Eds.), Preventing Crime and Violence. Springer International Publishing.

Suzuki, E., & VanderWeele, T. J. (2018). Mechanisms and uncertainty in randomised controlled trials: A commentary on Deaton and Cartwright. *Social Science & Medicine,* 210, 83-85.

Symonds, J. E., & Gorard, S. (2010). Death of mixed methods? Or the rebirth of research as a craft. *Evaluation & Research in Education*, 23(2), 121–136. https://doi.org/10.1080/09500790.2010.483514

Tilley, N. (2009). Sherman vs Sherman Realism vs rhetoric. *Criminology and Criminal*

Justice, 9(2), 135–144.

Weisburd, D., Lum, C. M., & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *The Annals of the American Academy of Political and Social Science*, 578(1), 50–70.

Wheeldon, J. (2012). Legacy and longing: Criminological education, understanding, and modesty. In American Society of Criminology Annual Meeting. Chicago, Illinois.

Wheeldon, J., Heidt, J., & Dooley, B. (2014). The trouble(s) with unification: Debating assumptions, methods and expertise in criminology. *Journal of Theoretical and Philsophical Criminology*, 6(2), 111–128.

White, H. (2009). Theory-based impact evaluation: principles and practice. *Journal of Development Effectiveness*, 1(3), 271–284.

White, H. (2013). *The use of mixed methods in randomised control trials. New Directions for Evaluation*, 2013(138), 61–73.

Wiggins, B. J. (2011). Confronting the dilemma of mixed methods. *The Journal of Theoretical and Philosophical Psychology*, 31(1), 44–60.