Four forms of 'offender' rehabilitation:
Towards an interdisciplinary perspective

Fergus McNeill, University of Glasgow

Address:
Fergus McNeill
Professor of Criminology & Social Work
Scottish Centre for Crime and Justice Research
University of Glasgow
Ivy Lodge
63 Gibson Street
Glasgow G12 8LR
UK
Email: Fergus.McNeill@glasgow.ac.uk
Abstract

This paper aims to advance the case for a more fully interdisciplinary understanding of offender rehabilitation, partly as a means of shedding light upon and moving beyond contemporary ‘paradigm conflicts’. It begins with a review of current arguments about what a credible ‘offender’ rehabilitation theory requires and by exploring some aspects of current debates about different theories. It goes on to locate this specific kind of contemporary theory-building in the context of historical arguments about and critiques of rehabilitation as a concept and in practice. In the third part of the paper, I explore the nature of the relationship between desistance theories and rehabilitation theories, so as to develop my concluding argument; i.e. that debates about psychological rehabilitation have been hampered by a failure to engage fully with debates about at least three other forms of rehabilitation (legal, moral and social) that emerge as being equally important in the process of desistance from crime.

Keywords

Rehabilitation, RNR, Desistance, Sentencing, Corrections
Introduction

I embark on writing this particular paper for this particular journal with a degree of trepidation. Firstly, I am not a psychologist and am naturally hesitant about contributing a paper to a distinguished psychological journal, which might be read as a critique of a disciplinary perspective that I understand only partially but which I respect and admire. Secondly, I am wary of being seen as issuing another salvo in a contemporary conflict between competing models of rehabilitation, variously cast as between the Risks-Needs-Responsivity model and the Good Lives Model model (see Andrews, Bonta and Wormith, 2011; Ward, Yates and Willis, forthcoming), or between ‘risk-based’ and ‘strengths-based’ approaches (Maruna and LeBel, 2003, 2009), or between ‘what works?’ and ‘desistance-based’ corrections (Farrall, 2002; McNeill, 2006). Though a great deal has been learned from these recent debates, both about the similarities and intersections between these various models and perspectives and about their genuine points of contention, I want to suggest instead that it is time to take a step back to questions that relate to the foundations of rehabilitation itself.

My central argument is that rehabilitation’s ‘paradigm conflicts’ arise in part from and are perpetuated by an inadequately articulated conception of what we mean by ‘offender rehabilitation’ itself. In essence, I argue that a psychological understanding of rehabilitation provides a necessary but insufficient basis for correctional or forensic psychology, and for anyone involved in the fields of sentencing and corrections. Both the concept and the project of rehabilitation compel and require the development of a more fully interdisciplinary perspective; one that combines insights from (at least) criminology, history, philosophy and the sociology of law and punishment, as well as from the more obviously related ‘psy’ disciplines of psychology, psychiatry and social work.

Paradigm conflicts and rehabilitation theories

This paper is of a somewhat different type from its companion paper in this issue by Devon Polaschek. Nonetheless, Polaschek’s excellent paper provides an obvious place to begin since her balanced and sympathetic critical analysis of the Risks-Needs-Responsivity model (Andrews and Bonta, 2010 – RNR) directly or indirectly raises many of the issues that I want to develop here. Her account of what types of theory-building are involved or implied in the development of RNR is particularly important. Polaschek points out that the adequacy of a theory should be evaluated at the explanatory level at which it is pitched, although we might add the caveat that this should also compel the theory’s proponents not to make claims or offer prescriptions that relate to other levels or types of analysis or theory-building. Drawing on work by
Ward and Hudson (1998), Polaschek argues that RNR should be understood and assessed as a Level I theory; that is, one which is:

global and multifactorial, necessarily broad in focus and lacking sufficient detail to directly shape the design of specific interventions. They are hybrids (Ward et al., 2007): informed partly by etiological theories, but also incorporating the underlying values and assumptions of intervention, therapy strategies, change processes, programme context and setting, and implementation, all in an abstract, ”high level” way. Their purpose is to provide general parameters in which rehabilitative endeavours will operate, and the support developments on the other two levels1 (Polaschek, this issue, p10).

Ward and Maruna (2007) draw on similar sources to suggest that a ‘good’ offender rehabilitation theory needs to combine three elements; a clear and explicit account of (1) the general principles underlying rehabilitation (metaphysical, epistemological, ethical and normative); of (2) its etiological assumptions (which by explaining the ‘causes’ of the offending behaviour guide understanding of ‘treatment goals’); and of (3) the intervention implications (the 'how to' element that guides effective interventions).

Drawing on a close reading of the latest articulations of the RNR theory (or family of theories), Polaschek prefaces her analysis of its strengths, weaknesses and implementation problems by noting its ‘overarching principles’. RNR is a theory based on respect for persons and for the normative context of rehabilitation; it is based on empirically validated psychological theory; and it asserts the importance and legitimacy of services that prevent crime, even outside the justice system. But, in terms of the relationships between these three overarching principles, and crucially for my purposes, she notes that:

‘Although values other than empirical standing are important in evaluating theory quality (see Ward, et al., 2006), empirical validity and practical application have primacy when evaluating intervention theories... their main purpose is application, rather than knowledge generation for its own sake (p12-13).

These are key points to which I will return, but first it may be helpful to briefly summarise Polaschek’s appraisal of RNR. In terms of its considerable strengths, she suggests that RNR has strong unifying power, external consistency and explanatory depth. In other words, it explains well why certain kinds of rehabilitative programmes work better than others, and why certain skills and practices are empirically associated with better outcomes. It also addresses a major weaknesses in ‘social-level etiological

---

1 These ‘other two levels’ include Level II theories which deal with a single etiological factor and Level III theories which are local theories of the offence process itself.
theories’; that is, their tendency to over-predict crime amongst those subject to social disadvantage and their failure to explain why some similarly socially situated individuals offend when others do not. RNR has also been subject to considerable empirical validation through evaluation research; its core principles have been refined by such evidence but never falsified. Moreover, it has proved to be a fertile theory in terms of generating numerous practical applications. Overall, it has made an ‘original substantive contribution to the development of criminal justice assessment, intervention, research, programme accreditation and programme integrity’ around the world (Polaschek, this issue, p15, emphasis in original). For the sake of clarity, let me make clear that I agree with that assessment.

But Polaschek also provides some thoughtful critique of RNR and notes several problems in its ‘translation’ into practice. For example, despite its merits, there are important limits in RNR’s explanatory depth; specifically, she argues that the responsivity principle, despite its centrality to the model both in terms of its humanity and its effectiveness, is under-articulated and theoretically unsophisticated – a ‘catch-all category’. One might add here that, whereas other aspects of RNR have become more elaborated and sophisticated with the passage of time, the responsivity principle has, at least in its implementation, become more vulnerable to over-simplification and misinterpretation (Robinson and Crow, 2009). Thus, whereas Andrews, et al., (1990: 20) initially defined responsivity in terms of the individualised matching of ‘[s]tyles and modes of service... to the learning styles and abilities of offenders’, the later distinction between individual and general responsivity (see, for example, Bonta and Andrews, 2011), and the subsequent over-emphasis on the latter, has too often allowed a complex and important principle to be reduced to the generalised injunction: ‘Employ social learning and cognitive behavioural strategies’ (Bonta and Andrews, 2011: 21).

In contradistinction, as Polaschek suggests, one of the benefits of strengths-based approaches has been their deeper engagement with individual motivations and their recognition of the importance of giving people (positive) ‘reasons to want to engage in desistance and change (e.g. Porporino, 2010: Ward and Maruna, 2007), not just the capacities to do so’ (p16). RNR also suffers from a ‘conceptual gap’ that exists between the main targets for change (that is, the ‘central eight’ risk factors) and the theoretical resources that are necessary to translate them into intervention (design, formulation, planning and monitoring). Moreover, drawing on the recent work of Mann, Hanson and Thornton (2010), Polaschek describes the ‘central eight’ themselves as proxies or placeholders for a series of more complex phenomena. A broad list of factors, she argues (following Ward and Maruna, 2007), is no substitute for a full understanding of the mechanisms behind offending, of the interrelationships between risk factors and of how change processes affect these interrelationships so as to moderate risk (though

---

2 I am indebted to Gwen Robinson for first making this point to me.
she suggests that such work is more appropriately a task for level II or level III theories).

With respect to the translation of RNR into practice, Polaschek is equally perceptive. She notes that the model is hard to comprehend and hard to communicate effectively, partly because of the use of complex technical language. Perhaps more importantly (though the problems are connected), she recognises some of the potential and actual mistranslations of RNR into the field that have unfairly affected the theory's standing. Thus some programmes in the RNR tradition have been accused of being focused solely on risk reduction (and thus the putative public good), while ignoring the legitimate (approach) goals and aspirations of offenders; as such they have failed to engage offenders and build the types of therapeutic alliance that positive change requires. They have focused on criminogenic needs to the exclusion of other needs that nonetheless require assistance and without which work on criminogenic needs may stall or never get started. They have worked to develop the capacities and skills of offenders without attending sufficiently to their motivations and their social opportunities (see McNeill and Whyte, 2007; McNeill, 2009). However, Polaschek, unlike less careful critics (including me, in some of my previous work, e.g. McNeill, 2001; 2004), is careful to distinguish between what RNR theory intends and how it has come to be applied in particular contexts for particular reasons.

Thus, while none of the criticisms mentioned in the last paragraph can be straightforwardly directed at RNR theory itself, they all have considerable bite when applied to examinations of its operationalization in some contexts (for a compelling and authoritative account of the largest scale attempt to operationalize RNR -- in England and Wales -- see Raynor and Robinson, 2009, and more generally Porporino, 2010). In such examinations, evidence of very significant problems has emerged; principal among these is perhaps the over-reliance on one style of intervention (‘structured, cognitive behavioural, closed-groups based treatment programming’), and with it the use of overly-scripted, manualised, homogenized approaches, which themselves seemed to neglect the very concept of individual responsivity itself. While it is true that more recent developments in programme design and delivery (for example, the increasing use of ‘rolling programmes’ that allow offenders to pursue more individualised programmes even within groups, thus still allowing for peer support and challenge) have worked much harder to engage offenders and to accommodate diversity, beyond these technical ‘implementation’ problems lies a deeper malaise to which Polaschek refers in her conclusion: ‘the more significant problems remain in actually implementing programmes in [the] unforgiving settings’ (Polaschek, this issue, p26) that contemporary criminal justice systems all too often provide.

Though we will return to some of these specific arguments below, in the next section I want to connect one of Polaschek’s first insights – the relative priority of empirical evidence as opposed to values-based arguments in building rehabilitation theory – with one of her last – the unforgiving nature of contemporary corrections. To
do so, however, requires a brief detour into the history of rehabilitation theory and practice.

**Revisiting rehabilitation's history**

It is, in one sense, hardly surprising that rehabilitation theories remain the subject of controversy; the concept of rehabilitation has always been both remarkably elastic and hotly contested. Bottoms (1980) has pointed out that whereas rehabilitation, at least on a Foucauldian reading, has sometimes been represented as being a project of 'coercive soul-transformation' (or better as being too readily corralled into such a project), a different meaning of the term had been identified in the work of the 18th century Classicists (e.g. Beccaria 1764/1963) who argued for the use of punishment as a way of 'requalifying individuals as... juridical subjects' (Foucault 1975/1977: 130). Though correctional psychologists might take issue with Foucault's reading of the ways in which they have exercised their professional power, the very term 'correctional' speaks to an essentially utilitarian conception of rehabilitation. This is a form of rehabilitation in which the ultimate aim is to change the offender so as to reduce harm (at least) and contribute to the public good, including the offender's best interests (at best). Beccaria clearly has in mind a different aim: the full restoration of the formerly errant citizen so that s/he can once again enjoy all of his/her rights and fulfil all of his/her responsibilities. The central distinction here is that Beccaria's notion of 'requalification' is not based principally on maximising social utility or welfare but on restoring duties and rights – hence, it is a more deontological conception of rehabilitation. I will return briefly to this important but quite different notion of rehabilitation as requalification in the final section of this paper.

The broadly correctional or utilitarian form of rehabilitation has, however, been expressed in very different ways. Two decades ago, Edgardo Rotman (1990) in a brilliant and brief introductory chapter to his book 'Beyond Punishment', summarised the history of rehabilitation as being represented in four successive models. For Rotman, the story begins with the rise of the penitentiary, as a place of confinement where the sinner is given the opportunity to reflect soberly on their behaviour, and on how to reform themselves, perhaps with divine help. This ideal stressed the reformative potential of both contemplation and work, sometimes in combination. But the religious ideas of rehabilitation expressed in the penitentiary evolved rapidly in the 19th and early 20th centuries with the emergence of the 'psy' disciplines. The idea that rehabilitation was about reforming the sinner, bringing them to acknowledgement of their sins, invoking repentance and requiring some penance before restoration, was progressively supplanted with a more scientific or medical model. Here, rehabilitation was understood as a form of treatment which could correct some flaw in individuals, whether physical or psychological, thus remedying the problem of their behaviour. In the later half of the 20th century, this more medical or therapeutic version of rehabilitation was itself displaced, to some extent, by a shift in emphasis towards a model based on social learning in which behaviours were understood as learned
responses that could be unlearned. In this context rehabilitation was recast not as a sort of quasi-medical treatment for criminality but as the re-education of the poorly socialised (Garland 1985, 2001).

This historical progression and reformulation of rehabilitation, of course, was spectacularly derailed in the 1970s. In this respect, an edited collection by Bottoms and Preston (1980), ominously entitled ‘The Coming Penal Crisis’, emerges as a remarkably prescient piece of work. In that collection, Tony Bottoms (1980), in a typically erudite and compelling chapter that deals with the collapse of the rehabilitative ideal, summed up its flaws and failings. First of all, rehabilitation came to be seen as being theoretically faulty in that it misconstrued the causes of crime as individual when they were coming to be understood as being principally social and structural, and it misconstrued the nature of crime, failing to recognise the ways in which crime is itself socially constructed. Secondly, rehabilitative practices had been exposed as being systematically discriminatory, targeting coercive interventions on the most poor and disadvantaged people in society. Third, rehabilitation was seen as being inconsistent with justice itself because judgements about liberty had come to be unduly influenced by dubious and subjective professional judgements hidden from or impenetrable to the offender. Through the development of the ‘psy’ disciplines, experts emerged with the supposed capacity to ‘diagnose’ what was wrong with the offender, and the offender was cast as a victim of his or her lack of insight. By implication, unless and until the offender was ‘corrected’ by the expert, s/he could not be treated as a subject. Fourth, it was argued that rehabilitation faced a fundamental moral problem rooted in its attempts to (psychologically) coerce people to change. Finally, at the time when Bottoms was writing, the empirical evidence seemed to suggest that, despite its scientific pretensions, rehabilitation did not seem to work.

Powerful though it is, there are flaws in this critique. Crucially, emerging evidence about ‘what works?’, not least evidence provided by RNR's progenitors and their colleagues, played a vital role in challenging the last point; thankfully, there is an evidence base for rehabilitation. But that evidence base, in and of itself, does not address the other four criticisms which I will label the problems of crime theory, of rehabilitation and injustice, of dubious expertise, and of coerced correction. Contemporary rehabilitation theories perhaps have more in common with a social learning or social psychological than a medicalised version of correctional rehabilitation (on the distinction, see Johnstone, 1996), although some critics discern in the concepts and language of RNR (specifically the oft-repeated and unhelpful use of terms like ‘dosage’) an implicit medical model (e.g. Kendall, 2004). But, even if rehabilitation is not now based on a strictly medical model, the full force of Bottoms’ (1980) four criticisms is only partly deflected.

Taking each problem in turn, while contemporary rehabilitation's underlying crime theories avoid the pathologising traps of individualistic positivism, they still have
to engage somehow with the problem that crime is (at least in part) a social construct. We do not choose to pursue all interpersonal or social harms through criminalisation; the explanandum of crime theories is itself socially conditioned. That single insight has profound consequences. It affects and infuses the normative contexts of rehabilitative work (including raising difficult questions about who and what gets selected for penal ‘correction’ and who and what does not) and it creates a series of complex methodological quandaries. Not least amongst these is the fact that our dependent variable of choice for evaluating rehabilitative strategies – reconviction – is revealed as being deeply and irrevocably flawed. It is not a straightforward measure of behaviour change (far less of positive shifts in identity); it is a measure of a series of interlocking social reactions to perceptions of behaviour (witnessing, reporting, detecting, prosecuting, sentencing, conviction). To extend a medical analogy, judging rehabilitative interventions by reconviction is a bit like judging the success of health interventions by whether anyone shows up at the doctor again (for any reason, not just in relation to the problem that has actually been treated); judging them by re-incarceration is like judging medical treatments by re-hospitalisation (again for any reason) after treatment. The reliance on reconviction also misses the point that even if reconviction were a good proxy for reoffending, the ‘mere’ absence of offending does not in and of itself signal progress towards long term or permanent desistance from crime or, putting it another way, long-term committed compliance with the law. If we have that longer term objective in mind, which is more significant: a minor lapse by someone committed to but struggling with desistance, or a crime-free lull while a committed re-offender regroups?

The other three problems also remain but for reasons of space need to be passed over more swiftly here. The problem of rehabilitation and injustice, or of rehabilitation’s place in or association with discriminatory justice practices, remains a challenging one, not just because of systematic (if unintended) biases in terms of who gets selected for ‘correction’ of which sorts and who gets defined as ‘incorrigible’, but also in terms of the extent to which rehabilitation’s resources (principally assessment tools and interventions approaches) are sufficiently sensitive to, for example, gender differences and cultural diversity (on which see, for example, Raynor, 2007; Robinson and Crow, 2009: Ch. 6). The problem of dubious expertise is only partly resolved by the development of a more robust evidence base for rehabilitation. While it is true that contemporary approaches, perhaps to varying degrees, begin to recognise the need to treat people as active subjects in their own rehabilitation (and not simply as the passive objects of expert intervention), there remains considerable force in critiques both of the professional power of the ‘psy’ disciplines (e.g. Rose, 1989) and, more specifically, of the increasing exercise of that power not in ‘treatment’ or ‘therapeutic’ decisions but in

---

3 As such, it violates ‘Sellin’s dictum’ which states that ‘[t]he value of criminal statistics as a basis for the measurement of criminality in geographic areas decreases as the procedures take us farther away from the offence itself’ (Sellin, 1931: 346). I am grateful to Elvinas Blazevicius for bringing Sellin’s dictum to my attention.
influencing decisions about justice itself, specifically in relation to sentencing or release-decision making. It is in the complex and fraught interfaces between these forms of judgement (psychological and legal) that we find the sources, for example, of burgeoning resentment amongst English prisoners towards what apparently seems to them to be the capricious and illegitimate exercise of ‘soft power’ by prison psychologists (Crewe, 2009; and more generally Maruna 2011).

Of course, this is closely related to the enduring problem of coerced correction (or ‘authoritarian’ rather than ‘anthropocentric’ rehabilitation – see Rotman, 1990); a problem exacerbated in systems where increasing confidence in rehabilitation’s effectiveness has coincided with the development of a ‘toughening up’ of penal policy. One apparently minor aspect of this conjunction in England and Wales was the withdrawal of any requirement for consent to accredited programme conditions within community sentences. These programmes therefore involve rehabilitative treatment without consent; a development that seems to have come to pass without much critique from psychologists whose ethical standards it might seem to compromise.

Fortunately, Bottoms’ (1980) also provides some suggestions about how we might best respond to these problems. Reflecting the pessimism of the times, four of his five options imply or require a shift away from rehabilitation, at least as traditionally conceived. However, he also argued that rehabilitation could be rescued through its own correction; that is, by attending more carefully to questions of consent, by committing adequate resources to make it more likely to be effective and by conducting our rehabilitative activities in ways which are more respectful of liberty. More specifically, we could ensure that the intrusions that rehabilitation imposes on the offender are never greater than is merited by their offending behaviour, placing rehabilitative requirements within the envelope of proportionality.

There is no space here to properly review the renaissance of rehabilitation in the 1980s and 1990s. However, leaving aside the better known story of the advancement of the evidence base, it is worth recalling that there was also, in the 1980s a brief flurry of writing about new normative or philosophical approaches to rehabilitation, including Rotman’s (1990) work (see also Cullen and Gilbert, 1982). The ‘new rehabilitationists’ (see Lewis, 2005), much as Bottoms’ suggested, proposed four principles to guide rights based rehabilitation: the assertion of the duty of the State to provide for rehabilitation; the establishment of proportional limits on the intrusions opposed; the principle of maximising choice and voluntarism in the process; and a commitment to using prison as a measure of last resort. However, as Robinson (2008) has shown, in policy and practice (in England and Wales at least) what emerged was not a right-based but a more profoundly utilitarian and correctional form of rehabilitation increasingly influenced by the preoccupation with public protection and risk reduction. Under this paradigm, probation officers were required to intervene with or treat the offender to reduce re-offending and to protect the public. What is critical about this paradigm was that the
‘client’ -- the person or social group that the probation service is serving -- is not the offender. Rather probation is trying to change offenders to protect the law-abiding (see McCulloch and McNeill, 2007, Robinson and McNeill, 2004). Within this paradigm, practice was rooted in professional assessment of risk and need governed by structured assessment instruments; the offender was less and less an active participant and more and more an object to be assessed through technologies applied by professionals and compulsorily engaged in structured programmes and offender management processes as required elements of legal orders imposed, as we have already noted, irrespective of consent (McNeill, 2006).

Though this brief account of rehabilitation’s recent development may seem some way distant from Polaschek’s discussion of RNR, in fact it speaks directly to her point about the relative import of questions of values and of questions of evidence, and to her point about the hostile correctional climate. Though the designers of correctional interventions are not directly responsible for the ways in which their ideas are refracted or distorted in the processes of transfer and implementation, anyone venturing to offer a rehabilitation theory to the world of practice has some responsibility for thinking critically (and sociologically) about how their ideas might be used and abused. It is manifestly reasonable to suggest that assessing the technical merits of a rehabilitation theory must rely primarily heavily on judgements about evidence. However, the lesson of rehabilitation’s troubled and troubling history is surely that the real-world and often unintended effects of correctional innovation must surely compel innovators to set some normative limits around their rehabilitative intentions, techniques and intrusions.

Since the history of rehabilitation is littered with adverse and unintended consequences (for offenders, victims and communities), it is essential to articulate some normative constraints around and within rehabilitation. To meet that obligation, rehabilitation theorists need to engage seriously with moral philosophy (which can help explore these necessary normative principles and constraints) and with the sociology of punishment (which can help to anticipate and expose the risks of unintended consequences). Perhaps the recent success of the GLM in engaging and inspiring practitioners owes something to that fact that its engagement with these questions has been both enthusiastic and sophisticated (Ward and Birgden, 2007).

Desistance theories and rehabilitation theories

Philosophy and sociology are not the only other disciplines with which rehabilitation theorists need to seriously engage. The brief discussion above of the problems posed by inadequate crime theories and flawed dependent variables suggests the need for a robust engagement with criminology too. This is, of course, already a well-developed dialogue, though not always a mutually respectful and helpful one (for an overview, see Hollin, 2002). But even where the dialogue has been reasonably productive in terms of rehabilitative ‘applications’ (for example, in the relationship
between risk assessment and developmental criminology), there remains a lot to talk about. One of the most interesting emerging conversations concerns the relationships between criminological evidence about the process of desistance from crime (that is, how and why people stop offending) and the theory and practice of rehabilitation. Space permits only a cursory engagement with this dialogue here; interested readers will find much more developed discussions in a number of recent books and articles (for example, Farrall, 2002; Laws and Ward, 2010; McNeill, 2006; McNeill and Whyte 2007; McNeill, 2009; McNeil, Raynor and Trotter, 2010; Ward and Maruna, 2007; Weaver and McNeill, 2010). Here, I will point to just some aspects of this dialogue that seem especially pertinent to the construction of rehabilitation theories and practices.

The central question here concerns how a body of evidence about how and why people change can and should speak to the practical challenge of helping them do so. Put that way, the relevance of desistance research for thinking about rehabilitation seems obvious; securing desistance is, after all, a central objective of rehabilitation. In a sense, of course, both RNR and GLM are (the latter more explicitly than the former: see Ward and Maruna, 2007) theories of ‘assisted desistance’; that is, they aim to explain how we can best help people to stop offending. Crucially though, RNR’s underlying approach is based on targeting the intervention on factors that predict criminal behaviour, not on factors that predict desistance. While we might tend to assume that these two sets of factors (risk factors and desistance factors) are mirror images of one another, or are simply opposing ends of a series of continua, recent evidence suggests that the picture is much more complex (Farrington, 2010). Explanations of why people offend may be a key part of building rehabilitation theories, but increasingly criminologists recognise the need for another kind of explanatory or theory concerned with the etiology not of crime but of desistance from crime.

There are three main theoretical perspectives in desistance research, which focus respectively on the significance of age and maturation, of life transitions and the social bonds associated with them, and of narrative changes in personal and social identity (Maruna, 2001). Increasingly, desistance theorists have drawn these strands together, arguing that the process of desistance is produced in the interplay between these three sets of factors (Farrall and Bowling, 1999; Farrall et al., 2011). One important development in the desistance literature is concerned with seeking a clearer account of the process of desistance, or of the different stages of the desistance process (though ‘aspects’ might be a better word than ‘stages’, since none of the researchers discussed below suggest a simple linear progression). Thus Giordano et al. (2002), in a particularly important contribution, offer a theory of ‘cognitive transformation’ in which developing ‘openness to change’ somehow meets with ‘…exposure to a particular hook or set of hooks for change’ (p1000) and ‘one’s attitude toward [such hooks]’ (p1001). This may in turn allow the development of a third stage in which the person begins envisioning ‘an appealing and conventional replacement self’ (and perhaps turning
away from a ‘feared self’, on which see Paternoster and Bushway, 2010). In somewhat similar vein, Maruna and Farrall (2004: 27-8) explain that:

‘a lull can turn into secondary desistance when two things happen. First, the person finds a source of agency and communion in non-criminal activities. They find some sort of “calling” -- be it parenthood, painting, coaching, chess or what Sennett (2003) calls “craft-love” -- through which they find meaning and purpose outside of crime... The second part of our desistance formula, like that of Lemert’s deviance theory, involves societal reaction. The desisting person’s change in behavior is sometimes recognized by others and reflected back to him in a “delabeling process”... ’ (Maruna and Farrall, 2004: 28).

In secondary desistance, it is not just that criminal behaviour stops, ‘existing roles become disrupted’ and a ‘reorganization based upon a new role or roles will occur (Lemert, 1951: 76)... desistance does involve identifiable and measurable changes at the level of personal identity or the ‘me’ of the individual’ (Maruna et al., 2004: 274). It is in secondary desistance perhaps that the person experiences or accomplishes the fourth stage in Giordano et al.’s (2002: 1002) model: ‘a transformation in the way the actor views the deviant behaviour or lifestyle itself’.

A recent elucidation of this process has been provided by Bottoms and Shapland (2011), drawing on the findings of the Sheffield desistance study of male young adult recidivists. Their model involves a seven-stage process in which current offending is influenced by a triggering event; which leads to the wish to try to change; which leads the offender to think differently about himself or his surroundings; which leads the offender to take action towards desistance. However, these fledgling attempts to desist may be threatened by obstacles, barriers and temptations, so the desister must find reinforcing factors (from within himself or more likely within his changing social relations) to maintain the change which, if successful, may ultimately enable with the establishment of a crime-free identity. Crucially, progression along this journey is affected, for better or worse, by both the ‘pre-programmed potential’ of the individual and his ‘social capital’ resources.

Drawing on these and other studies, a body of scholarship has emerged which, following Farrall’s injunction that practice should become ‘desistance-focused’, seeks to interpret desistance research for practice. This work tends to stress (albeit to varying degrees) six central themes:

1. Since desistance is an inherently individualised and subjective process, approaches to intervention must accommodate and exploit issues of identity and diversity. (Weaver and McNeill, 2010)
2. The development and maintenance not just of motivation but also of hope become key tasks for workers (Farrall and Calverley, 2006).

3. Desistance can only be understood within the context of human relationships; not just relationships between workers and offenders (though these matter a great deal) but also between offenders and those who matter to them (Burnett and McNeill, 2005; McNeill, 2006).

4. Although we tend to focus on offenders’ risk and needs, they also have strengths and resources that they can use to overcome obstacles to desistance – both personal strengths and resources and strengths and resources in their social networks. Supervision needs to support and develop these capacities (Maruna and LeBel, 2003).

5. Since desistance is about discovering agency, interventions need to encourage and respect self-determination; this means working with offenders not on them (McCulloch, 2005; McNeill, 2006).

6. Interventions should work on ‘human capital’ (or developing offenders’ capacities and skills) (Maguire and Raynor, 2006), but also on ‘social capital’ (developing relationships and networks that generate opportunities) (Farrall, 2002, 2004; McNeill and Maruna, 2007; McNeill and Whyte, 2007).

To some extent, the findings and prescriptions of desistance research first came to prominence in the UK as a critique of the narrow interpretations and overly managerialized implementations of ‘what works’ discussed above (see, for example, Farrall, 2002; McNeill, 2004, 2006). On my part at least, this was never intended as a direct critique of RNR as a rehabilitation theory – nor of any other rehabilitation theory. But perhaps, as an advocate of a ‘desistance paradigm’ (McNeill, 2006), I should hold myself to the standard I proposed in the last section and elaborate both the intended and the unintended consequences of my juxtaposition of a ‘what works’ paradigm with a desistance paradigm.

The tensions between narrow interpretations of ‘Evidence-Based Practice’ (not least in terms of what counts as evidence), the managerialized implementation of ‘what works’ and a desistance perspective are indeed serious and real; the tensions between a desistance-based perspective and RNR theory are perhaps somewhat less substantial. They require much more discussion than I can offer here (see McNeill et al., 2010; Porporino, 2010), but let me try to summarise some of the issues that need to be addressed, and also suggest some of the potential synergies, with reference to the notion of ‘Pasteur’s Quadrant’ (see figure 2, below).

‘Pasteur’s Quadrant’ is the title of a book by Donald Stokes (1997); he developed the concept in arguing against Vannevar Bush, FDR’s science advisor, who famously drew a firm distinction between basic research and applied research, seeing the latter
as being inherently inferior. Bush argued influentially that applied research (here represented by Thomas Edison) invariably drives our pure (here represented by the theoretical physicist Nils Bohr). But for Stokes, Pasteur epitomised the type of researcher whose work and impact refutes the distinction – whether in relation to the pasteurization process, the germ theory of disease, or the biology of fermentation, Pasteur developed entire fields of study at the same time as inventing improvements in health and commerce.

Using this model can perhaps help to clarify the relationships between desistance theories and rehabilitation theories. ‘Pure basic’ desistance research studies the process for its own sake, inverting Polaschek’s priorities (discussed above); it is not principally concerned with applications of the knowledge thus generated (but of course this is not to say that it doesn’t have applications). Indeed, it is possible that one reason for the muted impact of desistance research (until recently at least) is that it may have seemed to be more directed by scientific curiosity than by the quest to solve pressing social problems. Conversely, RNR and GLM – like most rehabilitation theories – do not represent pure basic research, but neither are they simple products of Edisonian, theory-less tinkering and experimenting. While they do draw on and depend on evaluation evidence (albeit to different extents), they are better understood as representing ‘use-inspired research’. However, the overly narrow anti-theory implementation of RNR as the ‘what works’ (or better ‘whatever works’) approach perhaps did come close to Edisonian ‘pure applied’ research. Though beyond the scope of this paper, there are even more alarming dangers of privileging a purely experimental criminology which commends and relies upon RCTs as the principal basis for innovation and development (on which see Hough, 2010; Maruna and Barber, 2011).

What we might call the emerging body of ‘assisted desistance’ research (which specifically intends to examine what helps people desist, and what frustrates their efforts) also fits into Pasteur’s Quadrant. It is clearly use-inspired, but it is also driven by pure basic research and theory-building around how and why people stop offending (not why they start or why they persist). In other words, its findings and prescriptions rely on analyses of the process that rehabilitative interventions exist to support as much (or more than) on evaluations of those interventions. To use an educational rather than a medical analogy, developing an approach to rehabilitation focused on understanding and supporting desistance is like developing an approach to teaching focused on understanding and supporting learning. By contrast, an approach focused only on evaluation evidence for different rehabilitative strategies is like an approach to teaching based only on studying which teaching methods work best. More particularly, the RNR analogue here would be an empirically validated approach to teaching based on understanding and targeting the correlates of learning problems, not of successful learning. The GLM, by contrast and in line with desistance-based perspectives, does have its roots in a theory of human flourishing.
As the analogy implies, a desistance perspective construes and influences intervention in a way which is subtly different from (but not incompatible with) most forms of rehabilitation theory. A good teacher, after all, needs to understand both how children develop and learn, and which teaching strategies work best (including to address learning problems). But where theory building starts out with a ‘remedial’ intervention in mind, it will almost inevitably tend to put intervention itself at the heart of the process of change, arguably misconstruing the remedy simply as its own ‘product’. By way of contrast, desistance-based perspectives stress that the process of change exists before, behind and beyond the intervention (McNeill, 2006, 2009); most people stop offending sometime with or without interventions, and sometimes even in spite of them— the practical challenge is to help them do so more swiftly and certainly. Rehabilitative interventions then, in desistance-based perspectives, are cast in an important but supporting role in the wider enterprise of desistance; a process that belongs to the desister. In other words, rehabilitative interventions do not cause change but they may support it. So, how we think about and construct intervention needs to be ‘embedded’ within an understanding of desistance (McNeill, 2006), just as thinking about teaching needs to be embedded in an understanding of learning and human development.

The analogy however, also exposes important limitations of desistance perspectives. Such perspectives will help rehabilitation theory developers and practitioners to better understand the process they are trying to support, and what rehabilitation ought to be trying to do, but they do not and indeed cannot tell developers or practitioners exactly what to do. Desistance perspectives do not and cannot in and of themselves offer empirically validated rehabilitation models nor even empirically testable ones – but they can and should inform the better development of such models. In such development work, the fourth quadrant in Stokes’ (1997) diagram (left blank in his model but marked with a question mark above) may also come into play. Other important forms of knowledge and expertise are routinely neglected in the business of rehabilitation; forms of knowledge and expertise that come from the life experiences of service users and the professional experience of practitioners. These much more local and personal forms of knowledge (McNeill, 2000) may not be a reliable basis for generalisation and grand theory-building, but they are nonetheless critical to the ‘co-production’ of credible, viable and productive plans and services for supporting desistance at the personal and local level (Weaver, 2011).

More recently, it has been argued that even desistance itself is not the ultimate concern (McNeill and Weaver, 2010) – it is, after all, still defined by (the absence of) offending. But people do not simply desist; they desist into something. Desistance is perhaps best understood as part of the individual’s ongoing journey towards successful integration within the community (Uggen et al., 2006). Rehabilitation, therefore, is not just about sorting out the individual’s readiness for or fitness for reintegration; it is as much about rebuilding the social relationships without which reintegration is
impossible. Any would-be supporter of rehabilitation has to do more than try to sort out ‘offenders’; s/he needs to mediate relationships between people trying to change and the communities in which change is impeded or impelled; s/he also has to mediate the role and limits of the state itself in the process.

In sum, my argument is that, viewed from the perspective of desistance theory and research, individual level psychological theories, analysis and prescriptions are best seen as a necessary but insufficient part of developing an integrated theory of rehabilitation. Such a theory also needs to draw on criminology’s developing understandings of how to assist desistance from crime, which in turn draws on both psychological and sociological research, as well as on the knowledge and insights of ex-offenders and practitioners.

Conclusion: Four forms of rehabilitation

‘To the extent that felons belong to a distinct class or status group, the problems of desistance from crime can be interpreted as problems of mobility – moving felons from a stigmatized status as outsiders to full democratic participation as stakeholders’ (Uggen et al., 2006, p283).

Rehabilitation is a social project as well as a personal one. Whether we cast it in deontological terms as being concerned with the requalification of citizens, or in utilitarian and correctional terms as being concerned with their re-education or re-socialization, it raises profound political questions about the nature of (good) citizenship, about the nature of society, about the relationship between citizenship, society and the state, and about the proper limits of legitimate state power.

Though these political questions and concerns are ultimately beyond the scope of this paper and may seem somewhat abstract, they are in fact cheek by jowl with the practical challenges of making rehabilitation work. At least some of rehabilitation’s practical problems (and its conceptual weaknesses) come from the failure of some of its proponents and practitioners to engage adequately with these moral and political questions. Such engagement requires ‘psychological rehabilitation’ (which is principally concerned with promoting positive individual-level change in the offender) to articulate its relationships with at least three other forms of rehabilitation.

The first of these concerns the practical expression of Beccaria’s concern with the requalification of citizens; this is the problem of ‘legal or judicial rehabilitation’ – when, how and to what extent a criminal record and the stigma that it represents can ever be set aside, sealed or surpassed. Maruna’s (2011b) has recently argued cogently

---

4 This section of the paper draws on part of a conference paper co-authored with Shadd Maruna (McNeill and Maruna, 2010). I am grateful to Shadd for permission to use some of that material here.
that efforts to sponsor rehabilitation and reform must address the collateral consequences of conviction – mostly notably its stigmatising and exclusionary effects -- or be doomed to fail. No amount of supporting offenders to change themselves can be sufficient to the tasks and challenges of rehabilitation and desistance, if legal and practical barriers to reintegration are left in place.

But these barriers are not just legal – they are moral and social too. A solely psychological conception of rehabilitation is inadequate to the moral and social offence that crime represents. In simple terms, doing something for or to the offender, even something that aims at somehow changing them as to reduce future victimisation, fails to engage with other key aspects of dispensing justice. Perhaps most importantly in moral terms, rehabilitation offers no moral redress per se; it operates only on the individual ‘offender’, not on the conflict itself and not on the victim or the community (Zedner, 1994). Critically, reparation – and reparative work in particular -- seems capable of fulfilling this function in ways in which rehabilitation cannot, perhaps principally because reparation seems better able to convey (not least visibly) that redress is being actively provided. Rehabilitation, by contrast, is typically a private and secretive business, incapable of responding to the late-modern re-emergence of appetites and demands for more expressive forms of justice (Freiberg, 2001; Pratt et al, 2005).

Reparation perhaps speaks to the insistence that moral demands have to be satisfied, and moral communication secured, before moral rehabilitation can be recognised (see also Duff, 2001). In simple terms, an offender has to pay back before s/he can trade up to a restored social position as a citizen of good character (McNeill and Maruna, 2010), or as Bazemore (1998) has argued, redemption needs to be earned. This is not necessarily bad news for rehabilitation; as the Scottish Prison Commission (2008, para 33) noted, ‘one of the best ways for offenders to pay back is by turning their lives around’. But it does mean that rehabilitation theories and practices need to engage much more explicitly with questions of justice and reparation.

Ultimately, even where psychological issues are tackled, legal requalification is confirmed and moral debts are settled, the question of ‘social rehabilitation’ remains. In European jurisprudence, the concept of ‘social rehabilitation’ entails both the restoration of the citizen’s formal social status and the availability of the personal and social means to do so (Van Zyl Smit and Snacken, 2009). But here, I mean instead something that is broader, deeper and more subjective; specifically, the informal social recognition and acceptance of the reformed ex-offender. This, rather than the advancement of the ‘science’ of personal reform, is perhaps the ultimate problem for rehabilitation in practice; it lies at the root of the hostile correctional climate that Polaschek bemoans in her paper and it lies behind the mistranslation, corruption and misuse of rehabilitation theories. The eminent sociologist and social theorist Zygmunt Bauman’s analysis of this broader social climate led him to the sobering if pessimistic
conclusion that ‘...the question of ‘rehabilitation’ is today prominent less by its contentiousness than by its growing irrelevance’ (Bauman, 2000: 210).

Though we may demur from Bauman’s dystopian conclusion, I have tried to show in this paper how too narrowly psychological a conceptualisation and theorisation of rehabilitation risks and indeed invites the irrelevance of which he speaks. Until and unless rehabilitation can renegotiate its relationships with its moral foundations and constraints, with the burgeoning evidence base about the desistance process it exists to support, and with its social, political and legal contexts, it will remain vulnerable to its co-optation and bastardisation as a mere technique for risk management and public protection. Such a fate is itself the likely prelude to rehabilitation’s displacement by more certain and more constraining forms of social control. Ultimately, this is why a more fully integrated and interdisciplinary theory and practice of rehabilitation urgently needs to be developed.

Acknowledgements

I am particularly grateful to Gwen Robinson and Shadd Maruna for their helpful and insightful comments on an earlier version of this paper.
References


Open University Press.


22


Figure 1: Two Models of Desistance Processes

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. General cognitive openness to change</td>
<td>1. Current offending is influenced by a triggering event</td>
</tr>
<tr>
<td>2. Exposure to ‘hooks for change’</td>
<td>2. The offender identifies a wish to try to change</td>
</tr>
<tr>
<td>3. Availability of an appealing conventional self</td>
<td>3. The offender thinks differently about himself and his surroundings</td>
</tr>
<tr>
<td>4. Reassessment of attitudes to deviant behaviour</td>
<td>4. The offender to take action towards desistance</td>
</tr>
<tr>
<td></td>
<td>5. He experiences obstacles, barriers, temptations</td>
</tr>
<tr>
<td></td>
<td>6. He attempts to maintain change and finds reinforcers</td>
</tr>
<tr>
<td></td>
<td>7. He develops a crime-free identity</td>
</tr>
</tbody>
</table>

Figure 2: Pasteur’s Quadrant (adapted)

<table>
<thead>
<tr>
<th>Pure basic research (Nils Bohr)</th>
<th>Use-inspired research (Louis Pasteur)</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>Pure applied research (Thomas Edison)</td>
</tr>
</tbody>
</table>

Relevance for advancement of knowledge

Relevance for immediate application

---

The Sheffield study is a study of young men’s desistance, hence the masculine pronoun.

---