

---

## **INCREASE: Development of an Inventory to Characterize Recorded Mental Health Recovery Narratives**

Llewellyn-Beardsley, Joy<sup>1</sup>  
Barbic, Skye<sup>2</sup>  
Rennick-Egglestone, Stefan<sup>1</sup>  
Ng, Fiona<sup>1</sup>  
Roe, James<sup>3</sup>  
Hui, Ada<sup>1</sup>  
Franklin, Donna<sup>4</sup>  
Deakin, Emilia<sup>1</sup>  
Hare-Duke, Laurie<sup>1</sup>  
Slade, Mike<sup>1</sup>

<sup>1</sup>School of Health Sciences, Institute of Mental Health, University of Nottingham, Nottingham, United Kingdom

<sup>2</sup>Department of Occupational Science and Occupational Therapy, Faculty of Medicine, University of British Columbia, Vancouver, Canada

<sup>3</sup>National Institute for Health Research, ARC East Midlands, University of Nottingham, Nottingham, United Kingdom

<sup>4</sup>NEON Lived Experience Advisory Panel, Nottingham, United Kingdom



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

### **KEYWORDS**

Mental Health, Recovery Narratives, Content Analysis, Diversity, Inventory, Reliability

---

### **Abstract**

**Objective:** Mental health recovery narratives are increasingly used in clinical practice, public health campaigns, and as directly-accessed online resources. No instrument exists to describe characteristics of individual recovery narratives. The aims were to develop and evaluate an inventory to characterize recorded recovery narratives.

**Research Design and Methods:** A preliminary version of the Inventory of Characteristics of Recovery Stories (INCREASE) was generated from an existing theory-base. Feasibility and acceptability were evaluated by two coders each rating

This article has been peer reviewed.

Corresponding Author: Joy Llewellyn-Beardsley, School of Health Sciences, University of Nottingham, Nottingham, United Kingdom. Email: [joy.llewellyn-beardsley@nottingham.ac.uk](mailto:joy.llewellyn-beardsley@nottingham.ac.uk). ORCID: 0000-0003-0525-6358

30 purposively-selected narratives. A refined version was produced and a formal evaluation conducted. Reliability was assessed by four coders each rating 95 purposively-selected narratives. Inter-coder reliability was assessed using Fleiss's kappa coefficients; test-retest reliability was assessed using intra-class correlation coefficients (ICCs).

**Results:** Multiple refinements to description, coding categories, and language were made. Data completeness was high, and no floor or ceiling effects were found. Inter-coder reliability ranged from moderate ( $k=0.58$ ) to perfect ( $k=1.00$ ) agreement. Test-retest reliability ranged from moderate ( $ICC=0.57$ ) to complete ( $ICC=1.00$ ) agreement. The final INCREASE comprises 77 items spanning five sections: Narrative Eligibility; Narrative Mode; Narrator Characteristics; Narrative Characteristics; Narrative Content.

**Conclusion:** INCREASE is the first evaluated tool to characterize mental health recovery narratives. It addresses current concerns around normative recovery narratives being used to promote compulsory wellness, e.g. by identifying narratives that reject diagnosis as an explanatory model and those with non-upward trajectories. INCREASE can be used to establish the diversity of a narrative collection and will be used in the NEON trials (ISRCTN11152837, ISRCTN63197153, ISRCTN76355273) to allow a recommender system to match narratives to participants.

## Introduction

The mental health recovery narratives of individuals (hereafter 'recovery narratives')<sup>1</sup> are a central component of clinical practice and other recovery-oriented interventions, for example narrative-based therapies,<sup>2-4</sup> Recovery College courses,<sup>5,6</sup> and peer support.<sup>7,8</sup>

Recovery narratives that have been recorded by others are increasingly used to augment this individual support. Recorded recovery narratives have been defined as first-person non-fiction accounts of recovery from mental health issues, including elements of adversity or struggle, and of self-defined strengths, successes, or survival.<sup>9</sup> They are widely available in the media, for example, memoirs and documentaries, and are used within mental health services in various forms, such as written and audio materials in bibliotherapy,<sup>10,11</sup> lived experience videos in digital interventions,<sup>12</sup> and blogs on national mental health organization websites. A publicly-available example fitting this definition is "Making Recovery Real in Dundee: Rona's Story," available at <https://youtu.be/7kGMazsGDJw>. Recovery narratives are also used in wider health contexts such as public health and anti-stigma campaigns.<sup>13,14</sup> Yet, until recently there has been little research on the impact of recovery narratives on those who access them (hereafter narrative 'recipients').

Evidence about impact on recipients is now emerging. A systematic review of five studies showed that recovery narratives can be both beneficial and harmful to recipients experiencing mental health issues.<sup>15</sup> A large-scale interview study found that positive impacts on recipients included feeling increased connection with others, validation of own experiences; empowerment and hopefulness about the future; greater life appreciation; changes in perspective; and a reduced sense of stigma. Negative impacts included a sense of inadequacy; disconnection from others; increased pessimism; and feeling emotionally burdened by others' distress.<sup>16</sup> Factors mediating whether a narrative had a beneficial or harmful impact on a recipient

included recognizing shared experiences or personal similarities with the narrator, which led to positive outcomes, and noticing narrator achievements, which led to hope and empowerment for some but inadequacy and disconnection for others.<sup>16,17</sup>

Clearly, not all narratives help all people at all times. How then can practitioners, public health campaigners, or individuals assess which narratives may be most beneficial, while avoiding potential harms? Although research has been undertaken in related areas, for example, whether content warnings are helpful in avoiding harms for some recipients,<sup>18,19</sup> no existing guidance describes what kinds of recovery narratives might have positive impacts for individuals, and how potential harms may be avoided.

Evidence is emerging, however, of the narrative characteristics that may affect a recipient's effective connection to a narrative, such as having shared experiences or the perceived authenticity of the narrator.<sup>20</sup> This evidence replicates findings in other fields. For example, reception theory, originating in communication and media studies, suggests that recipients are not passive but bring their own readings to narratives. A narrator may encode their narrative with various meanings, but the receiver may decode and interpret the narrative in different ways, relevant to their own circumstances such as individual cultural background and life experience.<sup>21</sup> Relatedly, narrative transportation theory suggests that recipients are more fully absorbed by narrative worlds if they can identify with aspects of the narrative or narrator.<sup>22,23</sup> And the more fully absorbed or transported they are, the more likely they are to be persuaded by a narrative's messages<sup>24</sup> – in this case, that recovery of a meaningful life may be possible for them.

Since characteristics of a narrative influence whether a recipient is likely to connect effectively with its message, an approach to identifying and describing the characteristics of a recovery narrative is needed. This will enable individuals to differentiate between narratives in advance of accessing them fully, and to choose narratives with characteristics that may be most helpful at a particular time.

No instrument currently exists to characterize recovery narratives. Available narrative-related instruments within mental health are designed to assess a narrator's mental processes or emotional states through their narratives, such as a sense of self in the Scale To Assess Narrative Development (STAND)<sup>25</sup> or narrator coherence in the Narrative Coherence Rating Scale.<sup>26</sup> Characterizing a narrative itself requires a different tool, drawing on methods used in other fields, such as the analysis of media content. An example is the PICMIN instrument (Picture of Mental Illness in Newspapers), which assesses mental illness stigma in print media.<sup>27</sup> PICMIN was developed using principles of content analysis, a systematic, replicable technique for compressing large amounts of text into fewer content categories based on explicit rules of coding, for use with qualitative and quantitative data.<sup>28</sup> Content analysis has been extended to health research, for example, in investigating stigmatizing attitudes towards mental health in social media<sup>29</sup> and exploring coping strategies for withdrawing from anti-psychotic medication.<sup>30</sup> It provides a set of procedures for the systematic coding of data, described by Robson<sup>31</sup> as a six-stage approach: identifying the research question; selecting the sample; defining the unit of analysis; constructing the categories for analysis; testing the categories on samples and assessing reliability; and conducting analysis. In this study, we adopted these procedures in modified form as design principles, omitting the final stage.

Content analysis was originally used to identify manifest (visible, surface-level) content, but has widened to include the identification of latent content, concerning underlying patterns or deep-level inferences.<sup>32,33</sup> Connection with a narrative is a complex process, involving recipient judgements about multifaceted phenomena such as authenticity.<sup>20</sup> Therefore, the characterization of latent as well as manifest content is needed, to ensure sufficient depth of characterization is achieved.

Any characterization of narratives risks adopting a reductive approach to its source material, and clearly a list of characteristics cannot replace the complexity of a narrative. What is needed is a 'condensation' - a process of shortening an artefact while still preserving its core<sup>32</sup> - to provide an instrument that may assist individuals and practitioners in finding their way to stories that might be most helpful for them.

Critical theorists and survivor/user researchers have also raised concerns about the commodification of recovery narratives within mental health practice,<sup>34</sup> for example when only certain "highly circumscribed" or "normatively successful" kinds of stories are promoted by services,<sup>35</sup> or when stories are used to promote "compulsory wellness."<sup>36</sup> Since evidence suggests that people connect with narratives they can identify with,<sup>22-24</sup> and since connection is the key mechanism of impact,<sup>16</sup> a tool that can characterize maximally diverse types of recovery narrative is required to create benefits for the greatest number of people. This study is therefore based on theory that uses the widest possible definitions of recovery narratives and which has been tested on diverse populations including marginalized groups and those currently under-served by mental health services.

The aim of this mixed methods study was to develop and evaluate an inclusive and in-depth inventory to characterize recorded recovery narratives, for use by mental health practitioners, individuals experiencing mental health distress, researchers, and curators of narrative-based collections or interventions. The objectives were to assess the inventory for its statistical reliability, and to assess its feasibility and acceptability to two groups: (1) a team of coders, as representatives of inventory target users; (2) a panel of people with lived experience of mental health difficulties, as representatives of the group whose stories would be characterized by the inventory.

## **Research Design and Methods**

### ***Ethical Considerations***

The research was undertaken as part of the Narrative Experiences Online (NEON) study. NEON aims to examine whether engagement with the recovery narratives of others can influence an individual's own recovery journey. Findings will inform three future trials (ISRCTN11152837, ISRCTN63197153, ISRCTN76355273).

Ethical Committee approval was obtained in advance for a protocol describing procedures for working with narratives and coders of narratives (London-West London REC and GTAC 18/LO/0991). The protocol for narratives described procedures for collecting consent for usage of narratives. The protocol for coders applied to all contributors not employed as NEON researchers. It described procedures for informed consent and to address the emotional and physical needs of research participants. All coders classified as research participants were provided with information sheets, and met with a member of the research team prior to study commencement to provide

written informed consent. Procedures for emotional and physical wellbeing took account of emotional labour, given the potentially distressing content of some recovery narratives, and the possibility of fatigue through coding large numbers of narratives.

### ***Stage 1: Development of Preliminary INCREASE***

A theory-based list of candidate inventory items and responses was generated through inspection of: a conceptual framework for recovery narratives;<sup>1,9</sup> change models of helpful and harmful recovery narrative impact<sup>15,16</sup> and recipients' experiences of hope and connection;<sup>20</sup> risk factors for developing mental health issues;<sup>37</sup> research on turning points in recovery narratives;<sup>38-44</sup> research on recovery factors;<sup>45</sup> and research on content warnings.<sup>18, 46-50</sup> Demographic item responses incorporated guidance from the UK Office of National Statistics for narrator ethnicity,<sup>51</sup> Stonewall UK for narrator sexuality,<sup>52</sup> NHS England for disability,<sup>53</sup> and the Diagnostic and Statistical Manual of Mental Disorders for items characterizing diagnosis.<sup>54</sup> This produced the preliminary inventory, INCREASE v1.

### ***Stage 2: Pilot Evaluation***

The aim of the pilot evaluation was to test the feasibility and acceptability of INCREASE v1 to coders. Feasibility was defined, modifying an existing definition, as 'suitable for use on a routine, sustainable and meaningful basis by target users when used in a specified manner for a specified purpose.'<sup>55</sup> The acceptability of a tool to its target users is important for successful implementation.<sup>56</sup> Use of appropriate language and a reasonable level of administrative burden have been identified as important factors in achieving this.<sup>55</sup> Acceptability was thus defined as 'capturing sufficient characteristics to provide an accurate summary of a recovery narrative, while representing reasonable administrative burden and using appropriate language.' Reasonable administrative burden was defined as each narrative taking <15 minutes to code. Appropriate language was defined as sensitive, inclusive, and acceptable both to coders and people with lived experience of mental health distress.

INCREASE v1 was tested by research team coders (JLB and FN) from sociology and psychology backgrounds, each of whom independently coded 30 narratives twice, two weeks apart. The narratives were a purposive sample,<sup>57</sup> selected for maximum variation in format (e.g., text, video), form (e.g., prose, poetry) and length. Characteristics of recovery narratives may differ according to form,<sup>1</sup> so text-based, audio and video narratives were chosen to test whether INCREASE could be applied equally to narratives across formats and forms. Narratives varied in length to test whether INCREASE sufficiently captured characteristics of both longer and shorter narratives, and whether administrative burden for longer narratives remained acceptable. Very long narratives such as documentaries and memoirs were excluded, due to limitations in coding team capacity to characterize these within the time constraints of the study. The 30 narratives for the pilot study comprised text (n=15), audio (n=5) and video-based (n=10) formats.

Coders completed written evaluations of feasibility and acceptability. Findings were analyzed in the first instance by the lead author (JLB) using semantic thematic analysis,<sup>58</sup> then discussed by the wider team, including researchers with a background in interaction design (SRE) and the development of psychometric measures (MS). This identified important elements missing from the inventory, and the impacts, both

physical and emotional, of coding large numbers of recovery narratives. Knowledge produced by evaluation work was used to create wellbeing training and instructions for external coders. No items from INCREASE v1 were deleted, but item descriptions were refined to produce INCREASE v2.

### ***Stage 3: Formal Evaluation***

The aim of the formal evaluation was to test (a) feasibility and acceptability to coders, (b) acceptability of INCREASE to people with lived experience of mental health difficulties and (c) the reliability of INCREASE. Twenty-two of the narratives used in Stage 2 were supplemented with 73 further narratives from the NEON Collection. The 95 narratives were a purposive sample chosen for maximum variation in format, form, length, gender, ethnicity, location, diagnosis, and collection source. This was to test whether INCREASE was applicable to the characteristics of narratives from as many different contexts as possible.

Four coders who had not previously used INCREASE and were not part of the NEON research team were recruited, comprising two from mental health research backgrounds and two from biomedical science and humanities research backgrounds, including coders with lived experience of mental health issues. The four coders were trained in the use of INCREASE v2, then each coded the same two narratives and discussed differences to assess concordance. The 95 recovery narratives were then independently coded by each coder using INCREASE v2 twice, two weeks apart (Round 1 and Round 2). Written evaluations were collected from the four coders on the feasibility and acceptability of the layout, items, responses, and descriptions in INCREASE v2. The acceptability of INCREASE v2 was reviewed by three NEON Lived Experience Advisory Panel (LEAP) members, comprising people with personal experience of mental health issues and recovery. INCREASE was finalized on the basis of statistical evaluation of poorly-performing items and further discussion of qualitative analysis findings within the wider research team.

## **Analysis**

### ***Qualitative Analysis***

Written evaluations of coders and LEAP members were analyzed using a semantic approach to thematic analysis, wherein data is searched to find repeated patterns of meaning in relation to the research questions.<sup>58</sup> In this study, the research questions related to the feasibility of INCREASE (e.g., aspects that were easy and difficult to use; areas for improvement) and the acceptability of INCREASE (e.g., appropriateness of language used; physical/emotional impacts on coders and people with lived experience).

### ***Statistical Analysis***

Data were analyzed using the following criteria: (1) data completeness, (2) response distributions for each item, including floor and ceiling effects (<10% of responses), and (3) exploring the scale score distributions at each time point (Round 1 and Round 2).

To assess item-level inter-coder reliability for agreement among all four coders, Fleiss's kappa coefficient<sup>59</sup> was computed on items with nominal values (un-weighted

kappa) or ordinal values (weighted kappa). Round 1 and Round 2 data were analyzed separately. Kappa value categories were defined in advance as 0.00–0.20 (None to slight); 0.21–0.40 (Fair); 0.41–0.60 (Moderate); 0.61–0.80 (Substantial); and 0.81–1.00 (Almost perfect agreement). Items with a kappa > 0.40 were identified as acceptable. To assess item-level test-retest reliability, the intra-class correlation coefficient (ICC) was calculated with a two-way random effects model. Quality of agreement was classified as: 0.00–0.19 (Poor), 0.20–0.39 (Weak), 0.40–0.59 (Moderate), 0.60–0.79 (Substantial), and 0.80–1.00 (Almost complete)<sup>60</sup>. Items with ICC > 0.40 were identified as acceptable. All statistical analyses were performed using SPSS (v26, IBM, Chicago, Illinois, USA).

## **Results**

### ***Stage 1: Development of Preliminary INCREASE***

INCREASE v1 comprised 68 items in five sections: Narrative Eligibility, Narrative Modality, Narrator Characteristics, Narrative Characteristics, and Narrative Content.

### ***Stage 2: Pilot Evaluation***

In relation to feasibility, coders evaluated instructions and descriptions of items as clear and thorough, with most items being straightforward to identify. Items in Section Four (Narrative Characteristics) were evaluated as more subjective and therefore more difficult to code. Some items could not be applied to all narratives, for example age of narrator in some audio-based narratives, and instructions were changed to reflect this. Other refinements to improve feasibility are shown in Table 1 (upper half).

**Table 1: Refinements to improve feasibility and acceptability following pilot evaluation**

INCREASE v1 Item	Issue identified	Refinement made in INCREASE v2
<b>FEASIBILITY</b>		
3. Is this a recovery narrative?	Coders applying differently - description contains two distinct constructs	Split into 2 items: 'Does the narrative contain elements of adversity or struggle?'; 'Does the narrative contain elements of success, strengths and/or survival?'
15-20. Disability categories	Rarely used - two items may be sufficient, e.g., 'narrator identifies/does not identify disabilities'	Candidates for deletion – reviewed in Stage 3
31-32. Relationship with Recovery and Stage of recovery	Found to be very similar, with 'stage' being more easily identifiable by coders within the narratives	Candidates for deletion – reviewed in Stage 3
40-46. Turning points	Coders applying differently	Instructions added. Items 41 & 46 renamed for greater clarity
59 & 60. Voluntary/involuntary use of mental health services	Insufficient items to characterize variety of mental health services used	Replaced with six items: 'Relationship with mental health professional'; 'Formal peer support'; 'Informal peer support'; 'Psychological services'; 'Involuntary uses of mental health services'; 'Hospitalization'
<b>ACCEPTABILITY</b>		
12. Ethnicity and 14. Sexuality	Available responses are insufficient to capture differences, and lead to 'othering' of narrators	Replaced with multiple response options
21-27. Diagnostic categories	Current items do not represent one coherent diagnostic model	DSM taxonomy chosen. 'Obsessive-compulsive related' item added. 'Rejects diagnosis' item added for narratives rejecting diagnosis.
Narrative content categories	Content frequently present but not captured by INCREASE	8 additional narrative content items added: Comparison with previous life experiences; spiritual/religious activities; experiences of stigma; creative activities; caring responsibilities; family experiences of mental health issues; diet/nutrition; and volunteering.



In relation to acceptability, administrative burden was evaluated as acceptable on the whole, although some negative impacts (fatigue, loss of concentration) were reported from coding large numbers of narratives without a break. Coder wellbeing was also both positively and negatively impacted by the process of coding recovery narratives. For example, coders reported both being inspired by narratives and finding it hard to read about the painful experiences of some narrators. To address this, a section on coder wellbeing was added to INCREASE instructions, wellbeing issues were highlighted and strategies suggested in INCREASE v2 training, and debriefs were offered to coders during the formal evaluation.

The language used in INCREASE v1 was found to be more appropriate for psychometric measures than an inventory of narratives and was changed to address this: for example, 'raters' was replaced with 'coders' and 'measure' became 'instrument.' Other refinements to improve acceptability are shown in Table 1 (lower half). All refinements were implemented to produce INCREASE v2.

### ***Stage 3: Formal Evaluation***

INCREASE v2 comprised 82 items in the same five sections. The sample narrative referred to in the introduction was coded as one of the 95 narratives. Worked examples of this and other narratives coded using INCREASE can be downloaded from [www.researchintorecovery.com/increase](http://www.researchintorecovery.com/increase).

In relation to feasibility, coders evaluated INCREASE v2 as easy to apply on the whole, due to the training provided and clear instructions. Items from Section 4 (Genre and Use of Metaphor), and some narrative formats (namely poems and images) were seen as more subjective and harder to code. Suggestions were made to amplify descriptions, for example including 'self-help/own learning' in the 'education' item of Section 5 (Narrative Content). Three sets of potentially overlapping items were reported in Section 5, between 'death/threatened death,' 'injury/threatened injury,' and 'self-inflicted injury/self-neglect;' between 'formal peer support' and 'informal peer support;' and between 'hobbies/interests' and 'creative activities.'

In relation to acceptability, coders identified fatigue as a negative effect of the administrative burden involved but reported having enough time to complete the task (95 narratives coded over 5 days, i.e., 19 per day). Therefore, INCREASE was deemed acceptable in terms of administrative burden. Coders identified that item labels describing content warnings were stark and could in themselves be distressing to encounter. Members of the LEAP independently identified the same issue. Item names were changed in the final version to reflect this, using alternatives generated by the LEAP.

In relation to statistical analysis, completeness of data was very high, with <1% missing data for each item. No ceiling or floor effects were identified. The inter-coder reliability and test-retest reliability for Sections 1 (Narrative Eligibility), 2 (Narrative Mode), 3 (Narrator Characteristics) and 4 (Narrative Characteristics) are shown in Table 2. The inter-coder reliability and test-retest reliability for Section 5 (Narrative Content) is shown in Table 3.

**Table 2:** Inter-coder and test-retest reliability for sections 1 to 4 of INCREASE v2

Item	Inter-coder reliability		Test-retest reliability (TRR)					
	Round 1	Round 2 2 weeks later	TRR by coder				TRR controlling for coder	TRR not controlling for coder
			Coder 1 <i>ICC</i>	Coder 2 <i>ICC</i>	Coder 3 <i>ICC</i>	Coder 4 <i>ICC</i>		
	<i>kappa</i>	<i>kappa</i>					<i>ICC</i>	<i>ICC (95% CI)</i>
1. Lived experience account?	0.79	1.00	0.46	0.48	1.00	0.48	0.66	0.79 (0.75-0.83)
2. Is this a narrative?	0.85	1.00	1.00	0.48	1.00	0.48	0.75	0.85 (0.82-0.88)
3. Narrator-defined adversity	0.62	0.76	<b>0.33</b>	1.00	1.00	<b>0.00</b>	0.47	0.64 (0.54-0.69)
4. Narrator-defined success	0.75	0.82	<b>0.35</b>	0.80	0.92	<b>0.32</b>	0.60	0.75 (0.62-0.79)
5. Written elements	0.99	1.00	0.97	1.00	1.00	0.98	0.97	0.99 (0.98-0.99)
6. Sound elements	0.99	1.00	1.00	1.00	0.95	1.00	0.98	0.99 (0.99-0.99)
7. Moving image elements	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00 (0.96-0.97)
8. Static image elements	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00 (0.98-0.99)
9. Total words (#)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
10. Total length (mins)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
11. Gender	0.90	0.90	0.99	0.91	0.93	0.76	0.83	0.90 (0.88-0.92)
12. Age	0.89	1.00	0.64	0.85	0.82	0.50	0.82	0.89 (0.99-1.00)
13. Ethnicity	0.95	1.00	0.85	0.95	0.77	0.90	0.91	0.95 (0.94-0.96)
14. Location of narrator	0.91	1.00	0.88	0.90	0.88	0.80	0.87	0.91 (0.89-0.93)
15. Sexuality	0.88	0.91	1.00	1.00	1.00	1.00	0.81	0.88 (0.84-0.91)
16. Visual difficulties	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00 (1.00-1.00)
17. Hearing difficulties	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00 (1.00-1.00)
18. Mobility difficulties	0.71	0.78	1.00	0.80	0.49	1.00	0.56	0.71 (0.64-0.76)
19. Memory difficulties	0.75	0.81	1.00	0.66	0.75	0.49	0.59	0.75 (0.69-0.79)
20. Self-care difficulties	0.76	0.81	0.66	1.00	0.49	1.00	0.63	0.76 (0.71-0.84)
21. Communication difficulties	0.86	0.86	1.00	1.00	0.80	0.39	0.76	0.86 (0.83-0.89)
22. Neuro-developmental related	0.57	0.70	0.66	1.00	1.00	1.00	0.50	0.57 (0.48-0.65)
23. Eating or food-related	0.94	1.00	1.00	1.00	1.00	1.00	0.86	0.94 (0.93-0.95)
24. Mood-related	0.87	1.00	0.72	0.91	0.83	0.48	0.77	0.87 (0.84-0.89)
25. Personality-related	0.89	0.89	1.00	0.85	0.66	0.71	0.51	0.89 (0.87-0.91)
26. Obsessive-compulsive related	0.75	0.80	0.66	1.00	0.49	1.00	0.62	0.75 (0.69-0.79)
27. Schizophrenia/psychosis	0.92	0.95	0.96	0.92	0.74	0.47	0.85	0.92 (0.90-0.93)
28. Stress-related	0.60	0.70	0.47	0.65	<b>0.01</b>	<b>0.00</b>	0.43	0.60 (0.51-0.61)
29. Substance-related	0.71	0.72	<b>0.37</b>	0.49	0.73	0.72	0.58	0.71 (0.64-0.76)
30. Rejects diagnosis	0.79	0.79	0.41	0.56	0.73	0.73	0.65	0.79 (0.74-0.82)
31. Genre	0.82	0.89	0.51	0.79	0.72	<b>0.34</b>	0.69	0.82 (0.78-0.85)
32. Positioning	0.85	0.90	0.54	0.74	0.65	0.70	0.72	0.85 (0.81-0.87)
33. Tone	0.78	0.78	0.60	0.81	0.81	0.41	0.64	0.78 (0.73-0.82)
34. Relationship with recovery	0.58	0.67	<b>0.38</b>	0.74	0.56	0.41	0.41	0.58 (0.49-0.65)
35. Stage of recovery	0.88	0.88	0.64	0.74	0.81	<b>0.26</b>	0.78	0.88 (0.85-0.90)
36. Trajectory	0.83	0.83	0.56	0.75	0.80	<b>0.35</b>	0.71	0.83 (0.79-0.86)
37. Use of metaphor	0.82	0.89	0.58	0.63	0.81	0.75	0.70	0.82 (0.78-0.85)

**Bold** = below predefined item-level threshold

**Table 3:** Inter-coder and test-retest reliability for section 5 of INCREASE v2

Item	Inter-coder reliability		Test-retest reliability (TRR)					
	Round 1	Round 2 2 weeks later	TRR by coder				TRR controlling for coder	TRR not controlling for coder
	<i>kappa</i>	<i>kappa</i>	Coder 1 <i>ICC</i>	Coder 2 <i>ICC</i>	Coder 3 <i>ICC</i>	Coder 4 <i>ICC</i>	<i>ICC</i>	<i>ICC (95% CI)</i>
38. Abuse or sexual violation	0.75	0.78	0.69	0.59	0.84	<b>0.30</b>	0.72	0.75 (0.69-0.79)
39. Death or threatened death	0.77	0.79	0.42	0.59	0.69	0.73	0.60	0.77 (0.72-0.81)
40. Self-inflicted injury	0.83	0.83	0.55	0.86	0.83	0.67	0.72	0.83 (0.79-0.86)
41. Injury or threatened injury	0.66	0.68	<b>0.39</b>	0.65	0.80	<b>0.36</b>	0.48	0.66 (0.58-0.72)
42. Injustice, prejudice, discrimination	0.85	0.89	0.83	0.81	1.00	0.42	0.73	0.85 (0.81-0.87)
43. Taking charge	0.62	0.72	<b>0.30</b>	0.50	0.57	<b>0.27</b>	0.41	0.62 (0.54-0.69)
44. Interventions/support from others	0.70	0.70	<b>0.25</b>	0.65	0.55	0.50	0.52	0.69 (0.62-0.74)
45. Self-acceptance	0.68	0.74	<b>0.10</b>	0.71	<b>0.24</b>	0.57	0.52	0.68 (0.61-0.74)
46. Spiritual	0.81	0.81	0.65	0.85	0.58	0.60	0.69	0.81 (0.76-0.84)
47. Rude awakening	0.97	0.94	1.00	1.00	0.74	1.00	0.95	0.97 (0.91-0.97)
48. Deciding to live	0.93	0.90	<b>0.25</b>	0.66	<b>0.14</b>	<b>0.14</b>	0.86	0.93 (0.91-0.93)
49. Shift in identify	0.63	0.76	<b>0.29</b>	0.88	<b>0.20</b>	0.54	0.47	0.63 (0.55-0.69)
50. Pregnancy/birth	0.99	1.00	0.77	0.77	0.66	0.66	0.98	0.98 (0.95-1.00)
51. Family	0.81	0.85	0.57	0.77	0.75	0.57	0.69	0.81 (0.77-0.84)
52. Care System	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00 (0.99-1.00)
53. Education	0.87	0.97	0.73	0.85	0.78	0.52	0.77	0.87 (0.84-0.89)
54. Friendships	0.79	0.84	0.53	0.78	0.74	0.49	0.64	0.78 (0.73-0.82)
55. Relationships	0.77	0.73	0.52	0.75	0.81	<b>0.33</b>	0.63	0.77 (0.72-0.81)
56. Housing	0.83	0.89	0.59	0.69	0.63	0.42	0.71	0.83 (0.79-0.86)
57. Income	0.84	0.85	0.47	0.92	0.79	0.49	0.72	0.84 (0.80-0.86)
58. Work	0.78	0.78	0.67	0.71	0.67	<b>0.21</b>	0.65	0.78 (0.73-0.82)
59. Criminal justice system	0.70	0.77	0.77	0.65	0.50	1.00	0.56	0.70 (0.65-0.76)
60. Diagnosis	0.80	0.86	0.60	0.71	0.57	0.51	0.66	0.80 (0.75-0.83)
61. Medication	0.86	0.85	0.80	0.82	0.70	0.60	0.75	0.86 (0.82-0.88)
62. Relationship with MH professional	0.83	0.83	0.63	0.88	0.69	0.59	0.71	0.83 (0.79-0.86)
63. Formal peer support	0.73	0.75	<b>0.39</b>	0.64	0.63	<b>0.33</b>	0.57	0.73 (0.67-0.74)
64. Informal peer support	0.78	0.82	0.44	0.62	0.69	0.61	0.65	0.78 (0.73-0.82)
65. Involuntary use of MH services	0.84	0.84	0.71	0.79	0.85	<b>0.36</b>	0.73	0.84 (0.81-0.87)
66. Hospitalization	0.78	0.73	0.78	0.83	0.67	<b>0.38</b>	0.65	0.78 (0.73-0.82)
67. Psychological services	0.78	0.75	0.63	0.78	0.53	0.43	0.63	0.78 (0.73-0.81)
68. Alternative therapies	0.82	0.84	0.61	0.73	0.84	0.65	0.70	0.82 (0.78-0.85)
69. Being in natural environments	0.71	0.78	0.50	0.72	0.50	0.48	0.56	0.71 (0.65-0.76)
70. Animals/pets	0.77	0.76	<b>0.38</b>	0.68	0.87	<b>0.00</b>	0.67	0.77 (0.71-0.81)
71. Community activities	0.63	0.63	<b>0.39</b>	0.56	0.56	0.48	0.46	0.63 (0.55-0.70)
72. Hobbies/interests	0.74	0.73	<b>0.31</b>	0.69	0.47	0.68	0.59	0.74 (0.68-0.78)
73. Physical activities	0.71	0.78	0.44	0.91	0.48	<b>0.38</b>	0.55	0.71 (0.65-0.76)
74. Political activities	0.61	0.72	0.43	0.43	0.66	0.41	0.44	0.61 (0.53-0.68)
75. Previous life circumstances	<b>0.26</b>	<b>0.31</b>	<b>0.09</b>	0.68	<b>0.23</b>	<b>0.18</b>	<b>0.14</b>	<b>0.26</b> (0.10-0.39)
76. Spiritual/religious activities	0.89	0.94	0.81	0.73	0.70	0.62	0.80	0.89 (0.86-0.91)
77. Stigma	0.79	0.84	0.74	0.79	0.60	<b>0.35</b>	0.66	0.79 (0.75-0.83)

78. Creative activities	0.82	0.82	0.64	0.80	0.63	0.69	0.70	0.82 (0.79-0.85)
79. Caring responsibilities	0.58	0.71	<b>0.38</b>	0.65	<b>0.39</b>	<b>0.35</b>	0.41	0.58 (0.49-0.66)
80. Family experiences of MH issues	0.87	0.88	0.65	0.65	0.42	0.49	0.77	0.87 (0.84-0.89)
81. Diet/Nutrition	0.81	0.88	0.59	0.62	0.56	<b>0.00</b>	0.68	0.81 (0.77-0.84)
82. Volunteering	0.89	0.94	0.74	0.91	0.85	0.68	0.82	0.89 (0.87-0.91)

**Bold** = below predefined item-level threshold

Inter-coder reliability for round 1 coding ranged from 0.58 (Moderate) to 1.00 (Almost perfect), and for round 2 coding ranged from 0.63 (Substantial) to 1.00 (Almost perfect), with the exception of one outlier item #75 (Previous life circumstances) which had only Fair agreement. Coder 4 had the most variance in coding from round 1 and round 2. When controlling for coder 4, eight items showed anomalous behaviour (items #22, #28, #34, #41, #44, #45, #49 and #71). Overall, 126 (79%) of the 160 item-level ICCs were  $\geq 0.75$ , indicating inter-coder reliability is strong.

Test-retest reliability (TRR) controlling for coder ranged from 0.41 (Moderate) to 1.00 (Almost complete), with the exception of the same outlier item #75 which had only Poor agreement. TRR not controlling for coder ranged from 0.57 (Moderate) to 1.00 (Almost complete), again with the exception of item #75 which had only Weak agreement. Three items performed less well in terms of TRR: #22, #34 and #38. Overall, most items had a test-retest reliability of greater than 0.75, indicating test-retest reliability is strong.

Coders had different levels of test-retest agreement, with the highest for coder 2 (mean 0.78, SD=0.25), followed by coder 3 (mean 0.71, SD=0.22), coder 1 (mean 0.64, SD=0.25) and coder 4 (mean 0.57, SD=0.27). Coder 4 was the coder with least mental health-specific experience or training.

All item names, responses, and descriptions were reviewed and refined, with a particular focus on items identified as problematic in the statistical analysis. For example, item #75 (Previous life circumstances) performed poorly in terms of both inter-coder and test-retest reliability so was deleted from the final version. For item #34 (Relationship with recovery), coding by the two coders with non-mental health research backgrounds was less reliable, so a definition of the specific meaning of recovery in mental health was added to the final version to address this.

All refinements made at items and response level to produce the final 77-item version of INCREASE are shown in Table 4. The final version of INCREASE can be downloaded from <http://www.researchintorecovery.com/increase>.

**Table 4:** Refinements to improve feasibility, acceptability and reliability following pilot evaluation

INCREASE v2 Item	Issue identified	Refinement made in final version of INCREASE
<b>FEASIBILITY</b>		
18. Memory difficulties	Items are too similar	Merged into one item: 'cognitive difficulties'
20. Communication difficulties		
30. Rejects diagnosis	Does not capture narratives which also use non-diagnostic terms, but do not reject diagnosis	Renamed to 'uses a non-diagnostic framework'
34. Relationship with recovery	Current item responses do not capture narratives which contain active resistance to the recovery model	Added 'rejects recovery' responses to both items
35. Stage of recovery		
39. Death/threatened death	Descriptions contain potentially overlapping examples	Item names and descriptions revised into more discrete categories
40. Self-inflicted injury/self-neglect		
41. Injury/threatened injury		
63. Formal peer support	Not possible to identify whether peer support is formal or informal in some narratives	Merged into one item: 'peer support'
64. Informal peer support		
72. Hobbies/interests	Items are too similar	Merged into one item: 'hobbies/interests/creative activities'
78. Creative activities		
<b>ACCEPTABILITY</b>		
35. Stage of recovery	May be perceived by users as too similar to 34 - Relationship with Recovery. This item based on the Stages of Recovery (STORI) instrument, thus refers more to narrator than narrative characteristics.	Moved to Section 3 'Narrator Characteristics'
38-42. Content warnings	Item labels too stark - potentially distressing in themselves	Renamed
41. Self-harm	Does not clearly include eating disorder behaviours, specifically identified in previous research as potentially harmful content	Renamed to 'Self-harm including eating disorders'
74. Political activities	Item label inaccurate, as description relates to certain political activities and not others e.g., standing for election	Item renamed to 'Activism' to better reflect recovery narrative content
<b>RELIABILITY</b>		
48. Deciding to live	Poor performance for TRR	Item removed; description incorporated into other turning point items.
75. Previous life circumstances	Poor performance for ICR and TRR	Item removed

## **Discussion**

INCREASE has satisfactory acceptability, feasibility, and reliability across time and between coders. It comprises 77 items across five sections: Narrative Eligibility, Narrative Mode, Narrator Characteristics, Narrative Characteristics, and Narrative Content. 71 items characterize manifest content and six characterize latent content (Stage of recovery; Genre; Positioning; Tone; Relationship with recovery; and Trajectory). The inclusion of these six items links this study to other content analysis-based studies within mental health research which incorporate latent content in addition to the more traditional focus on manifest content alone.<sup>33,61,62</sup>

Using INCREASE can impact on the wellbeing of coders, with both positive and negative effects being reported. This links to previous work on researcher wellbeing when exploring distressing topics,<sup>63-65</sup> on vicarious traumatization among mental health professionals and researchers,<sup>66-68</sup> and on the wellbeing of online content moderators.<sup>69,70</sup> Organizations utilizing INCREASE should select approaches that support coder wellbeing. Future research may focus on identifying the most effective approaches to supporting coder wellbeing when using INCREASE or similar tools, for example, by experimentally comparing different approaches to reducing harms.

### ***Strengths and limitations***

Strengths of this study include an inclusive definition of recovery narratives, a strong theoretical basis for each included item, the use of coders from diverse backgrounds, thorough ethical consideration of both coder wellbeing and the issues raised by characterizing narratives, the coding of a large number of narratives, the inclusion of a pilot stage as well as a formal evaluation, and the testing of both inter-coder and test-retest reliability.

Limitations include that INCREASE is not suitable for use with group or collective narratives. INCREASE was not tested on longer recovery narratives such as documentaries or memoirs, so validating the inventory against such material could be a future research study. No mechanism was included in the study to assess the effects of coder fatigue; future research may explore this. Lastly, some items within the inventory are more difficult to apply to image-based recovery narratives and to more symbolic forms of prose, such as poetry. Other tools may need to be developed which can characterize such narratives more meaningfully. This is particularly important as evidence suggests that more symbolic forms may better support meaning-making for particular forms of mental health distress, as a recent conceptual review on understanding psychosis through poetry suggests.<sup>71</sup>

### ***Implications***

Five uses for the Inventory of Characteristics of Recovery Stories (INCREASE) are envisaged.

First, INCREASE can be used to create catalogues of narratives for both mental health practitioners and individuals experiencing mental health distress to draw on. This would facilitate the choosing of narratives judged likely to be most helpful at a given time, with the aim of maximising benefit and minimizing harm for the recipient.

Second, INCREASE can characterize collections of narratives for use in narrative-based online interventions. Within the NEON study INCREASE will be used to characterize a large database of narratives called the NEON Collection, to enable participants to browse stories by categories. In addition, a machine learning algorithm will use the data produced by INCREASE to choose individual narratives to present to participants opting to be 'matched' to stories.<sup>72</sup> Future research is planned to assess the effectiveness of the instrument in assisting individuals to find helpful stories.<sup>72</sup>

Third, as its name suggests, INCREASE can be used as a tool for organizations using recovery narratives in campaigns or on websites, and for curators of recovery narrative collections to consider the level of diversity within the narrative collection. The inventory will identify gaps, enabling the scope of the collection to be established and under-represented narrators or narrative types to be targeted for inclusion as appropriate. This may increase the likelihood of a positive connection with narratives for a wider variety of recipients.

Fourth, INCREASE can be used as a tool for researchers to assess the extent to which narratives available from organizations span a diverse range of narrative characteristics. For example, stories on health service websites (e.g., [www.likemind.nhs.uk/your-stories](http://www.likemind.nhs.uk/your-stories)) could be evaluated in relation to whether they include narratives from across the range of recovery narrative characteristics, including people who use a non-diagnostic explanatory framework (item #30), or whose recovery trajectory is not upwards (item #36). This would allow the extent to which concerns that "personal stories from consumer/survivors have been harnessed by mental health organizations to further their interests" are valid<sup>73(p.86)</sup> to be empirically investigated.

Fifth, INCREASE can be used to provide evidence about particular characteristics that are more likely to be beneficial or to trigger distress in recipients. One way in which avoiding potential harm to recipients has been attempted in this field is the contested issue of providing content warnings, also known as trigger warnings, within collections of recovery narratives, particularly those online. Graphic description of subjects such as violence, death, or abuse are thought potentially to trigger distress similar to the original trauma in those with existing mental health conditions such as post-traumatic stress disorder. Content warnings have been widely used online and within education and professional training contexts<sup>74,75</sup> to help people avoid the perceived harms which may be caused by certain narrative content. Although some empirical research is emerging,<sup>18,19</sup> there is minimal empirical evidence on whether trigger warnings are helpful or not in managing distress for people who have experienced trauma. INCREASE could be used in future research to identify particularly beneficial or harmful characteristics for particular populations and to generate evidence-based guidance on the use of narratives and their potential benefits and harms.

This study also carries implications for designers of content warning systems. Reliability of content warning items in INCREASE ranged from moderate to almost perfect/almost complete agreement but were not perfectly reliable. Other content warnings systems may thus also not be perfectly reliable; hence, designers of content warning systems should assess and manage any lack of reliability, either through use of multiple coders, or by providing recipients with the information that the categorization may not be complete.



## Conclusion

INCREASE is a standardized and reliable inventory to characterize mental health recovery narratives, which is feasible for use by coders without a narrative research background, and acceptable to both coders and people with lived experience of mental health issues. It has the potential to be used not only to help individuals and practitioners to find the stories which may be most helpful to them, but also to enable curators and researchers to identify bias or lack of diversity within collections and narrative-based campaigns or interventions.

## Acknowledgements

This article is independent research funded by the NIHR under its Programme Grants for Applied Research Programme (Personal experience as a recovery resource in psychosis: Narrative Experiences Online (NEON) Programme, RP-PG-0615-20016). Mike Slade acknowledges the support of Center for Mental Health and Substance Abuse, University of South-Eastern Norway and the NIHR Nottingham Biomedical Research Centre. James Roe was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East Midlands (CLAHRC EM), now recommissioned as NIHR Applied Research Collaboration East Midlands (ARC EM). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The authors declare there are no conflicts of interest regarding this article.

## Tables Legend

Table 1: Refinements to improve feasibility and acceptability following pilot evaluation

Table 2: Inter-coder and test-retest reliability for Sections 1 to 4 (shown as groups) of INCREASE v2

Table 3: Inter-coder and test-retest reliability for Section 5 of INCREASE v2

Table 4: Refinements to improve feasibility, acceptability and reliability following pilot evaluation

## References

1. Llewellyn-Beardsley, J., et al., *Not the story you want? Assessing the fit of a conceptual framework characterising mental health recovery narratives*. *Social Psychiatry and Psychiatric Epidemiology*, 2019. **55**: p. 295-308. DOI: 10.1007/s00127-019-01791-x
2. Hansson, L., A. Lexén, and J. Holmén, *The effectiveness of narrative enhancement and cognitive therapy: a randomized controlled study of a self-stigma intervention*. *Social Psychiatry and Psychiatric Epidemiology*, 2017. **52**(11): p. 1415-1423.
3. White, M., et al., *Narrative means to therapeutic ends*. 1990, New York: WW Norton & Company.
4. Yanos, P.T., et al., *A randomized-controlled trial of treatment for self-stigma among persons diagnosed with schizophrenia-spectrum disorders*. *Social Psychiatry and Psychiatric Epidemiology*, 2019. **54**: p. 1363-1378. DOI: 10.1007/s00127-019-01702-0
5. Hall T., et al., *A process and intermediate outcomes evaluation of an Australian recovery college*. *Journal of Recovery in Mental Health*, 2018. **1**(3): p. 7-20.
6. Nurser, K.P., et al., *Personal storytelling in mental health recovery*. *Mental Health Review Journal*, 2018. **23**(1): p. 25-36.
7. Ibrahim, N., et al., *A systematic review of influences on implementation of peer support work for adults with mental health issues*. *Social Psychiatry and Psychiatric Epidemiology*, 2019. DOI: 10.1007/s00127-019-01739-1

8. Mancini, M.A., *Strategic storytelling: An exploration of the professional practices of mental health peer providers*. Qualitative Health Research, 2019. **29**(9): p. 1266-1276.
9. Llewellyn-Beardsley, J., et al., *Characteristics of mental health recovery narratives: systematic review and narrative synthesis*. PloS One, 2019. **14**(3) DOI: 10.1371/journal.pone.0214678.
10. Fanner, D., and C. Urquhart, *Bibliotherapy for mental health service users part 1: A systematic review*. Health Information & Libraries Journal, 2008. **25**(4): p. 237-252.
11. Peterkin, A., and S. Grewal, *Bibliotherapy: The therapeutic use of fiction and poetry in mental health*. International Journal of Person Centered Medicine, 2018. **7**(3): p. 175-181.
12. Williams, A., et al., *Recovery after psychosis: Qualitative study of service user experiences of lived experience videos on a recovery-oriented website*. JMIR Mental Health, 2018. **5**(2) DOI: 10.2196/mental.9934
13. Kosyluk, K., et al., *Examining the impact of This Is My Brave on mental illness stigma and willingness to seek help: A pilot study*. Community Mental Health Journal, 2018. **54**(3): p. 276-281.
14. Pinfold, V., et al., *Active ingredients in anti-stigma programmes in mental health*. International Review of Psychiatry, 2005. **17**(2): p. 123-131.
15. Rennick-Egglestone, S., et al., *Mental Health Recovery Narratives and Their Impact on Recipients: Systematic Review and Narrative Synthesis*. Canadian Journal of Psychiatry, 2019. **10**: p. 669-679. DOI: 10.1177/0706743719846108
16. Rennick-Egglestone, S., et al., *The impact of mental health recovery narratives on recipients experiencing mental health issues: qualitative analysis and change model*. PloS One, 2019. **14**(12) DOI: 10.1371/journal.pone.0226201.
17. Truong, C., et al., *The role of self-disclosure by peer mentors: Using personal narratives in depression care*. Patient Education and Counseling, 2019. **102**(7): p. 1273-1279.
18. Bellet, B.W., P.J. Jones, and R.J. McNally, *Trigger warning: Empirical evidence ahead*. Journal of Behavior Therapy and Experimental Psychiatry, 2018. **61**: p. 134-141.
19. Sanson, M., D. Strange, and M. Garry, *Trigger warnings are trivially helpful at reducing negative affect, intrusive thoughts, and avoidance*. Clinical Psychological Science, 2019. **7**(4): p. 778-793.
20. Ng, F., et al., *The mechanisms and processes of connection: Developing a causal chain model capturing impacts of receiving recorded mental health recovery narratives*. BMC Psychiatry, 2019. **19**(413): p. 1-15. DOI: 10.1186/s12888-019-2405-z.
21. Hall, S., *Encoding/decoding, in media and cultural studies: Keywords*, M.G. Durham and D.M. Kellner, Editors. 2006, Blackwell: Oxford. p. 163-173.
22. Caputo, N.M., and D. Rouner, *Narrative processing of entertainment media and mental illness stigma*. Health Communication, 2011. **26**(7): p. 595-604.
23. Slater, M.D., and D. Rouner, *Entertainment—education and elaboration likelihood: Understanding the processing of narrative persuasion*. Communication Theory, 2002. **12**(2): p. 173-191.
24. Green, M.C., and T.C. Brock, *The role of transportation in the persuasiveness of public narratives*. Journal of Personality and Social Psychology, 2000. **79**(5): p. 701.
25. Lysaker, P.H., et al., *The scale to assess narrative development: Association with other measures of self and readiness for recovery in schizophrenia spectrum disorders*. The Journal of Nervous and Mental Disease, 2006. **194**(3): p. 223-225.
26. Lysaker, P.H., et al., *Insight and personal narratives of illness in schizophrenia*. Psychiatry: Interpersonal and Biological Processes, 2002. **65**(3): p. 197-206.
27. Rukavina, T.V., et al., *Development of the PICMIN (picture of mental illness in newspapers): Instrument to assess mental illness stigma in print media*. Social Psychiatry and Psychiatric Epidemiology, 2012. **47**(7): p. 1131-1144.
28. Krippendorff, K., *Content analysis: An introduction to its methodology*. 2018, London: SAGE Publications.
29. Robinson, P., et al., *Measuring attitudes towards mental health using social media: Investigating stigma and trivialisation*. Social Psychiatry and Psychiatric Epidemiology, 2019. **54**(1): p. 51-58.
30. Larsen-Barr, M., et al., *Attempting to stop antipsychotic medication: Success, supports, and efforts to cope*. Social Psychiatry and Psychiatric Epidemiology, 2018. **53**(7): p. 745-756.
31. Robson, C., *Real world research*. Chichester, John Wiley & Sons, 2011. **3**
32. Graneheim, U.H., and B. Lundman, *Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness*. Nurse Education Today, 2004. **24**(2): p. 105-112.
33. Hicks, J., et al., *Middle school bullying: Student reported perceptions and prevalence*. Journal of Child and Adolescent Counseling, 2018. **4**(3): p. 195-208.

34. O'Donnell, A., L. Sapouna, and L. Brosnan, *Storytelling: An act of resistance or a commodity?* Journal of Ethics in Mental Health, 2019. **10**(6): p. 1-13.
35. Woods, A., A. Hart, and H. Spandler, *The recovery narrative: Politics and possibilities of a genre.* Culture, Medicine, and Psychiatry, 2019. p. 1-27.
36. Carton, T., *A Welcome antidote to the evangelism of compulsory optimism and resilience.* Advances in Applied Sociology, 2017. **7**(2): p. 83.
37. World Health Organisation. *Risks to mental health: An overview of vulnerabilities and risk factors.* 2012 [cited 2019 Nov 24]. Available from: [https://www.who.int/mental\\_health/mhgap/risks\\_to\\_mental\\_health\\_EN\\_27\\_08\\_12.pdf](https://www.who.int/mental_health/mhgap/risks_to_mental_health_EN_27_08_12.pdf).
38. Allott, P., and L. Loganathan, *Discovering hope for recovery from a British perspective: A review of literature.* 2002, Centre for Community Mental Health, University of Central England: Birmingham.
39. Anich, C.L., *The nature and meaning of lived experiences on 'the becoming of self' for people with schizophrenia,* 1998. Utah: University of Utah.
40. Brown, W., and N. Kandirikirira, *Report on narrative investigation of mental health recovery.* 2007, Scottish Recovery Network: Glasgow.
41. Deegan, P.E., *Recovery as a self-directed process of healing and transformation.* Occupational Therapy in Mental Health, 2002. **17**(3-4): p. 5-21.
42. Roberts, G., and P. Wolfson, *The rediscovery of recovery: Open to all.* Advances in Psychiatric Treatment, 2004. **10**(1): p. 37-48.
43. Smith, M.K., *Recovery from a severe psychiatric disability: Findings of a qualitative study.* Psychiatric Rehabilitation Journal, 2000. **24**(2): p. 149.
44. Snelling, E., *Hungry researchers: The tensions and dilemmas of developing an emancipatory research project with members of a hearing voices group.* Journal of Social Work Practice, 2005. **19**(2): p. 131-147.
45. Leamy, M., et al., *Conceptual framework for personal recovery in mental health: Systematic review and narrative synthesis.* The British Journal of Psychiatry, 2011. **199**(6): p. 445-452.
46. Beverly, E.A., et al., *Students' perceptions of trigger warnings in medical education.* Teaching and Learning in Medicine, 2018. **30**(1): p. 5-14.
47. Boysen, G.A., *Evidence-based answers to questions about trigger warnings for clinically-based distress: A review for teachers.* Scholarship of Teaching and Learning in Psychology, 2017. **3**(2): p. 163.
48. Gainsburg, I., and A. Earl, *Trigger warnings as an interpersonal emotion-regulation tool: Avoidance, attention, and affect depend on beliefs.* Journal of Experimental Social Psychology, 2018. **79**: p. 252-263.
49. Kostouros, P., and J. Wenzel, *Depictions of suffering in the postsecondary classroom.* Traumatology, 2017. **23**(3): p. 250.
50. Shannon, P.J., et al., *Experiences of stress in a trauma treatment course.* Journal of Social Work Education, 2014. **50**(4): p. 678-693.
51. Gov.UK. *List of Ethnic Groups.* 2019 [cited 2019 Jan 14]; Available from: <https://www.ethnicity-facts-figures.service.gov.uk/ethnic-groups>
52. Ashworth, A., and A.V. Vliet, *Using monitoring data: Making the most of sexual orientation data collection.* Stonewall Workplace Guides 2012 [cited 2019 Jan 20]; Available from: [https://www.stonewall.org.uk/sites/default/files/using\\_monitoring\\_data.pdf](https://www.stonewall.org.uk/sites/default/files/using_monitoring_data.pdf).
53. England, N., *Accessible information standard review: Diversity monitoring questions.* 2017 [cited 2019 Jan 20]; Available from: <https://www.england.nhs.uk/wp-content/uploads/2017/01/ais-review-diversity-monitoring-questions.docx>.
54. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders.* 5th ed. 2013, Arlington, VA: American Psychiatric Publishing, Inc.
55. Slade, M., G. Thornicroft, and G. Glover, *The feasibility of routine outcome measures in mental health.* Social Psychiatry and Psychiatric Epidemiology, 1999. **34**(5): p. 243-249.
56. Proctor, E., et al., *Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda.* Administration and Policy in Mental Health and Mental Health Services Research, 2011. **38**(2): p. 65-76.
57. Teddlie, C., and F. Yu, *Mixed methods sampling: A typology with examples.* Journal of Mixed Methods Research, 2007. **1**(1): p. 77-100.
58. Braun, V., and V. Clarke, *Using thematic analysis in psychology.* Qualitative Research in Psychology, 2006. **3**(2): p. 77-101.

59. Fleiss, J.L., and J. Cohen, *The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability*. Educational and Psychological Measurement, 1973. **33**(3): p. 613-619.
60. Lee, K.M., et al., *Pitfalls and important issues in testing reliability using intraclass correlation coefficients in orthopaedic research*. Clinics in Orthopedic Surgery, 2012. **4**(2): p. 149-155.
61. Carlozzi, N.E., et al., *A latent content analysis of barriers and supports to healthcare: perspectives from caregivers of service members and veterans with military-related traumatic brain injury*. The Journal of Head Trauma Rehabilitation, 2018. **33**(5): p. 342-353.
62. Thyme, K.E., et al., *Qualitative content analysis in art psychotherapy research: Concepts, procedures, and measures to reveal the latent meaning in pictures and the words attached to the pictures*. The Arts in Psychotherapy, 2013. **40**(1): p. 101-107.
63. Hanna, E., *The emotional labour of researching sensitive topics online: Considerations and implications*. Qualitative Research, 2018. **19**(5): p. 524-539.
64. Moncur, W., *The emotional wellbeing of researchers: Considerations for practice*. 2013. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
65. Velardo, S., and S. Elliott, *Prioritising doctoral students' wellbeing in qualitative research*. The Qualitative Report, 2018. **23**(2): p. 311-318.
66. Kadambi, M.A., and D. Truscott, *Vicarious trauma among therapists working with sexual violence, cancer and general practice*. Canadian Journal of Counselling and Psychotherapy, 2004. **38**(4): p. 260-276.
67. McCann, I.L., and L.A. Pearlman, *Vicarious traumatization: A framework for understanding the psychological effects of working with victims*. Journal of Traumatic Stress, 1990. **3**(1): p. 131-149.
68. Van der Merwe, A., and X. Hunt, *Secondary trauma among trauma researchers: Lessons from the field*. Psychological Trauma: Theory, Research, Practice, and Policy, 2019. **11**(1): p. 10-18.
69. Gerrard, Y., *Behind the screen: Content moderation in the shadows of social media*. 2019, London: SAGE Publications.
70. Gillespie, T., *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. 2018, London: Yale University Press.
71. Pearson, M., S. Rennick-Egglestone, and G. Winship, *How can poetry support the understanding of psychotic experiences? A conceptual review*. Journal of Recovery in Mental Health, 2020. **3**(1): p. 39-53.
72. Rennick-Egglestone S., et al., *Impact of receiving recorded mental health recovery narratives on quality of life in people experiencing psychosis, people experiencing other mental health problems and for informal carers: Narrative Experiences Online (NEON) study protocol for three randomised controlled trials*, Trials, 2020. **21**, 661.
73. Costa, L., et al., *"Recovering our stories": A small act of resistance*. Studies in Social Justice, 2012. **6**(1): p. 85-101.
74. Bass, S.A., and M.L. Clark, *The gravest threat to colleges comes from within*. Chronicle of Higher Education, 2015. **62**(7): p. A26-A27.
75. Gust, O., *I use trigger warnings—but I'm not mollycoddling my students*. 2016 [cited 2019 Sep 19]; Available from: <https://www.theguardian.com/higher-education-network/2016/jun/14/i-use-trigger-warnings-but-im-not-mollycoddling-my-students>.