Psychiatry Research 217 (2014) 240–247

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Development of item bank to measure deliberate self-harm behaviours: Facilitating tailored scales and computer adaptive testing for specific research and clinical purposes

Shane Latimer [a], Tanya Meade [a,*], Alan Tennant [b]

[a] School of Social Sciences and Psychology, University of Western Sydney, Penrith, New South Wales, Australia
[b] Department of Rehabilitation Medicine, Faculty of Medicine and Health, University of Leeds, Leeds, UK

ARTICLE INFO

Article history:
Received 26 September 2013
Received in revised form
7 February 2014
Accepted 11 March 2014
Available online 20 March 2014

Keywords:
Deliberate self-harm
Self-harm
Non-suicidal self-injury
Rasch measurement

ABSTRACT

The purpose of this study was to investigate the application of item banking to questionnaire items intended to measure Deliberate Self-Harm (DSH) behaviours. The Rasch measurement model was used to evaluate behavioural items extracted from seven published DSH scales administered to 568 Australians aged 18–30 years (62% university students, 21% mental health patients, and 17% community members). Ninety four items were calibrated in the item bank (including 12 items with differential item functioning for gender and age). Tailored scale construction was demonstrated by extracting scales covering different combinations of DSH methods but with the same raw score for each person location on the latent DSH construct. A simulated computer adaptive test (starting with common self-harm methods to minimise presentation of extreme behaviours) demonstrated that 11 items (on average) were needed to achieve a standard error of measurement of 0.387 (corresponding to a Cronbach's Alpha of 0.85). This study lays the groundwork for advancing DSH measurement to an item bank approach with the flexibility to measure a specific definitional orientation (e.g., non-suicidal self-injury) or a broad continuum of self-harmful acts, as appropriate to a particular research/clinical purpose.

1. Introduction

Early theorists have defined Deliberate Self-harm (DSH) (also referred to as self-harm) as an act of bodily self-harm that is intentional, direct and immediate (Babiker and Arnold, 1997; Kreitman, 1977) with a non-fatal outcome (Morgan, 1979). In more recent times, several specific DSH definitions have been proposed that can be distinguished according to the dimensions of method, intent, lethality and outcome (Ougrin and Zundel, 2009), ranging from a narrow set of visible tissue damage acts performed in the absence of a desire to die (Nock, 2010) to broad spectrums of self-injurious behaviours with multiple intentions (Skegg, 2005). There is on-going debate about the relative merits of the various definitions (De Leo, 2011; Wilkinson and Goodyer, 2011), and some researchers argue that no clear picture of the epidemiology of DSH can be gained until clinicians and researchers agree on a definition of DSH (Rodham and Hawton, 2009).

An example of a narrow DSH definition (covering visible tissue damage acts with non-suicide intent) is Non-Suicidal Self-Injury (NSSI) (Nock, 2010). An independent mental health disorder based on the NSSI definition has been included in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013) as a condition requiring further study. The argument for NSSI as a disorder is based, in part, on the proposition that the methods of DSH most associated with NSSI (viz., mild to moderate forms of visible tissue damage) (Wilkinson and Goodyer, 2011) may form a distinct grouping on a latent continuum of self-harmful behaviours (Ougrin and Zundel, 2009).

An example of a less restricted definition of DSH includes self-poisoning and self-injury, irrespective of suicide intent or type of motivation (Hawton and James, 2005). Many researchers and clinicians argue the merits of adopting a broader definition of DSH because of the difficulties in reliably measuring intent (Ougrin and Zundel, 2009), the lack of support for inferring intent from choice of method (Nada-Raja et al., 2004), and the frequent situation of suicide and non-suicide related self-harm occurring in the same individual (Nock et al., 2006).

Regardless of the specific definition of DSH, it is common practice to measure engagement in self-harm by interview

* Correspondence to: School of Social Sciences and Psychology, University of Western Sydney, Locked Bag 1797, Penrith South DC, New South Wales 1797, Australia. Tel.: +61 2 9772 6266; fax: +61 2 9772 6736.
E-mail address: t.meade@uws.edu.au (T. Meade).

http://dx.doi.org/10.1016/j.psychres.2014.03.015
0165-1781/

schedules and self-report tests that include a behavioural scale comprised of specific methods of self-harm (Borschmann et al., 2012). Scale developers select items that accord with their adopted conceptualisation of DSH. Scales that conform to the narrow NSSI definition are generally restricted to items related to visible tissue damaging acts (e.g., cutting, bruising, scratching and burning) (Gratz, 2001). Scales that are consistent with less restrictive definitions of DSH are likely to combine visible tissue damaging methods with highly dangerous methods (e.g., swallowing dangerous substances) (Croyle and Waltz, 2007), self-harm without visible injury (e.g., exercised an injury to cause harm) (Vrouva et al., 2011), lack of self-care (e.g., taking too little medication to cause harm) (Nixon et al., 2002), and deliberate recklessness to cause harm (e.g., driven recklessly to cause harm) (Sansone et al., 1998).

Most published DSH behavioural scales exhibit good content validity (at least within their particular definitional orientation), reliability, and external validity (Borschmann et al., 2012). Our recent investigation of six frequently used DSH scales found strong evidence for the unidimensionality of the item sets contained in each scale (Latimer et al., 2013). It should be noted that some published scales exhibit high levels of local dependency (i.e., a person's response on one item influences their response on another item) (Latimer et al., 2013) thereby inflating estimates of reliability (Boyle, 1991).

Consistent with their overall sound psychometric properties, published DSH behavioural scales have made a major contribution to advancing our understanding of DSH. They assist clinicians to accurately identify the range of past methods which is the recommended first step in the assessment of DSH (Skegg, 2005), and they are preferred to open-ended questions which are likely to under-estimate the range of self-harm acts (Nock, 2010). In research studies, they inform DSH prevalence rates which are based on the endorsement of at least one method of self-harm (Plener et al., 2009), and a count of the different methods over a person's lifetime appears to provide a very useful estimate of a person's location on a latent DSH construct with high scores indicating a progression to more extreme forms of DSH (Latimer et al., 2013). Their application in both clinical and research settings is likely to increase with the emerging evidence that a count of past methods of DSH has the strongest association with psychopathology (compared to frequency and recency of any one method) (Nock et al., 2006), and it appears to be the best single predictor of future DSH (Glenn and Klonsky, 2011).

However, there are several shortcomings associated with published DSH behavioural scales. First, existing scales cover different combinations of methods (as a consequence of being orientated to specific definitions of DSH) thus preventing the comparison of prevalence rates across studies and equating of clinical cut-off scores (Gratz, 2001; Rodham and Hawton, 2009). Second, some scales are based on narrow definitions of DSH and provide less information about the range of past and present methods of DSH compared to scales based on more expansive definitions (Latimer et al., 2012). Third, concerns have been raised about the potential for scales to cause distress (or arouse curiosity) as a result of the presentation of a fixed set of methods (a necessary feature of traditional pencil and paper formats) to all respondents regardless of their level of experience in DSH (Patton et al., 1997; Zetterqvist et al., 2013). Consequently, ethics committees may be reluctant to approve research designs involving a checklist of specific methods (Swannell et al., 2014) and more research is required to fully inform the risks associated with in-depth DSH assessments (Reynolds et al., 2006; Whitlock et al., 2013). Fourth, many DSH scale developers (Croyle and Waltz, 2007; Nock et al., 2007) do not report Cronbach's Alpha for behavioural items thereby making it difficult to researchers and clinicians to select a scale with the required level of measurement precision for group level research (Cronbach's Alpha greater than 0.70) or for individual level research/clinical assessment (Cronbach's Alpha greater than 0.85) (Ponterotto and Ruckdeschel, 2007).

It may be possible to advance the measurement of specific DSH methods by the development of a unidimensional item bank based on modern test theory (Hambleton et al., 2005). This approach has the potential to maintain the significant past achievements of the published scales while addressing at least some of the limitations that impact on their use in both clinical and research settings. First, a DSH item bank may allow the extraction of tailored scales (covering different combinations of DSH) with the same raw score for each person location on the latent DSH construct (Boekkooi-Timminga, 1990; Hambleton et al., 2005) thus allowing for the direct comparison of prevalence rates and clinical cut-off scores. Second, it may be possible to tailor extracted scales to a particular clinical purpose (e.g., escalation from mild to extreme forms of visible tissue damage) (Favazza, 2006) or a specific research task (e.g., relationship between visible tissue damage acts and highly dangerous methods) (Nada-Raja et al., 2004) rather than to a specific definition of DSH. Third, a DSH item bank may facilitate computer adaptive testing (CAT) applications (Hambleton et al., 2005) to minimise the presentation of items that are beyond the experience of the person taking the test (e.g., extreme methods that she/he has not even thought about). This innovation would be consistent with the recommendation for a graduated assessment of DSH (Nock, 2010), and would address ethical concerns about pencil and paper formats (Patton et al., 1997; Zetterqvist et al., 2013). Fourth, it may be possible to configure DSH CAT applications to provide specific levels of measurement precision (Choi et al., 2012a) that correspond to the values of Cronbach's Alpha required for various research and clinical tasks (Ponterotto and Ruckdeschel, 2007).

A unidimensional DSH item bank (and associated applications of tailored scale extraction and CAT) requires the calibration of a sufficiently expansive set of DSH methods on the same latent DSH construct (Hambleton et al., 2005). We have successfully co-calibrated six published DSH scales on a common measurement metric for the purpose of providing raw score conversion tables (Latimer et al., 2012). An additional outcome from the co-calibration is a pool of candidate items for a DSH item bank. The item pool covers all three accepted groupings of DSH behaviours proposed by Skegg (2005), namely, DSH by the use of dangerous methods, DSH by the use of visible tissue damage, and DSH by the use of other methods without visible injury (including lack of self-care and risky behaviours to cause harm). All such items are expected to relate the same latent DSH construct consistent with a continuum model of DSH behaviours (Nock, 2010).

The major objective of the present study was therefore to develop an item bank covering specific methods of self-harm extracted from published tests of DSH. A second objective was to investigate the utility of the item bank by extracting tailored scales and by simulating a CAT application.

2. Methods

2.1. Selection of candidate items

In total, 98 candidate items related to specific methods of self-harm were extracted from seven published DSH scales, namely: Self-Injury Questionnaire Treatment Related (SIQTR) (Claes and Vandereycken, 2007), Self-Injurious Thoughts and Behaviours Interview (SITBI) (Nock et al., 2007), Deliberate Self-Harm Inventory (DSHI) (Gratz, 2001), Inventory of Statements About Self-Injury (ISAS) (Klonsky and Glenn, 2009), Self-Harm Information Form (SHIF)