This is a repository copy of *How Do We Measure Success in Homelessness Services? : Critically Assessing the Rise of the Homelessness Outcomes Star*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/102904/

Version: Published Version

Article:
Johnson, Guy and Pleace, Nicholas orcid.org/0000-0002-2133-2667 (2016) How Do We Measure Success in Homelessness Services? : Critically Assessing the Rise of the Homelessness Outcomes Star. European Journal of Homelessness. pp. 31-51. ISSN 2030-3106

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

How Do We Measure Success in Homelessness Services? Critically Assessing the Rise of the Homelessness Outcomes Star

Guy Johnson and Nicholas Pleace

Centre for Applied Social Research, RMIT University, Melbourne Australia

Centre for Housing Policy, The University of York

› **Abstract_** *The Homelessness Outcomes Star (HOS) is probably the most widespread form of outcome measurement employed by homelessness service providers. Developed in the UK, the HOS is now being used by homelessness services in other European countries and Australia, while being promoted internationally as a validated set of key performance indicators. This paper examines the ideological framework that underpins the HOS, as well as the theoretical and methodological approaches that inform its operation. The review concludes that while there is some utility in the measurement of relative progress for individual service users, the HOS has important limits, both as a means of comparative outcome analysis and as a validated measure of homelessness service outcomes.*

› **Keywords_** *Homelessness, outcome measures, motivational interviewing, stages of change*

ISSN 2030-2762 / ISSN 2030-3106 online

Introduction

New Managerialism, which arose in the public sector (Exworthy and Halford, 1999), has become increasingly evident in the homelessness sector in the UK and Australia (Bullen, 2015). New Managerialism emphasizes the role of efficiency and productivity within a conceptual framework derived from a particular view of what constitutes efficient Capitalism, rather than, for example, defining organisational worth only in terms of public good. This approach, sometimes characterised as the organisational manifestation of neoliberalism, produces an emphasis on market principles in the delivery of public services. For the homelessness sector, it creates a radically new and challenging context in which 'value' is no longer derived simply from the public good of preventing and reducing homelessness, but must instead be assessed and reported upon in terms of efficiency, effectiveness and making the best use of public money.

In this context, homelessness agencies increasingly have to demonstrate their "accomplishments and inherent worth" (Greenway, 2001, p.217). Indeed, as NGOs have come to realize the importance of documenting their impact or risk losing their funding, there has been increased attention on measuring the social and economic outcomes they achieve. This is a significant shift. Whereas in the past, funding arrangements were often based on inputs (the amount of funding) or outputs (number of clients served; the services they received), outcome measurement increasingly focuses attention on the benefits that organisations produce to improve the quality of life for individuals (and communities). While better outcome measurement has the potential to benefit governments, commissioners, service providers and the people they serve, it is a demanding and complex task that poses major challenges.

New Managerialism requires a tangible means of measuring impact in order to function properly. If efficiency and effectiveness are to be demonstrated, then performance must somehow be *recorded*. For homelessness services this means they need to record what happens at the individual level if they are to demonstrate that they are working efficiently. This creates an emphasis on recording how an individual is 'positively changed' by a homelessness service intervention (Lyon-Callo, 2000; Dordick, 2002; Löfstrand, 2010; Hansen-Löfstrand and Juhila, 2012). The requirements of New Managerialism thus combine with a wider political and cultural tendency to reduce homelessness to individual pathology, downplaying or dismissing possible structural causation. Homelessness services are also increasingly defined as successful if they move homeless people towards actions deemed productive by the state (Wacquant, 2009). Outcome measurement is thus driven by New Managerialism, but also in ways that reflect mainstream political and social views of who and what homeless people are.

In the last decade, homelessness service providers, service commissioners and policy-makers in the UK and elsewhere have focused on developing tools that can reliably measure service and programme outcomes (Homeless Link, 2007). One tool that has gained a considerable amount of exposure is the Homelessness Outcomes Star (HOS), which is designed to "both support and measure change when working with [homeless] people" (MacKeith, 2011, p.2). Despite the rapid take-up of the HOS, there has been "no formal research on the usefulness of the Star" (op cit., p.1), nor has the theoretical model that the HOS is predicated on been subject to critical examination.

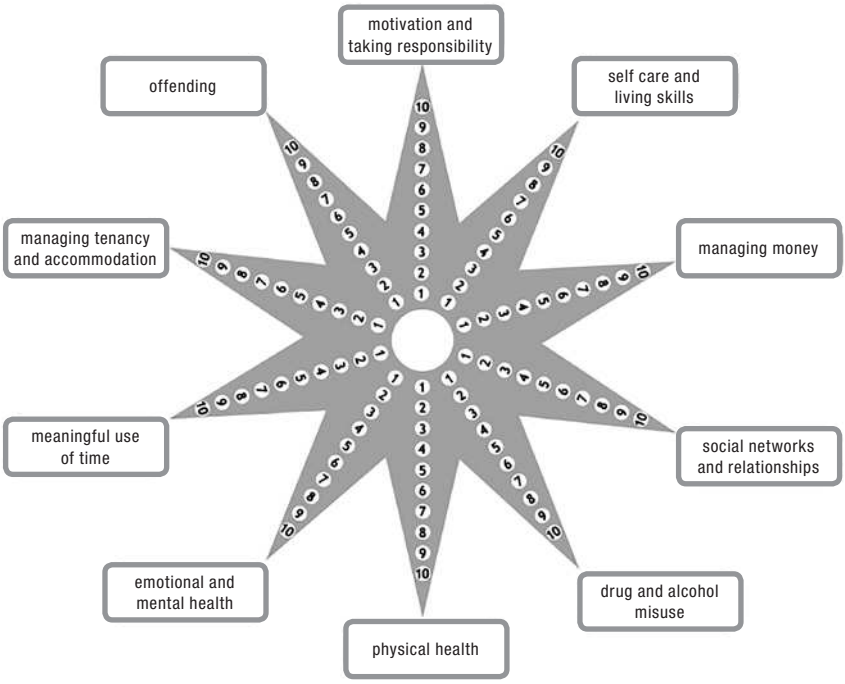
In this paper we examine the HOS. We start by describing the HOS. We then examine the theoretical tenets on which it is based, before turning our attention to the empirical approach it uses to measure outcome results. With respect to the theoretical framework, we argue that both the Journey of Change stage model that underpins the HOS, and the application of Motivational Interviewing (MI) as the key technique for facilitating behavioural change, lack clear empirical support. With respect to the measurement of client outcomes, our view is that claims to the effect that the HOS is a valid outcome measurement tool are greatly overstated – we found no empirical support for the psychometric properties of the HOS. The lack of clear evidence is a significant problem but it is not the only issue. A further concern is that the way the HOS is conceptualized and implemented appeals to a particular conception of human behaviour that assumes change is the result of "careful (cognitive) consideration of alternatives and their consequences" (Littell and Girvin, 2002, p.251). Individuals who score poorly on the HOS (or whose score does not improve) are in danger of being labelled unmotivated and irresponsible, while those who score well provide support for policies that consider ameliorating homelessness as best achieved by reforming individuals. The overall intent of the HOS to promote greater respect and understanding of homeless people as service users is a positive one. However, homelessness agencies that use the HOS and focus solely on changing individual behaviour, risk reinforcing an overly simplistic discourse that sees individual pathology as the root cause of homelessness.

The Homelessness Outcome Star (HOS)

The HOS was developed in the United Kingdom by Triangle Consulting. Triangle Consulting was originally commissioned by St Mungo's, a homelessness service provider in London. The HOS was developed in an attempt to improve the metrics available to the UK homelessness sector, with the intent being to enhance internal management data and to give homelessness service providers viable outcome measures to secure and sustain funding. At the time of writing there are

over 20 versions of the outcomes star, and all bar the HOS are owned and licensed exclusively by Triangle Consulting. The HOS is widely used in the UK, with over 20% of homelessness agencies surveyed by Homeless Link using the Star (Homeless Link, 2011 cited in MacKeith, 2011). The HOS is also being used internationally, with countries such as Australia, France, Italy and Denmark reportedly using it (MacKeith, 2011).

Figure 1: The Homelessness Outcomes Star



Copyright: Triangle Consulting Social Enterprise Limited and the London Housing Foundation

The HOS is based on 10 items, which are rated on a 10-point scale (Figure 1). The 10 items are: motivation and taking responsibility; self-care and living skills; managing money and personal administration; social networks and relationships; drug and alcohol misuse; physical health; emotional and mental health; meaningful use of time; managing tenancy and accommodation; and offending. Individuals make an initial assessment on the 10-point scale. Individuals and their caseworker then discuss and score their subsequent progress over time. Individual scores are then calculated for each domain and the scores are summated and averaged to

provide an overall measure of change. Scores within set ranges are banded into five classifications: 'stuck' (1-2), 'accepting help' (3-4), 'believing' (5-6), 'learning' (7-8) and 'self-reliance' (9-10).

The goals of the HOS are defined as threefold. First, it is designed to actively inform and monitor casework, so that the benefits of support provided by a homelessness service can be monitored by people using a service, frontline staff and managers. Second, at the management level, the HOS is intended to generate benefits in organisational learning, which may in turn lead to service design modifications. Third, the HOS is designed to enhance relationships with commissioners by providing 'statistical' proof of service effectiveness (Burns *et al.,* 2008). This third function reflects the UK origins of the HOS, where, as a direct result of Thatcherism and the implementation of New Managerialism in the public and charitable/NGO sectors, homelessness services are funded through competitive commissioning processes.

Using simple metrics, such as whether or not a formerly homeless person is housed or employed, can make a homeless service look inefficient. This is because progress can take time – i.e., someone may still not be sustainably housed, or in work, a year after starting to use a service, but they may be much *closer* to those goals (Pleace and Quilgars, 2013). Alternative metrics, which are at the core of the HOS, can be used to show funders that even if desired end goals have not been reached, *progress* is being made (Burns *et al.*, 2008).

Theoretical Foundations

The HOS is based on principles drawn from Participatory Action Research. It is explicitly "rooted in a conception of the person receiving the service as an active agent in their own life, not a passive sufferer of an affliction that the professional with their expertise and knowledge will cure" (MacKeith, 2011, p.6). The HOS is thus consistent with a client-centred approach, as has been at the heart of social work practice since the 1950s.

The client-centred approach that underpins the HOS builds on the idea of self-determination, first articulated as one of the seven core casework principles by Felix Biestek in his seminal text *The Casework Relationship* (1957). Indeed, as Harris and Andrews (2013, p.1) note, the HOS explicitly acknowledges the "significance of personal motivation and agency for a service user in achieving sustainable change in their journey towards independence and choice in critical areas of their lives." In this context, the use of Motivational Interviewing (MI) to facilitate behavioural change is consistent with, and builds on, earlier ideas of self-determination. MI was first described by Miller (1983) and further elaborated by Miller and Rollnick (1991)

as "a person-centered, goal orientated approach for facilitating change through exploring and resolving ambivalence" (Miller, 2006, p.138). MI is a collaborative, non-judgmental, strengths-based approach that seeks to enhance "intrinsic motivation to change" (Wahab, 2010, p.198).

In the context of the HOS, MI is strongly influenced by the transtheoretical model of Prochaska and DiClemente (1982; 1983; 1984), which gained widespread popularity in the fields of health psychology and addiction in the 80s and 90s. The transtheoretical model conceptualizes behaviour change as a process with various stages. Stages represent distinct categories along a "continuum of motivational readiness" (Wahab, 2010, p.198). As noted, the HOS identifies five stages of change – stuck, accepting help, believing, learning and self-reliance – that broadly correspond to the five stages identified by Prochaska and DiClemente. According to the HOS, stage one (1-2 on the self-report scale) is the stage in which an individual is 'stuck' and not considering any possibility of change. Accepting help (3-4 on the self-report scale) is the stage defined by the recognition that they need 'someone else to sort things out'. Believing (5-6 on the self-report scale) is a state characterized by an 'internal shift towards taking responsibility'. Stage four is characterized by learning how to do things independently (7-8 on the self-report scale), and stage five (9-10 on the self-report scale) is defined by an individual's capacity to manage without any assistance (see Burns *et al.,* 2013). The intention of MI is to support people to move from being 'stuck to being self-reliant and independent'. Self-reliance is defined in global terms, and the ultimate goal, at least theoretically, is achieving 9 or 10 for each of the ten HOS domains (Burns *et al.*, 2008).

While MI and the transtheoretical model are not necessarily the same thing, both focus on individuals as the key agents of change. The client focus of the HOS is an important contribution, precisely because it seeks to empower and motivate individuals to improve their circumstances. In theory at least, this can assist agencies to more thoughtfully and actively engage and support people in the "co-production of their own futures" (MacKeith, 2011, p.2). Interventions in which agencies provide information and guidance to assist individuals have been described as a form of 'weak paternalism'. Their use has been justified where the intervention assists individuals to achieve their own objectives (Parsell and Marston, 2016). But the HOS could also be seen as reflecting a stronger form of paternalism, in that by predetermining areas of change, as well as ultimate goals, it effectively *determines* "what people see as their own interests" (op cit., p.3).

The Homelessness Outcome Star in Practice

Questions of neutrality in outcome goals and measurement

The HOS is sometimes presented as an unqualified success by those who advocate for its use (Burns *et al.*, 2008; MacKeith, 2014). In the UK, the HOS has been described as enjoying 'enormous popularity' (MacKeith, 2009). In 2014, the range of outcome stars that followed in the wake of the HOS were described as making 'new conversations possible' and as giving 'new hope' to service users (MacKeith, 2014).

A number of studies on the use of the HOS and related outcome stars have been carried out. Those that focused specifically on the HOS include Australian research (Harris and Andrews, 2013), a UK study of 25 service providers, 11 of which self-identified as working with single homeless people (Burns *et al.*, 2008) and a small American study (Petersen *et al.,* 2014). All of these pieces of research come to similar conclusions. The HOS is presented as offering meaningful metrics for monitoring progress over time. Further, the HOS is described as providing management information that enables service providers to monitor how well they are performing. Finally, the HOS is also shown as offering data that show service efficiency (Burns *et al.*, 2008). Much of this research is qualitative, centring on reports of how the HOS has benefitted organisations and individuals.

Indeed, the HOS places great emphasis on an individual's perception of 'where they are at' in relation to a series of *specific* goals. They should, on achieving a score of 9 or 10, be using their time meaningfully, demonstrate good emotional, physical and mental health, have positive social networks and relationships, be motivated and be taking responsibility. Homelessness, on these measures, starts to look like a matter of individual pathology that can only be addressed by changing behaviour in *set* ways – i.e., being an economically productive and socially engaged consumer (Carlen, 1996; Dordick, 2002; Wacquant, 2009). Thus, the use of specific pre-determined goals potentially disrupts "long standing values on one's right to decide what constitutes a good life and how one ought to live" (Parsell and Marston, 2016, p.3).

Thus, a key test of the HOS is the extent to which progress against the goals set by the ten points in the star reflects and relates to the kind of progress that homeless people actually wish to make in their lives. Another test is whether HOS delivers meaningful management information both for individual workers and at management level. And finally, HOS must be assessed on whether it provides outcome monitoring that is convincing to commissioners, donors and governments.

We now examine the application of the HOS to two domains – housing and drug use. We use the following discussion to argue that if the link between behaviour and social context is ignored, the heavy focus on individual motivation, while important, can have potentially negative effects and reinforce prevailing images of homeless people as incompetent, wilful and dysfunctional individuals. Indeed, as much as self-determination is a key principle guiding social work practice, so too is the recognition that external conditions influence individual behaviour. While the HOS is presented and seen as means of empowerment for homeless people (Burns *et al*., 2008), we argue that the language, concepts and approach within the HOS may actually undermine empowerment by ignoring or downplaying the structures and systems that contribute to individual problems.

In the domain of 'managing tenancy and accommodation', people are 'stuck' (stage 1) because they "are not *able* or not *willing* to comply with the rules and regulations" (Burns *et al.,* 2013, p.23; our italics). Through the application of MI, homeless people are supposedly empowered to make changes that eventually lead to the ability to live independently. However, what is notable in the HOS is that individuals that get a low score on this measure are 'stuck' – effectively depicted either as irresponsible because they are 'not willing to comply', or incompetent because they are 'not able to comply' with the 'rules and regulations'. As an individual progresses further along the 'Journey of Change', improvements in their housing circumstances occur *only* because they realize they have to 'make changes, and are motivated to do so'.

What is missing from the HOS is any sensitivity to housing and labour markets. The *image* – and it is the image that is important here – in the HOS is of homeless people as individuals who have to be made 'housing ready', in the sense of being willing to change their behaviour. In short, the focus on empowerment as a method by which people gain control over their lives and secure independent living can minimise, if not entirely obscure, the connection between individual housing problems and the way that social inequality and power differentials play out in external domains such as the housing market and the labour market (Dordick, 2002; O'Sullivan, 2008; Busch-Geertsema *et al.*, 2010; Lee *et al.*, 2010; Johnson *et al.*, 2015).

Some of the first experiments with resettlement of long-term homeless people into ordinary housing reported boredom, isolation and the need for treatment as risks to housing sustainment – not a widespread 'inability' to comply with rules or behave in acceptable ways (Dant and Deacon, 1989). There is, moreover, only scant evidence of a significant need for training in how to run a home among most homeless people (Jones *et al*., 2001). If the successes of Housing First (Pleace, 2016) or, indeed, what much of the homelessness sector in the UK regards as good practice tell us anything about housing sustainment, it is that success in housing sustainment *centres* on maximising individual choice and control (Hough and Rice, 2010).

There are similar issues in the domain 'drug and alcohol use'. Homeless people are defined as stuck when they "deny they have a problem" (Burns *et al.*, 2013). From there, 'empowered' and 'motivated' individuals progress towards self-reliance and independence whereby drug use is no longer problematic. We examine the effects of MI on substance misuse behaviour in subsequent pages, but here draw attention to a body of work that specifically examines the nexus between substance misuse and homelessness.

The prevalence of substance misuse is high among some homeless populations, such as young homeless people and those experiencing sustained and recurrent homelessness, with estimates ranging from 20 to 45 percent (Neale, 2001; Fountain and Howes, 2002; Kemp *et al*., 2006). We also know that substance use is often a consequence rather than a cause of homelessness and can exist prior to, during and following homelessness (Johnson *et al.,* 1997; Johnson and Chamberlain 2008; Pleace, 2008). Service outcomes, from abstinence-based services through to Housing First, have never been perfect (Pleace and Quilgars, 1996; Neale, 2001; Padgett, 2007; Pleace, 2008; Pleace and Quilgars, 2013; Rog *et al.,* 2014; Rae and Rees, 2015). However, what is and remains abundantly clear is that abstinence-based interventions have consistently proven to be relatively *ineffective* (Pleace, 2008). It seems that in the HOS, there is no real place for harm reduction, in which choice and control remain with homeless people and which can generate comparably good, if not perfect, outcomes (Pleace and Quilgars, 2013). Again, the absence of change in substance use behaviour or in relapses is framed by the HOS as a failure of *individual motivation*. What is missing is a recognition that the "social and personal resources a person has are instrumental in overcoming dependence" (Hser *et al.,* 2010, p.181).

In each of the remaining eight domains, similar issues are evident – lacking skills or the correct attitudes to take care of themselves, to manage their money effectively, create social networks, improve their physical and emotional health, to meaningfully use their time and to cease offending are all problems to be "overcome through motivation and empowerment, whereby the individual is to accept responsibility for change" (Hansen-Lofstrand and Juhlia, 2012, p.57). In every domain, the effects of structural, biographical and situational factors are ignored.

The meaning of outcomes in the Homelessness Outcomes Star

One aspect of the focus on the individual within the HOS is related to the use of MI. While the focus of MI on enabling individuals to take positive choices is constructive, concerns about the efficacy of MI have led some to question whether its "popularity… may have outstripped its effectiveness" (McMurran, 2009, p.85). Part of the challenge of evaluating MI lies in the fact that it can mean very different things, it is applied in very different ways and it can also have very different aims.

There is evidence that when used to improve engagement with services, MI can work well (Lundahl *et al.,* 2010). For example, using MI to encourage long-term homeless people to use mental health, drug and alcohol services can have good results and there is some evidence around positive behavioural changes resulting from MI (Lundahl *et al.,* 2009; McMurran, 2009).

The key problem with ascribing behavioural change to MI is that we know that behaviour is driven by multiple, fluid variables and that "intention, motivation and behaviour change may fluctuate independently, in various ways and in no particular order" (Littell and Girvin, 2002, p.249). A settled, stable home can have an independent positive effect, which makes homeless people start to behave like other citizens without being 'motivated' to change (Pleace and Quilgars, 2013). Further, there is no evidence to suggest MI is unambiguously effective when used in relation to substance misuse or offending, or in tackling other needs among homeless people (Project MATCH Research Group, 1997; Peterson *et al.,* 2006; Baer *et al.,* 2007; Van Wormer, 2007; McMurran, 2009; Wain *et al.,* 2011).

This last point is important. We suspect part of the reason that studies of MI fail to report consistent results is that MI fails to account adequately for external factors. This point is particularly relevant to homelessness. No matter how motivated individuals are, many factors can remain outside their control (Dordick, 2002).

Finally, a key aspect of the HOS is that it is explicitly informed by a 'coherent theory of change' (Triangle Consulting, 2014a). The Journey of Change is clearly articulated and firmly embedded in the HOS (Planigale and HomeGround Services, 2011). The validity of the Journey of Change rests on the idea of *distinct* stages of change. Some argue the model is evidenced (Prochaska and DiClemente, 1984; Morera *et al.,* 1998; MacKeith, 2011); others argue that the delineation between the stages is not clear (Sutton, 1996; Andresen *et al.,* 2003). We are inclined to agree with Littell and Girvin (2002, p.253) who make the following critique of stages of change models:

> The search for a generic, underlying structure of behaviour change has led to unnecessary reductionism, reliance on a set of categories that do not reflect qualitatively different states, and adherence to assumptions about stage progression that have not been supported.

Following from this, there are two potential problems with the HOS. The first is the *precise* meaning of progression towards a score of 9 or 10 on the ten points of the star. In their study of 10 homeless people, Petersen *et al*. (2014) reported an average progression of 2.02 steps. Citing the guidance for HOS (MacKeith *et al.,* 2008), the authors suggest this would be a "very significant step" (Petersen *et al.,* 2014, p.33). However, the *empirical basis* for this claim is unclear.

The HOS has been praised as a way of tracking progression for homelessness service users and there is a fairly detailed description of what each score means – a three is 'I have had enough of living like this and want things to change', a five is 'I see that I need to do things for myself to get where I want to be' (Burns *et al.,* 2013). But other than a higher score suggesting progression, what – in qualitative terms and, particularly, in quantitative terms – is the precisely measurable *consistent* difference between scores? Does a movement on the HOS from an average score of two to an average score of three indicate genuine change – that, on average, people in a project are moving from not discussing or accepting help with an issue to accepting help with it, as MacKeith argues (2011, p.3-4)? Or, might it simply reflect measurement error, as suggested by Beazley (2011)? Finally, what does it actually mean when someone is, for example, scoring seven on two points of the star, two on seven points, and 10 on one point: how can their total progress be assessed compared to, say, someone scoring four on everything?

At first glance it might seem that questions of reliability and validity have little direct relevance to the HOS. Reliability and validity are tools of an "essentially positivist methodology" (Golafshani, 2003, p.598), but the HOS is positioned as "existential phenomenological approach… [that] challenges the assumptions of absolute truth and objectivity of the traditional positivist, science paradigm" (MacKeith, 2011, p.8).

However, the issues of reliability and validity are important and the reasons for this are quite simple. First, the HOS is described as "tried and tested" and as intended to "support as well as measure change" (MacKeith, 2011, p.8). Second, the HOS uses what is called an 'objective' self-report scale. Outcomes are reported and scored on a scale of 1-10. Progress is reported in changes in scores over time. The collection of apparently *quantitative* data aligned to pre-determined categories (the five stages in the Journey of Change) is a hallmark of a positivistic approach. Further, establishing the validity and reliability of the HOS is crucial, given that the authors suggest it can measure more than just individual outcomes. Star data:

> … can be aggregated for all service users within a project to provide project level outcomes. It can also be aggregated and compared across groups or projects, or nationally (MacKeith, 2011, p.3).

If the outcome results of the HOS are to be trusted, irrespective of what level the outcome measures are applied to, they need to be both credible and defensible. In short, the HOS needs to be able to demonstrate that its 10 measures measure what they are intended to (validity), and the extent to which the results are consistent and stable (reliability).

The issue of validity has not escaped the attention of the developers of the HOS. Indeed, in various publications the developers of the tool make the claim that there is a "growing body of evidence… demonstrating that the outcomes star is… valid as an outcomes measure" (Triangle Consulting, 2014a, p.1), and that "research into the psychometric properties of the star has shown that it performs well as an outcomes measure" (op. cit., p.2).

With respect to the HOS, we suspect these claims are greatly overstated for two reasons. First, we could not find a single peer-reviewed study of the HOS that examined its reliability and validity, or any evidence among commissioned research. Another limitation is that there are no statistical data on how the 10 domains were selected, how they might interrelate, how they take into account the impact of parallel interventions, or how they might relate to hard outcomes such a securing housing, which are of considerable interest to programme funders. Nor is there any published data that relates changes in the measures to various demographic characteristics (e.g., age, gender, ethnicity) or other pertinent factors, such as complexity of need or the duration of homelessness – factors that can significantly influence what services can accomplish.

Finally, we were struck by the absence of any clear indication of how much time should elapse between data collection. MacKeith only suggests "some time later" (2011, p.3). It is unclear if agencies are collecting data using a similar timeframe or, indeed, what an appropriate timeframe is. Thus, any comparisons between projects or organisations that use different timeframes are likely to be flawed. This is a particularly pertinent point when the length of time for which services are offered for varies so much. There is also the question of how much change it is reasonable to expect in a given period of time.

While the evidence supporting the reliability and validity of the HOS is limited (Burns *et al*., 2008; Petersen *et al*., 2014), two peer reviewed papers have been published on the Mental Health Recovery Star, examining its psychometric properties. In a study of 203 working age adults with moderate to severe mental illness who undertook two Recovery Star readings (113 did a third reading), Dickens *et al.* (2011) found that the Recovery Star had high internal consistency but made no comment on its validity other than to state that "little is currently known" and further research into the psychometric properties is "warranted" (p.49). Killaspy *et al.*'s (2012) study of 172 service users and 120 staff from in-patient and community services reported that staff found it to be acceptable to service users and useful for care planning. However, while they found that the tool had good test-retest reliability for the same staff members, inter-rater reliability between different staff members was 'inadequate'. They note that this is a "serious problem in mental health services where staff turnover and multidisciplinary working mean that different members of staff

need to be able to assess service users reliably" (Killaspy *et al.,* 2012, p.69). Because of this, they conclude that they could not support its "recommendation for use as a clinical outcome tool at present" (p.70).

Our second reason for questioning the validity of the HOS is that it relies on self-report data. Self-report tools are popular in the behavioural sciences – they are cheap and relatively easy to administer. However, self-report data are subject to a number of problems. Research suggests that people have different ways of responding to scales (Pollio *et al*., 2006; Tsemberis *et al*., 2007). Additionally, people can lack introspective ability, particularly if they are in crisis, and when it comes to drug use or offending, research shows that people may conceal problems (Pleace, 2008). Finally, there is the problem of social desirability – a tendency to tell workers what they want to hear, both to please them and also to achieve a better response from the service being used (Hutson and Liddiard, 1994; Lyon-Callo, 2000; Dordick, 2002; Pleace, 2008).

It is argued by the authors and advocates of HOS that the interpretative approach to validation means that we have to think of validity in a different way (Triangle Consulting, 2014b). Qualitative researchers have argued persuasively that the ideas of reliability and validity have different meanings in the qualitative research paradigm (Dickens *et al.,* 2011; Killaspy *et al.,* 2012). The problem is that HOS measures outcomes using a basic self-report scale and consequently should be subject to the same sort of scrutiny as any similar quantitative tool.

In short, until independent research demonstrates the psychometric properties of the HOS, it is best to treat the HOS outcome results with caution. That is not to say the HOS does not have a role, and a potentially important one at that. The strength of the tool may not be as an outcomes measurement tool but rather as case management tool. However, case management and outcome measurement are very different things indeed.

Conclusion

The HOS was developed – with good intentions – to enable homelessness services to show efficiency and effectiveness in a new, very challenging, context (Burns *et al*., 2008; MacKeith, 2009; 2014). It promotes a number of important ideas about management information and outcome monitoring. The HOS presents the ideas of consistent, regulated, comparable and, importantly, outcome monitoring that tracks individual progress over time.

Looking at the positives of HOS, it can be said that it attempts to measure many important outcomes. There is a reasonable amount of evidence to say that money management, social networks and relationships, drug and alcohol use, mental and

physical health, offending and threats to housing sustainment are things to watch if someone with complex needs is to exit recurrent or sustained homelessness (Dant and Deacon, 1989; Busch-Geertsema *et al.*, 2010).

If we were to apply the HOS to the evaluation of Housing First services, there would be interesting results. The evidence base says Housing First is good at ending homelessness, but results around health, mental health, drug and alcohol use, nuisance and offending behaviour and social integration can be more mixed (Pleace and Quilgars, 2013). Using the HOS to measure the performance of Housing First approaches, there is a fair chance the HOS would report good results on housing sustainment, but much less success in relation to the other outcomes (Padgett, 2007). Yet, we quickly run into three significant problems. The first issue is measurement consistency. Second, what exactly the measurements mean. Third, and most importantly, the conceptualisation of 'success' in HOS. Housing First is influenced by MI and emphasises active engagement and a recovery orientation, but ultimately it is not *telling* homeless people to change. Further, ideas at the core of Housing First – from harm reduction, choice and control, through to housing as a human right rather than something to be 'earned' (Pleace, 2016) – take Housing First out of sync with the HOS. The personalisation, co-production and choice-led innovations at the core of what is regarded as best practice in the HOS country of origin, the UK, are also examples of disconnects with the internal logic of the HOS.

Imagine if the worker collecting HOS data on a single individual changed. Based on the limited research available, it is likely that the interpretation of the HOS recorded for that individual would be different, irrespective of whether the assessment was done independently by the worker or in collaboration with a homeless person. This foreshadows a deeper problem, which is whether it is really possible to delineate between a HOS score of three and a score of five in a meaningful and robust way. The empirical support for its theoretical framework is ambiguous, there is no contextual data, there is no allowance for needs outside those within the 10 points of the HOS, and there is no allowance for the possibility that attitude, behaviour and willingness to change are *not* the areas that explain homelessness or that need to be changed. Above all, it is not clear, comparing one homelessness service user with another, or one homelessness service with another, what the HOS scores actually *mean*.

Another issue is how *useful* the HOS are for external purposes. From a political perspective, the truly tangible still matters. The homelessness service that will get the funding is the one that sustainably ends long-term homelessness and that has statistics to show that happening – or, better still, experimental research that shows it outperforming the usual homelessness services. That is the primary lesson from the inexorable rise of Housing First at global level, even if that evidence base is not

as solid as it is sometimes presented (Pleace and Quilgars, 2013). In the country in which HOS originated and is most widely used, the UK, the 'evidence' provided by HOS has not prevented deep cuts to homelessness services and the emergence since 2010 of an existential threat to the sector (Homeless Link, 2015). HOS cannot provide robustly evidenced statistical demonstrations of effectiveness, nor, importantly, can it be used to demonstrate systematically that a homelessness service is cost-effective or delivers cost offsets for other services.

As we have discussed in this paper, there are some serious questions to ask about the theoretical tenets that shape data collection in the HOS. Further, as we have highlighted, issues with the quality, meaning and comparability of the information collected require further investigation. The worries are methodological but also cultural, ideological and political, because HOS is ultimately posited on an assumption that individual pathology is the root cause of homelessness, and behavioural modification the only answer. The evidence indicates the complex nature of homelessness and homelessness causation (Kuhn and Culhane, 1998; O'Flaherty, 2010; Culhane *et al.*, 2013), and when held up to the light, ideas of *entirely* behavioural causation and 'cultures' of homelessness fall apart (Burt, 2001; O'Sullivan, 2008; Lee *et al*., 2010). We know that homelessness varies markedly in size, nature and scope between different welfare systems and cultures (Busch-Geertsema *et al.*, 2010; Fitzpatrick and Stephens, 2014; Benjaminsen, 2015). We also know that interventions like Housing First, which emphasize choice and control for homeless people with complex needs, are more effective than those that attempt to regulate or dictate behaviour (Pleace, 2008; 2016). To suggest that individual pathology is unimportant may be a leap too far, but to suggest that individual pathology is the only thing that matters in understanding, preventing and stopping homelessness is, frankly, nonsense.

Imprecision in measurement, including poor delineation between scores, and likely inconsistency in interpretation of scores lead to problems in using the data collected by the HOS in a comparative way. When this is combined with a flawed conceptualisation of homelessness and a theoretical framework that lacks empirical support, claims that the HOS is an "evidence-based tool for supporting and measuring change" (Triangle Consulting[1]) appear exaggerated.

---

1 http://www.outcomesstar.org.uk/homelessness/[16.06.2016]

› References

Andresen, R., Oades, J. and Caputi, P. (2003) The Experience of Recovery from Schizophrenia: Towards an Empirically Validated Stage Model, *Australian and New Zealand Journal of Psychiatry* 37(5) pp.586-594.

Baer, J., Garrett, S., Beadnell, B., Wells, E. and Peterson, P. (2007) Brief Motivational Intervention with Homeless Adolescents: Evaluating Effects on Substance Use and Service Utilization, *Psychology of Addictive Behaviors* 21(4) pp.582-586.

Beazley, P. (2011) The Recovery Star: Is it a Valid Tool?, *The Psychiatrist* 35 pp.196-197.

Benjaminsen, L. (2015) Homelessness in a Scandinavian Welfare State: The Risk of Shelter Use in the Danish Adult Population, *Urban Studies* DOI: 0042098015587818.

Biestek, F. (1957) *The Casework Relationship* (Chicago: Loyola University Press).

Bullen, J. (2015) Governing Homelessness: The Discursive and Institutional Construction of Homelessness in Australia, *Housing, Theory and Society* 32(2) pp.218-239.

Burns, S., MacKeith, J. and Graham, K. (2008) *Research Report. Using the Outcomes Star: Impact and Good Practice* (London: Homeless Link and Triangle Consulting).

Burns, S., Graham, K. and MacKeith, J. (2013) *Outcomes Star: The Star for People with Housing and Other Needs. Users' Guide, 3rd Edition* (Wilbury Villas, Hove: Triangle Consulting Social Enterprise Ltd.).

Burt, M.R. (2001) Homeless Families, Singles and Others: Findings from the 1996 National Survey of Homeless Assistance Providers, *Policy Debate* 12(4) pp.737-780.

Busch-Geertsema, V., Edgar, W., O'Sullivan, E. and Pleace, N. (2010) *Homelessness and Homeless Policies in Europe: Lessons from Research* (Brussels: European Commission).

Carlen, P. (1996) *Jigsaw: A Political Criminology of Youth Homelessness* (Buckingham: Open University Press).

Culhane, D.P., Metraux, S., Byrne, T., Stino, M. and Bainbridge, J. (2013) The Age Structure of Contemporary Homelessness: Evidence and Implications for Public Policy, *Analyses of Social Issues and Public Policy* 13(1) pp.228-244.

Dant, T. and Deacon, A. (1989) *Hostels to Homes? The Rehousing of Single Homeless People* (Aldershot: Avebury).

Dickens, G., Weleminsky, J., Onifade, Y. and Sugarman, P. (2011) Recovery Star: Validating User Recovery, *The Psychiatrist* 36 pp.45-50.

Dordick, G.A. (2002) Recovering from Homelessness: Determining the 'Quality of Sobriety' in a Transitional Housing Program, *Qualitative Sociology* 25(1) pp.7-32.

Exworthy, M. and Halford, S. (1999) Professionals and Managers in a Changing Public Sector: Conflict, Compromise and Collaboration?, in: M. Exworthy and S. Halford (Eds.) *Professionals and the New Managerialism in the Public Sector*, pp.1-17. (Buckingham: Open University Press).

Fitzpatrick, S. and Stephens, M. (2014) Welfare Regimes, Social Values and Homelessness: Comparing Responses to Marginalised Groups in Six European Countries, *Housing Studies* 29(2) pp.215-234.

Fountain, J. and Howes, S. (2002) *Home and Dry? Homelessness and Substance Use in London* (London: National Addiction Centre).

Golafshani, N. (2003) Understanding Reliability and Validity in Qualitative Research, *The Qualitative Report* 8(4) pp.597-607.

Greenway, M. (2001) The Emerging Status of Outcome Measurement in the Nonprofit Human Services, in: P. Flynn and V. Hodgkinson (Eds.) *Measuring the Impact of the Nonprofit Sector*, pp.217-229. (New York: Klumer Academic/Plenum Publishers).

Hansen-Lofstrand, C. and Juhlia, K. (2012) The Discourse of Consumer Choice in the Pathways Housing First Model, *European Journal of Homelessness* 6(2) pp.47-68.

Harris, L. and Andrews, S. (2013) *Implementing the Outcomes Star in a Multi-dsiciplinary Environment* (St. Kilda: The Salvation Army).

Homeless Link (2007) *The Use of Outcomes Measurement Systems within Housing and Homelessness Organisations* (London: Homeless Link).

Homeless Link (2015) *Support for Single Homeless People in England: Annual Review 2015* (London: Homeless Link).

Hough, J. and Rice, B. (2010) *Providing Personalised Support to Rough Sleepers: An Evaluation of the City of London Pilot* (York: Joseph Rowntree Foundation).

Hser, Y., Longshore, D. and Anglin, A. (2010) The Lifecourse Perspective on Drug Use: A Conceptual Framework for Understanding Drug Use Trajectories, in: K. Van Wormer and B. Thyer (Eds.) *Evidence Based Practice in the Field of Substance Abuse: A Book of Readings,* pp.171-196. (Los Angeles: Sage).

Hutson, S. and Liddiard, M. (1994) *Youth Homelessness: The Construction of a Social Issue* (London: Macmillan).

Johnson, G. and Chamberlain, C. (2008) Homelessness and Substance Abuse: Which Comes First?, *Australian Social Work* 61(4) pp.342-356.

Johnson, T., Freels, S., Parsons, J. and Vangeest, J. (1997) Substance Abuse and Homelessness: Social Selection or Social Adaptation, *Addiction* 92(4) pp.437-445.

Johnson, G., Scutella, R., Tseng, Y. and Wood, G. (2015) *Entries and Exits from Homelessness: A Dynamic Analysis of the Relationship between Structural Conditions and Individual Characteristics* (Melbourne: Australian Housing and Urban Research Institute).

Jones, A., Quilgars, D. and Wallace, A. (2001) *Life Skills Training for Homeless People: a Review of the Evidence* (Edinburgh: Scottish Homes).

Kemp, P., Neale, J. and Robertson, M. (2006) Homelessness among Problem Drug Users: Prevalence, Risk Factors and Trigger Events, *Health and Social Care in the Community* 14(4) pp.319-328.

Killaspy, H., White, S., Taylor, T. and King, M. (2012) Psychometric Properties of the Mental Health Recovery Star, *The British Journal of Psychiatry* 201(1) pp.65-70.

Kuhn, R. and Culhane, D. (1998) Patterns and Determinants of Public Shelter Utilization among Homeless Adults in Philadelphia and New York City, *Journal of Policy Analysis and Management* 17(1) pp.23-43.

Lee, B., Tyler, K. and Wright, J. (2010) The New Homelessness Revisited, *Annual Review of Sociology* 36 pp.501-521.

Littell, J. and Girvin, H. (2002) Stage of Change: A Critique, *Behavior Modification* 26(2) pp.223-273.

Löfstrand, C. (2010) Reforming the Work to Combat Long-term Homelessness in Sweden, *Acta Sociologica* 53(1) pp.19-34.

Lyon-Callo, V. (2000) Medicalizing Homelessness: the Production of Self-blame and Self-governing within Homeless Shelters, *Medical Anthropology Quarterly* 14(3) pp.328-345.

Lundahl, B., Kunz, C., Brownwell, C., Tollefson, D. and Burke, B. (2010) A Meta-analysis of Motivational Interviewing: Twenty-Five Years of Empirical Studies, *Research on Social Work Practice* 20(2) pp.137-160.

MacKeith, J. (2009) Why Hard-Nosed Agencies Measure Soft Outcomes, *Sitra Bulletin* 221 pp.10-11.

MacKeith, J. (2011) The Development of the Outcomes Star: A Participatory Approach to Assessment and Outcome Measurement, *Housing Care and Support*: *A Journal on Policy, Research and Practice,* 14(3) pp.1-14.

MacKeith, J. (2014) Assessing the Reliability of the Outcomes Star in Research and Practice, *Housing, Care and Support: A Journal on Policy, Research and Practice* 17(4) pp.188-197.

MacKeith, J., Burns, S. and Graham, K. (2008) *The Outcomes Star Users Guide* (2nd Ed) (UK: Triangle Consutlting Social Enterprise).

McMurran, M. (2009) Motivational Interviewing with Offenders: A Systematic Review, *Legal and Criminal Psychology* 14(1) pp.83-100.

Miller, W. (1983) Motivational Interviewing with Problem Drinkers, *Behavioral Psychotherapy* 11(2) pp.147-172.

Miller, W. (2006) Motivational Factors in Addictive Behavior, in: W. Miller and K. Carroll (Eds.) *Rethinking Substance Abuse: What the Science Shows*, pp.134-150. (New York: Guilford Press)

Miller, W. and Rollnick, S. (1991) *Motivational Interviewing: Preparing People to Change Addictive Behaviour* (New York: Guilford Press).

Morera, O., Johnson, T., Freels, S., Parsons, J., Crittenden, K., Flay, B. and Warnecke, R. (1998) The Measure of Stages of Readiness to Change: Some Psychometric Considerations, *Psychological Assessment* 10(2) pp.182-186.

Neale, J. (2001) Homelessness Amongst Drug Users: A Double Jeopardy Explored, *The International Journal of Drug Policy* 12(4) pp.353-369.

O'Flaherty, B. (2010) Homelessness as Bad Luck: Implications for Research and Policy, in: I. Gould Ellen and B. O'Flaherty (Eds.) *How to House the Homeless* (New York: Russell Sage Foundation).

O'Sullivan, E. (2008) Pathways Through Homelessness: Theoretical and Policy Implications, in: J. Doherty and B. Edgar (Eds.) *'In My Caravan, I Feel Like Superman': Essays in Honour of Henk Meert, 1963-2006* (Brussels: FEANTSA / University of St. Andrews: Centre for Housing Research).

Padgett, D. (2007) There's no Place like (a) Home: Ontological Security among Persons with a Serious Mental Illness in the United States, *Social Science and Medicine* 64(9) pp.1925-1936.

Parsell, C. and G. Marston (2016) Supportive Housing: Justifiable Paternalism?, *Housing, Theory and Society* 33(2) pp.195-216.

Peterson, P., Baer, J., Wells, E., Ginzler, J. and Garrett, S. (2006) Short-term Effects of Brief Motivational Intervention to Reduce Alcohol and Drug Risk among Homeless Adolescents, *Psychology of Addictive Behaviors* 20(3) pp.254-264.

Petersen, V., Ellis, P., Lorenz, R. and Armbrecht E. (2014) An Interventional Study of Guiding Homeless Persons to Self-reliance using the Outcomes Star for Homelessness, *Critical Scholars Review* 7(1) pp.30-41.

Planigale, M. and HomeGround Services (2011) *Measurement of Client Outcomes in Homelessness Services: Literature Review* (Melbourne: HomeGround Services).

Pleace, N. (2008) *Effective Services for Substance Misuse and Homelessness in Scotland: Evidence from an International Review* (Edinburgh: Scottish Government).

Pleace, N. (2016) *Housing First Guide Europe* (Brussels: FEANTSA).

Pleace, N. and Quilgars, D. (1996) *Health and Homelessness in London: A Review* (London: The Kings Fund).

Pleace, N. and Quilgars, D. (2013) *Improving Health and Social Integration through Housing First: A Review* (Brussels: FEANTSA).

Pollio, D., North, C., Eyrich, K., Foster, D.A. and Spitznagel, E.L. (2006) A Comparison of Agency-based and Self-report Methods of Measuring Services across an Urban Environment by a Drug-abusing Homeless Population, *International Journal of Methods in Psychiatric Research* 15(1) pp.46-56.

Prochaska, J. and Diclemente, C. (1982) Transtheoretical Therapy: Toward a More Integrative Model of Change, *Psychotherapy: Theory, Research and Practice* 19(3) pp.276-288.

Prochaska, J. and Diclemente, C. (1983) Stage and Processes of Self-change of Smoking: Towards an Integrative Model of Change, *Journal of Consulting and Clinical Psychology* 51(3) pp.390-395.

Prochaska, J. and Diclemente, C. (1984) *The Transtheoretical Approach: Crossing Traditional Boundaries of Therapy* (Malabar, Fl: Krieger).

Project MATCH Research Group (1997) Matching Alcoholism Treatment to Client Hetrogeneity: Project Match Posttreatment Drinking Outcomes, *Journal of Studies on Alcohol* 58(1) pp.7-29.

Rae, B. and Rees, S. (2015) The Perceptions of Homeless People regarding their Healthcare Needs and Experiences of Receiving Health Care, *Journal of Advanced Nursing* 71(9) pp.2096-2107.

Rog, D., Marshall, T., Dougherty, R., George, P., Daniels, A., Ghose, S. and Delphin-Rittmon, M. (2014) Permanent Supportive Housing: Assessing the Evidence, *Psychiatric Services* 65(3) pp.287-294.

Sutton, S. (1996) Can 'Stages of Change' Provide Guidance in the Treatment of Addictions? A Critical Examination of Prochaska and DiClemente's model, in: G. Edwards and C. Dare (Eds.) *Psychotherapy, Psychological Treatment and Addictions*, pp.189-205. (Cambridge: Cambridge University Press).

Triangle Consulting (2014a) *Briefing: The Outcomes Star: Unpacking the Evidenc*e (UK: Triangle Consulting Social Enterprise).

Triangle Consulting (2014b) *The Validity of the Outcomes Star as a Tool for Promoting and Measuring Change* (London: Triangle Consulting Social Enterprise).

Tsemberis, S., McHugo, G., Williams, V., Hanrahan, P. and Stefancic, A. (2007) Measuring Homelessness and Residential Stability: The Residential Time-line Follow-back, *Inventory Journal of Community Psychology* 35(1) pp.29-42.

Van Wormer, K. (2007) Motivational Interviewing: A Theoretical Framework for the Study of Human Behavior and the Social Environment, *Advances in Social Work* 8(1) pp.19-29.

Wacquant, L. (2009) *Punishing the Poor: The Neoliberal Government of Social Insecurity* (Durham, NC: Duke University Press).

Wahab, S. (2010) Motivational Interviewing and Social Work, in: K. Van Wormer and B. Thyer (Eds.) *Evidence Based Practice in the Field of Substance Abuse: A Book of Readings*, pp.197-224. (Los Angeles: Sage).

Wain, R., Wilbourne, P., Harris, K., Pierson, H. and Teleki, J. (2011) Motivational Interview Improves Treatment Entry for Homeless Vets, *Drug and Alcohol Dependence* 115(1-2) pp.113-119.