



Laing, Christianne (2026) *Emotional schema: a systematic review of measures and development of a new measure*. D Clin Psy thesis.

<https://theses.gla.ac.uk/85909/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>  
[research-enlighten@glasgow.ac.uk](mailto:research-enlighten@glasgow.ac.uk)



**Emotional Schema:  
A Systematic Review of Measures and  
Development of a New Measure**

Christianne Marie Laing, PhD MSc MRes BSc (Hons)

Submitted in partial fulfilment of the requirements for the degree of  
Doctorate in Clinical Psychology

School of Health and Wellbeing  
College of Medical, Veterinary and Life Sciences  
University of Glasgow

April 2026

## Table of Contents

<b>List of Tables</b> .....	<b>3</b>
<b>List of Figures</b> .....	<b>4</b>
<b>Acknowledgements</b> .....	<b>5</b>
<b>Chapter 1: Reliability and Validity of Current Measures of Emotional Schema: A Systematic Review</b> .....	<b>6</b>
Abstract.....	7
Introduction.....	9
Methods.....	14
Results.....	18
Discussion.....	30
References.....	35
<b>Chapter 2: Development of a New Measure of Emotional Schema</b> .....	<b>40</b>
Plain Language Summary.....	41
Abstract.....	42
Introduction.....	43
Methods.....	48
Results.....	53
Discussion.....	63
References.....	69
<b>Appendices</b> .....	<b>73</b>
Appendix 1.1 PRISMA Checklist.....	73
Appendix 1.2 Search Strategy.....	77
Appendix 1.3 Data Extraction Checklist.....	81
Appendix 1.4 COSMIN Risk of Bias Quality Appraisal Tool.....	82
Appendix 1.5 Further Information on Included Scales and Their Utility.....	83
Appendix 2.1 Revised Research Proposal.....	86
Appendix 2.2 Ethical Approval Letter.....	87
Appendix 2.3 Phase 2 Materials.....	88
Appendix 2.4 Phase 3 Questionnaire Items.....	89
Appendix 2.5 Phase 3 Materials.....	92
Appendix 2.6 Revised Analysis Plan.....	93
Appendix 2.7 Mood Questionnaires Phase 3.....	94

Appendix 2.8 Syntax File.....96

Appendix 2.9 Data Availability Statement.....97

## List of Tables

<b>Table 1.1</b>	COSMIN definitions of measurement properties
<b>Table 1.2</b>	Overview of Included Studies
<b>Table 1.3</b>	Measure Characteristics
<b>Table 1.4</b>	Evaluations of included studies according to the COSMIN guidelines
<b>Table 2.1</b>	Demographic Characteristics of Participants in Phase 2
<b>Table 2.2</b>	Demographic Characteristics of Participants in Phase 3
<b>Table 2.3</b>	Comparison of descriptive statistics for PHQ-9, GAD-7 and BAEQ, scores by self-reported mental health experience
<b>Table 2.4</b>	Summary of ESSeNcE Factors, Sample Items and Reliability
<b>Table 2.5</b>	Promax-Rotated Loadings for ESSeNcE questionnaire items in Sample 3 (n = 232)
<b>Table 2.6</b>	Correlations among ESSeNcE Subscales, PHQ-9, GAD-7 and BAES
<b>Table 2.7</b>	Average Inter-Item Correlations (AICs) for the three ESSeNcE Subscales
<b>Table 2.8</b>	Internal Consistency of ESSeNcE Subscales
<b>Table 2.9</b>	Reliability Statistics for the ESSeNcE Scale (Split-Half Method)
<b>Table 2.10</b>	Total Descriptive Statistics for ESSeNcE Subscales

## List of Figures

**Figure 1.1** PRISMA Flow Diagram

## Acknowledgements

Firstly, I would like to acknowledge and thank my academic supervisors, Dr. Andy Siddaway and Professor Hamish McLeod for their expertise and guidance throughout the course of my doctoral research. I am most grateful for their insight and constructive feedback, which have been fundamental in shaping this work and made it all possible.

I would also like to express deepest thanks to all of the participants who took the time to complete my survey and share their feedback. I am especially grateful to Professor Andrew Ortony, who gave so kindly of his time and provided thoughtful input amidst what I am sure is an exceptionally busy schedule.

By no means least, I offer my sincerest and heartfelt thanks to three very special people - Dr. Gavin Richardson, Dr. Andrew MacDougall and Dr. Mary Reid - for their unwavering support during the most challenging times of my doctoral journey. Their generosity, kindness and understanding provided the strength I needed to persevere and complete this work. I am profoundly grateful to them all and genuinely could not have done this without them.

My appreciation also extends to all the lecturers and administrative staff on the DClinPsy for their help and support throughout the programme. I am also deeply grateful to Paul Cannon and Chris O'Malley, librarians at the University of Glasgow and NHS Highland respectively, for their invaluable assistance with the systematic review. My thanks also go to Dr. Alice Thomson for her assistance and being a first-rate second-rater.

On a personal note, I owe a debt of thanks to Anna, Maria and Tara, whose messages, phone calls and voice notes deserve their own doctorate for services rendered in friendship, support and stress relief. Thank you, ladies, for making the hard days easier and the good days better. Also, my sister, Sarah; I am so delighted to have reconnected with you, and it is a reminder that positives can emerge even from the toughest of times. Love you lots and jelly tots!

I am also grateful to my mother; a force of nature whose strength and determination defined her until the very end.

Thanks also go to my partner, Steve, for his genuine pride in me and steadfast commitment to the provision of perfectly chilled fizzy wine at the end of a working week (and sometimes in the middle too!)

Finally, I wish to express deepest love and gratitude to my daughter, Kirsten, and granddaughter, Lottie-Marie, who fill my heart every single day with love and laughter. Their presence - and huggles - have been a constant source of purpose, strength and inspiration throughout this journey. I love you both so much and am proud beyond words.

---

*This thesis is dedicated to my Parents:  
Warwick Charles Minnery: 15<sup>th</sup> August 1947 - 9<sup>th</sup> December 2025*

*Ann Borthwick: 17<sup>th</sup> November 1948 - 9<sup>th</sup> April 2026*

---

## **Chapter 1: Systematic Review**

### **Reliability and Validity of Current Measures of Emotional Schema: A Systematic Review**

Prepared in accordance with Clinical Psychology Review: Guide for authors - [Guide for authors - Clinical Psychology Review - ISSN 0272-7358 | ScienceDirect.com by Elsevier](#)

## Abstract

### *Background:*

Cognitive and metacognitive models propose that the ways in which individuals understand, interpret, and respond to their emotions - often conceptualised as emotional schemas - play a fundamental role in both the development and maintenance of psychological distress. Emotional schemas influence how people perceive the causes, controllability, and consequences of their emotions, which in turn affects their capacity for regulation and coping. Reliable and valid measures are therefore essential to capture how individuals conceptualise and manage their feelings, to deepen theoretical understanding, and to inform the development of more targeted and effective psychological interventions.

### *Method:*

A systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Guidelines. Searches were performed across five electronic databases - CINAHL, MEDLINE, PsycINFO, PsychArticles and Psychology & Behavioural Sciences (via EMBASE) - to identify studies evaluating measures of emotional schemas or related constructs. Inclusion and exclusion criteria were applied, and reference lists of relevant papers were also screened to ensure comprehensive coverage of the literature.

### *Results:*

Eleven studies met inclusion criteria and provided sufficient data for extraction and synthesis. These studies evaluated a range of instruments designed to assess emotional schemas, beliefs about emotions, or related cognitive-affective processes. Across the included studies, evidence generally supported acceptable levels of internal consistency and factorial validity. However, methodological limitations and variability in reporting standards reduced the strength of the overall evidence base.

### *Conclusions:*

Although existing measures of emotional schemas demonstrated adequate internal consistency and structural validity, there was limited evidence regarding reliability over time, cross-cultural applicability, criterion-related validity, and responsiveness to clinical change. Overall, findings highlight the need for further psychometric development and evaluation of emotional schema measures. Future research

should prioritise longitudinal, cross-cultural, and clinical validation studies to ensure these tools are robust, generalisable, and sensitive to therapeutic change.

**Word count:** 298

**Keywords:** *Emotional schema; Measurement; Reliability; Validity; Psychometrics.*

## Introduction

Emotions are an integral part of the human experience, influencing how we perceive ourselves and the world around us. They guide our behaviours, shape relationships, and affect overall psychological well-being (Edwards & Wupperman, 2018). From the moment we experience emotions such as happiness, sadness, fear, and anger, we begin to interpret, evaluate and respond to them. These emotional experiences are not purely physiological reactions; they are profoundly influenced by our thoughts, beliefs and perceptions of these emotions, which can be collectively referred to as 'emotional schemas'.

### *Conceptualising Emotional Schemas and Their Meta-Cognitive Regulation*

Emotional schemas are cognitive structures and beliefs about one's emotions and their functions, which organise emotional experiences and guide interpretations of future emotional events. They play a central role in emotion regulation, influencing both psychological well-being and distress depending on their adaptiveness (Leahy, 2002; Manser, Cooper, & Trefusis, 2012). How individuals appraise their emotions affects not only the strategies they use to regulate them, but also the likelihood of maintaining psychological equilibrium over time, highlighting the importance of these schemas for mental health outcomes (Leahy, 2015). Although emotions such as anxiety, fear, anger and sadness are universal and normative, they can become dysfunctional when they are persistent, intense, or disproportionate to the triggering context. Maladaptive patterns of emotional experience, compounded by ineffective regulation strategies such as rumination, can contribute to the onset and maintenance of chronic psychological difficulties, including anxiety, depression, and personality disorders (Leahy, 2002; Gross, 1998; Manser et al., 2012).

Maladaptive schemas - such as beliefs that emotions are inherently dangerous, uncontrollable, or a sign of weakness - are associated with ineffective coping strategies, including suppression, avoidance, and emotional withdrawal, which can exacerbate distress and interfere with adaptive functioning (Edwards & Wupperman, 2018). In contrast, adaptive schemas characterised by acceptance, understanding, and recognition of emotions as meaningful experiences, supporting effective emotional processing, flexible coping, and psychological resilience.

The meta-cognitive perspective provides a useful framework for understanding emotional schemas. Emotional schemas represent the content of appraisals - the beliefs and interpretations individuals hold

about their emotions - while metacognitive processes guide how this content is processed, evaluated, and acted upon (Wells, 2000). These processes include monitoring one's emotional state, evaluating its significance and implementing strategies for managing or responding to emotions. Negative metacognitive beliefs, such as perceiving emotions as dangerous, uncontrollable, or unacceptable, often trigger repetitive thinking patterns, including rumination and worry. Rumination, defined as persistent reflection on the causes and consequences of emotional experiences, amplifies negative affect and reinforces maladaptive patterns, frequently leading to ineffective regulation strategies such as avoidance or suppression (Wells, 2000; Manser, Cooper & Trefusis, 2012). Over time, this can create a self-perpetuating cycle: dysfunctional beliefs trigger rumination, which promotes maladaptive coping and maintains emotional distress, making it increasingly difficult to restore balance and resilience.

Closely related to rumination is worry, which is also conceptualised as a form of repetitive negative thinking but differs in its time-related focus and functional role. Whereas rumination is predominantly past- or present-focused and involves dwelling on perceived losses, failures or distress, worry is typically future-orientated and characterised by repetitive thoughts about potential threats, uncertainties or negative outcomes (Borkovec et al., 1983; Watkins, 2008). Worry has been described as a process involving linear, internally verbal thought, aimed at anticipating and preventing future harm, often functioning as a form of cognitive avoidance (Borkovec et al., 2004). Despite these distinctions, rumination and worry share core features, including repetitiveness, abstract processing and a perceived sense of uncontrollability, and are increasingly viewed as sub-types of a broader construct of repetitive negative thinking (Ehring & Watkins, 2008). From an emotional schema framework, both processes represent maladaptive attempts to manage emotional experiences in the context of negative beliefs about the meaning, duration and tolerability of emotions. Differentiating between rumination and worry is therefore clinically valuable, as it clarifies how distinct emotional schemas give rise to specific forms of self-focused thought and informs the selection of targeted, schema-focused intervention strategies

### *Dimensions of Emotional Schemas*

Emotional schemas encompass a wide range of beliefs and appraisals regarding emotional experiences. Key dimensions include:

- **Intensity and duration of emotions:** whether emotions are experienced as overwhelming, persistent, or uncontrollable.
- **Causes and consequences:** how emotions are attributed, including internal versus external explanations.
- **Implications for self-concept:** whether emotions are interpreted as reflecting personal weakness, failure, or strength.
- **Appropriate regulation strategies:** beliefs about whether emotions should be expressed, suppressed, or avoided, and the acceptability of different regulatory behaviours (Edwards & Wupperman, 2018).

These beliefs shape how individuals experience, interpret and respond to emotions. Maladaptive emotional schemas - such as viewing emotions as harmful, uncontrollable, or shameful - can contribute to emotional dysregulation and the development of psychological disorders. Conversely, adaptive schemas characterised by acceptance and constructive interpretation of emotions, facilitate flexible coping, effective emotional processing, and greater resilience. Understanding these schemas and their interaction with metacognitive processes is essential for mental health research and clinical practice, particularly when emotional dysregulation plays a central role.

### *Measuring Emotional Schemas*

Given their importance in the regulation of emotions and mental health, emotional schemas are often assessed through self-report scales. One of the most widely used tools for assessing emotional schemas is the Leahy Emotional Schema Scale (LESS), which was developed to measure individuals' beliefs and attributions about their own and others' emotions. The LESS captures a broad range of emotional schema dimensions, such as validation, guilt, acceptance, and blame. These dimensions are critical for identifying negative emotional concepts related to emotional disorders, such as anxiety and depression (Suh et al., 2010). The LESS has been shown to be particularly useful for examining how emotional schemas relate to emotion regulation strategies, as maladaptive schemas often lead to ineffective coping approaches, such as avoidance or rumination (Leahy, 2002).

In addition to the LESS, several other tools have been developed to measure related constructs. For example, the Emotion Beliefs Questionnaire (EBQ; Becerra et al., 2020), the Beliefs about Emotions Questionnaire (BAEQ; Manser, Cooper, & Trefusis, 2012), and the Beliefs about Emotions Scale (BES; Rimes & Chalder, 2010), which capture aspects of the emotional experience, particularly meta-cognitive beliefs about emotions. However, all of these instruments, including the LESS, exhibit limitations. Existing measures of emotional schemas frequently suffer from restricted cultural and ecological validity, in part because they rely on small, homogenous, or otherwise non-representative samples - often undergraduate student populations - which limits external validity and can compromise the stability and generalisability of the resulting factor structures (Rossellini & Brown, 2021; Worthington & Whittaker, 2006). These limitations are compounded by a lack of qualitative data in the scale-development process; without participants' lived experiences to guide item generation, quantitative analyses alone are insufficient to produce comprehensive or inclusive assessments of emotional schema content (Rowan & Wulff, 2007; Worthington & Whittaker, 2006; Padgett, 2016). Consequently, many existing tools fail to capture the full range of ways in which individuals experience, interpret, and evaluate their emotions, thereby constraining their ability to provide a thorough and clinically meaningful assessment.

### *Rationale for the Current Review*

Despite growing interest in emotional schemas and their role in psychological functioning, there is currently no comprehensive systematic review critically appraising measures of emotional schemas and related constructs in adults. Whilst a handful of reviews exist, they either focus on different age groups, evaluate broader concepts or include some non-self-report measures. Emotional schemas or evaluations of own or others' emotions, by their nature, necessitate the use of self-report instruments.

Myers et al. (2019), synthesised evidence on the Metacognitions Questionnaire (MCQ; Cartwright-Hatton & Wells, 1997) and its derivatives in children and adolescents. It highlighted the variability in reliability and validity across versions and the need for further psychometric evaluations. Focusing on adult populations is particularly important, however, as emotional schemas are theoretically and empirically more consolidated compared with younger populations (Nook & Somerville, 2019), and are central to clinical interventions and research on emotions. More recently, Peter and colleagues (2025) conducted a systematic review of emotion beliefs measures, evaluating validity and reliability across

multiple instruments designed to assess people's beliefs about emotions, using the Quality Assessment with Diverse Studies (QuADS; Halle et al., 2016). They found that existing measures are useful, but highlighted the need for more consistent frameworks that can capture emotion beliefs at different stages of emotional and cognitive development, rather than applying adult-focused questionnaires to younger populations. However there were some important differences between Peter et al.'s review and the present work. The authors imposed no restrictions on study design or research paradigm, including cross-sectional, longitudinal, experimental, qualitative, and mixed methods studies; they also did not limit the specific dimensions of emotion beliefs examined, i.e. whether the measures addressed general or personal beliefs about emotions. Furthermore, no restrictions were placed on the age of the participants included in the studies, or whether included measures were self-report.

Although these reviews provide valuable insights, they are limited in scope and do not focus specifically on evaluation of emotions, as conceptualised by emotional schema theory. The present review addresses this gap by synthesising available instruments and evaluating their psychometric properties. Specifically, it examines methodological rigour and psychometric properties of self-report tools assessing emotional schemas in individuals aged 16 and above. By highlighting the strengths, limitations and areas requiring further development, this review aims to inform both future research and clinical applications.

#### *COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)*

The COSMIN methodology is the only consensus-based approach specifically designed to evaluate patient-reported outcome measures (PROMs) (Mokkink et al., 2018; Mokkink et al., 2010a, 2010b). It is a framework that helps to assess the reliability, validity, responsiveness and interpretability of health measurement tools (such as questionnaires, scales, and surveys) used in clinical trials or observational studies. It provides standards for assessing measurement instruments, guidelines on how to evaluate the measurement properties of instruments and is a checklist for researchers to assess the quality of studies evaluating PROMs. COSMIN's goal is to promote standards for the quality evaluation of measurement instruments used in health research, ensuring that the tools used to collect data are scientifically robust and valid. Its methodology not only provides clear definitions and guidelines for evaluating various measurement properties, but also includes detailed criteria for assessing content

validity, which is regarded as one of the most important characteristics of a measurement instrument - the degree to which an instrument reflects the construct it is intended to measure (Terwee et al., 2018). This process is essential to ensuring that the measure accurately captures concepts of interest and is acceptable to the target population for which it is designed.

This review therefore aimed to fill an important gap in the literature by providing a comprehensive evaluation of the tools available to measure emotional schema. By identifying strengths, weaknesses and areas for improvement in these tools, this review sought to inform both future research and clinical practice, highlighting the most effective measures for understanding emotional schema and their role in the regulation of emotions.

## **Method**

### *Protocol Registration*

To verify that the review questions had not already been answered, a search was conducted on PROSPERO, the Database of Abstracts of Reviews of Effects, and the Cochrane Database of Systematic Reviews for any similar existing or ongoing reviews. The review protocol was created following the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. (PRISMA; Page et al., 2021); full details of the PRISMA Checklist and PRISMA Abstract Checklist are provided in Appendix 1.1. The review was registered on PROSPERO (Project ID: PROSPERO 2025 CRD420250619288) and is accessible online at <https://www.crd.york.ac.uk/PROSPERO/view/CRD420250619288>.

### *Search Strategy*

The search terms for this review were developed based on a comprehensive appraisal of existing literature on the development and validation of emotional schema measurement tools, as well as measures assessing beliefs or evaluations of emotions. In conducting this systematic review, the COSMIN search terms related to psychometric properties, and validation and measurement tools were employed to ensure a comprehensive evaluation of the relevant instruments. Other key terms were also identified from relevant papers that discussed the conceptualisation, development and validation of such instruments across clinical and non-clinical populations. The search strategy employed in this

systematic review was developed and reviewed with the assistance of librarians from the University of Glasgow and NHS Highland to ensure comprehensive coverage and methodological rigor (see Appendix 1.2 for an outline of the full search strategy). The refined search terms ensured that the search would capture a wide range of studies related to emotional schema measures, including their validity, reliability, and other psychometric properties. This collaborative approach helped ensure a thorough and systematic search process, aimed at identifying the most relevant and high-quality PROMs.

Published literature was retrieved from the following electronic databases: CINAHL, MEDLINE, PsycINFO, PsychArticles and Psychology and Behavioural Sciences (PBS). The search period covered publications from 1st January 1970 until the extraction date, on 29<sup>th</sup> August 2025. The start date was selected to ensure comprehensive coverage of the emergence modern psychometric theory and the development of self-report psychological measures. Although emotional schemas were defined later, earlier work on beliefs and evaluations of emotions is conceptually relevant and may underpin the development and validation of current instruments, ensuring comprehensive coverage of relevant PROMs. All retrieved records were uploaded into EndNote X9 referencing software for de-duplication and initial screening.

Articles were initially screened for eligibility based on title and abstract. The second reviewer (AT) screened 10% of randomly selected papers at the title and abstract stage, with an agreement of 95% between both reviewers. For full-text papers, 10% were independently reviewed by AT yielding 100% agreement between raters.

In addition to database searches, a manual search of the reference lists of included papers and a forward-search of their citations were conducted in order to identify further relevant studies. This process led to the discovery of two additional eligible papers. This systematic approach, incorporating both database searches and manual methods, ensured that the review captured all relevant studies concerning the development and validation of measures for metacognition of emotions.

### *Eligibility Criteria*

Studies meeting the following criteria were included for review:

- Those featuring the development of a measure of beliefs, appraisals or concepts about one's own or others' emotions
- Measures which were self-report
- For comparative purposes, studies which were quantitative in design
- Studies with adult samples aged 16 or over
- Studies published in peer reviewed journals, in English language

Exclusion criteria were:

- Studies where the developed measure was related to expression of emotion
- Book chapters, review articles, responses to articles (e.g. letters/ reviews) and dissertations

Whilst this review focused on metacognitions of emotions in adults, a post hoc broadening of the inclusion criteria to include samples where the age range was from 16 years enabled two additional studies to be included in the review.

#### *Data Extraction*

Characteristics of the studies (e.g., sample, study setting, location) and included measures (e.g., mode of completion, number of emotional schema items, were collated. See Appendix 1.3 for full data extraction checklist. Statistical and methodological information was extracted and evaluated to rate the quality of the measures using the manualised COSMIN approach, see Table 1.1 below for COSMIN definitions of measurement properties (Mokkink et al., 2018; Mokkink et al., 2010a, 2010b; Prinsen et al., 2018; Terwee et al., 2018). See Appendix 1.4 for full table.

**Table 1.1: COSMIN definitions of measurement properties**

<b>Measurement Property</b>	<b>Definition</b>
<i>Internal Consistency</i>	The degree of the interrelatedness among the items
<i>Test-retest</i>	The degree to which individual's scores have not changed and are the same for repeated measurement over time
<i>Measurement Error</i>	The systematic and random error of an individual's score that is not attributed to true changes in the construct to be measured
<i>Content Validity</i>	The degree to which the content of an instrument is an adequate reflection of the construct to be measured
<i>Structural Validity</i>	The degree to which the scores of an instrument are an adequate reflection of the dimensionality of the construct to be measured
<i>Hypothesis Testing for Construct Validity</i>	The degree to which the scores of an instrument are consistent with hypotheses (for instance regarding internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the instrument reliably measures the construct to be measured
<i>Cross-cultural Validity</i>	The degree to which the performance of the items on a translated or culturally adapted instrument are an adequate reflection of the performance of the items of the original version of the instrument
<i>Criterion Validity</i>	The degree to which the scores of an instrument are an adequate reflection of a 'gold standard'
<i>Responsiveness</i>	The ability of an instrument to detect meaningful changes over time

*Methodological quality of included instruments*

The COSMIN manual was used to assess the methodological quality and measurement properties of the included studies. The risk of bias aided methodological quality assessment through the evaluation of the strength of psychometric properties: reliability (internal consistency, test-retest and measurement error), validity (content/face validity, structural validity/hypothesis testing, cross-cultural validity) and responsiveness. The development process was judged using criteria from the COSMIN manual (Mokkink et al., 2018) and each of these psychometric properties was rated as 'very good', 'adequate', 'doubtful', 'inadequate' or 'not applicable'. The results were summarised qualitatively to determine the overall psychometric quality of each tool. To ensure consistency, the second reviewer evaluated the psychometric quality of a random 20% sample of the included studies. There was 87.5% agreement between raters, with any discrepancies resolved through discussion and a consensus reached.

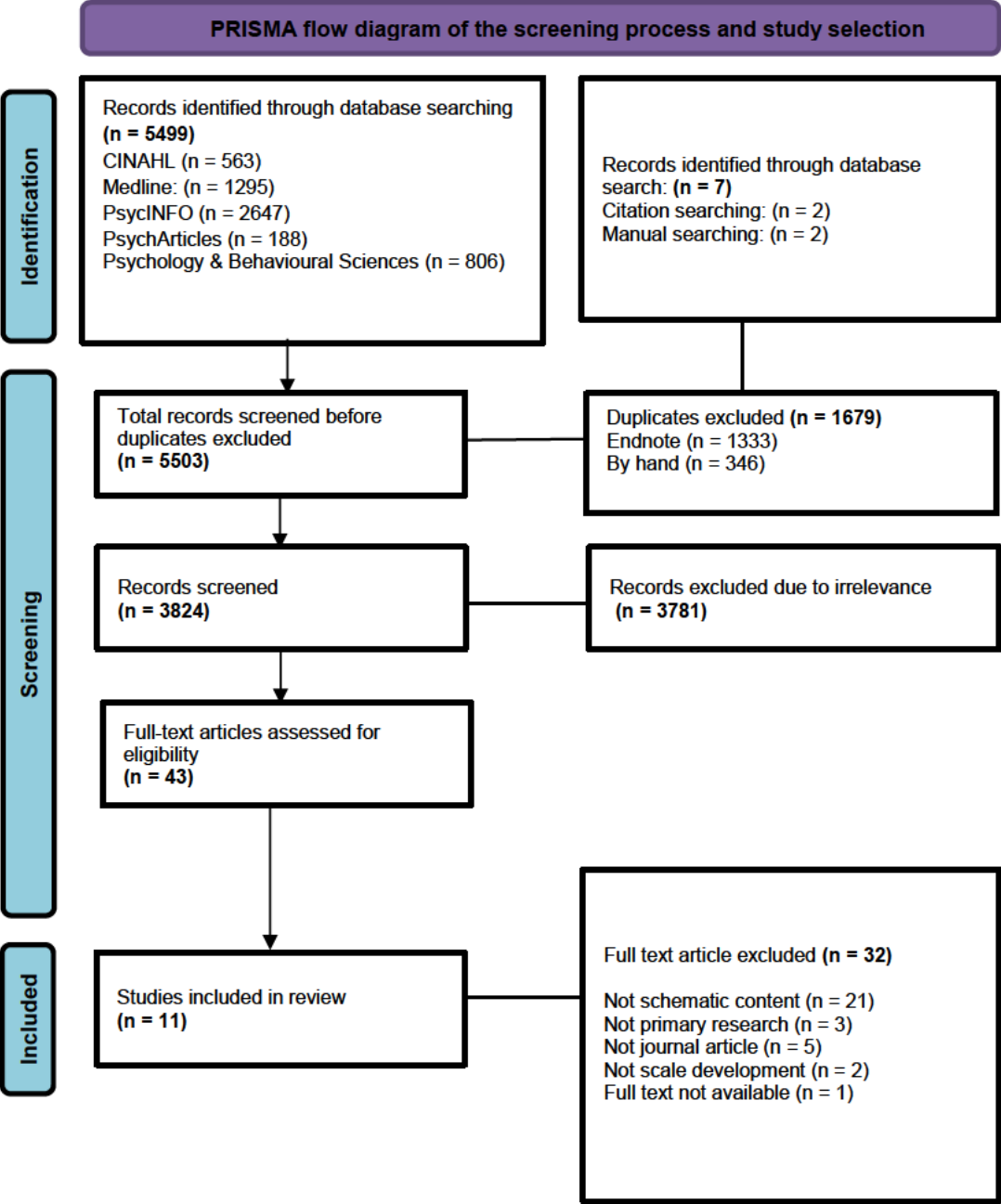
## Results

### *Overview of Included Studies*

5499 records were identified from the database search and a further four were found from the reference lists and citations of eligible articles. After removal of 1679 duplicates, 3824 were screened at title/abstract stage, of which 3781 were excluded, leaving 43 papers for full-text review. Of these, 32 were excluded for not meeting eligibility criteria, resulting in a total of eleven studies being included in the review. Included studies are summarised in Table 1.2. The table is colour-coded to show the source from which each study was identified (i.e. database, citation or hand search).

Six studies were conducted in Europe (UK:  $n=4$ , Austria  $n=1$ , Germany:  $n=1$ ) and the remaining were conducted in the United States ( $n=3$ ) and Australia ( $n=2$ ). Across all studies, the pooled sample size was 6966, with sample sizes ranging from 53-1175. Gender breakdown, this comprised of 2550 males (36.6%); 4358 females (62.6%) and 58 who identified as 'other' (0.8%). Of the studies included in this review, six exclusively utilised clinical samples, reflecting populations engaged with mental health services. An additional four studies incorporated mixed samples, comprising both clinical and non-clinical participants, thereby enabling comparisons across differing levels of psychological need. Of the studies utilising clinical samples, two were made up of participants with mental health problems, one sample had a physical health condition (chronic fatigue syndrome) and the remaining study comprised of a mix of psychological and somatic illnesses (e.g. mood/anxiety disorders, diabetes, chronic pain and eating disorders). One study (Leahy, 2002) recruited a sample made up solely of individuals receiving psychiatric care, with no inclusion of non-clinical participants. Mean participant ages ranged from 19.4-47.3. All of the included studies had a larger number of female to male participants. Three of the eleven studies utilised university students.

Figure 1.1 *Prisma Flow Diagram*



**Table 1.2: Overview of Included Studies**

Authors	Measure	Location	No of items	N	Population	Age M (SD) Range	Gender Male/ Female/ Other
Becerra et al (2019)	Perth Emotional Reactivity Scale (PERS)	Australia	30	183	Non-clinical	32.36 (10.05) 18-67	72/ 111
Becerra et al (2020)	Emotion Beliefs Questionnaire (EBQ)	US	40	161	Non-clinical	47.03 (18.02) 18-83	77/ 84
Gratz et al (2004)	Difficulties in Emotional Regulation Scale (DERS)	UK	36	Study 1: 357 Study 2: 194	University students (undergrad psychology) University students	23.10 (5.67) 18-55 25.95 (8.94) 18-48	97/ 260 74/ 120
Greenberg et al (2021)	Brief- Mentalized Affectivity Scale (BMAS)	US	12	Study 1: 978 Study 2: 230	Non-clinical Community mental health clinic	36.20 (11.46) 18-65 28.36 (9.38) 18-65	349/ 623/ 6 74/ 52/ 4
Kasper et al (2024)	Mentalizing Emotions Questionnaire (MEQ)	Germany	16	Study 1: 510 Study 2: 509	Non-clinical Current and previous mental health problems	43.3 (13.8) 18-65 44.0 (13.2) 18-65	252/ 255/ 3 235/ 271/ 2
Leahy (2002)	Leahy Emotional Schema Scale (LESS)	UK	50	53	Psychiatric patients	40.2** 23-69	21/ 32
Manser et al (2012)	Beliefs About Emotions Questionnaire (BAEQ)	UK	102	288	Non-clinical	28.05 (10.99) 16-**	74/ 214
Mitmansgruber et al (2009)	Meta-Emotion Scale (MES)	Austria	28	Study 1: 334 Study 2: 222 Study 3: 297	Students Students Mixed clinical population (psychological and physical health issues)	22.37 (2.81) 19-42 22.8 (3.8) 19-50 45.0 (11.4) **	153/ 181 43 92/ 130 54/ 243

Preece et al (2018)	Perth Emotion Regulation Competency Inventory (PERCI)	Australia	32	Study 1: 231	Non-clinical	41.52 (16.93) 18-65	80/ 151
				Study 2: 1175	Non-clinical	43.22 (16.58) 18-88	356/ 815
Rimes & Chalder (2010)	The Beliefs about Emotions Scale (BES)	UK	12	194	121 (CFS)	38.8 (11.9) **	29/ 92
					73 (Non-clinical)	37.6 (11.7) **	24/ 49
Veilleux et al (2015)	Emotion and Regulation Beliefs Scale (ERBS)	US	21	Study 1: 255	Non-clinical	34.05 (12.45) **	117/ 138
				Study 2: 281	Student/Non-clinical	19.4 (2.47)/ 34.54 (13.09) **	90/ 191 230/346

**\*\*not available**

	Database search
	Citation search
	Hand search

### *Measure Characteristics*

Table 1.3 below provides an overview of key information about the characteristics of each included measure characteristics. (See Appendix 1.5 for further information on included scales and their utility).

### *Constructs Measured*

While all of the measures addressed beliefs about or attitudes towards emotions, there were a variety of specific sub-constructs measured within this broad domain. Specifically, five out of the eleven tools explored beliefs about the controllability of emotions, highlighting how individuals perceive their ability to manage emotional experiences. Two of the measures focused on the perceived utility or usefulness of emotions, assessing whether emotions are seen as helpful or valuable. Five measures examined beliefs regarding the acceptability of emotions, capturing attitudes about whether it is appropriate or acceptable to experience certain emotions. Additionally, three measures addressed beliefs related to the intensity of emotions, while two assessed beliefs about the duration or persistence of emotional experiences.

Beyond beliefs, five measures evaluated individuals' capacity to reflect upon and understand their own emotions, emphasising emotional awareness and insight. Although most measures concentrated on self-directed emotions, two instruments also incorporated attitudes toward the emotions of others, acknowledging the social and interpersonal aspects of emotional experience. Other important constructs covered included maladaptive beliefs about emotions, such as perceptions of emotions as dangerous or harmful (measured in two of the studies), and mentalising abilities, which refers to the capacity to understand and interpret one's own and others' emotional states (also examined in two tools).

### *Number of Items*

The BMAS and BES contained the smallest number of questionnaire items (12), and the longest scale was the BAEQ at 102 items. Most of the other scales contained 20 to 50 items.

### *Scoring Method/Response Options*

Most of the scales used an agreement- or frequency-based Likert response scale, thereby allowing for some flexibility in response granularity, depending on the construct being measured. Likert scales

generally range between five and seven options. Five of the measures used five-point Likert scales; three measures used six-point Likert scales, and four measures incorporated seven-point Likert scales. The latter provides a more refined and varied response range, thereby increasing sensitivity and detecting subtle differences in attitudes or beliefs. The response anchors across all scales generally fell into categories such as agreement, frequency, or self-descriptive judgment (e.g. 'Strongly disagree' to 'Strongly agree', 'Very unlike me' to 'Very like me' and 'Almost never' to 'Almost always').

#### *Factor Analysis Methods*

Across the eleven measures reviewed, most were developed using exploratory analytic techniques. Specifically, four studies employed a combination of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), while two used CFA alone. Of the remaining measures, two relied solely on EFA, one used EFA alongside common factor analysis, and two applied principal components analysis (PCA). EFA, specifically common factor analysis, is used to examine the underlying latent structure of data. Although PCA is also exploratory, it is not a true factor analytic technique. Instead, PCA is a data reduction method that transforms observed variables into components that account for as much variance as possible, without modelling latent factors. This contrasts with CFA which is theory-driven and used for hypothesis testing and model confirmation (Brown, 2015; Kline, 2016)

**Table 1.3 Measure Characteristics**

Measure	Reference	Construct Measured	Number of items	Scoring Method	Response options	Factor Analysis Methods
Perth Emotional Reactivity Scale (PERS)	Becerra et al (2019)	Emotional reactivity across different emotions, capturing intensity, duration, and speed of emotional responses	30	5-point Likert scale	1= Very unlike me 5 = Very like me	EFA/CFA
Emotion Beliefs Questionnaire (EBQ)	Becerra et al (2020)	Beliefs about emotions, specifically focusing on beliefs about the controllability and usefulness of emotions	40	6-point Likert scale	1 = Strongly disagree 6 = Strongly agree	CFA
Difficulties in Emotional Regulation Scale (DERS)	Gratz et al (2004)	Multiple dimensions of emotion regulation difficulties, including awareness, clarity, acceptance of emotions, impulse control, goal-directed behaviour, and access to effective regulation strategies	36	5-point Likert scale	1 = Almost never 5 = Almost always	EFA and Common Factor Analysis
Brief- Mentalized Affectivity Scale (BMAS)	Greenberg et al (2021)	The capacity to reflect on, understand, and regulate one's own emotions of others, within a mentalising framework	12	7-point Likert scale	1 = Strongly disagree 7 = Strongly agree	PCA
Mentalizing Emotions Questionnaire (MEQ)	Kasper et al (2024)	The capacity to understand and interpret one's own and others' emotional states within a mentalising framework.	16	7-point Likert scale	1 = Almost never 7 = Almost Always	EFA/CFA
Leahy Emotional Schema Scale (LESS)	Leahy (2002)	Individuals' emotional schemas beliefs and attitudes about emotions, including how they understand, manage, and react to their own emotional experiences	50	6-point Likert scale	1 = Very untrue of me 6 = Very true of me	EFA
Beliefs About Emotions Questionnaire (BAEQ)	Manser et al (2012)	Assessing maladaptive beliefs about emotions, such as beliefs that emotions are unacceptable, uncontrollable, or harmful	102	5-point Likert scale	1 = Strongly disagree 5 = Strongly agree	EFA
Meta-Emotion Scale (MES)	Mitmansgruber et al (2009)	How people think and feel about their own emotions, specifically their attitudes toward accepting or judging their emotions	28	6-point Likert scale	1 = Not at all true for me 6 = Completely true for me	EFA/CFA
Perth Emotion Regulation Competency Inventory (PERCI)	Preece et al (2018)	Emotion regulation difficulties in terms of how well individuals understand, accept, and manage both positive and negative emotions	32	7-point Likert scale	1 = Strongly disagree 7 = Strongly agree	CFA
The Beliefs about Emotions Scale (BES)	Rimes & Chalder (2010)	Beliefs people hold about the acceptability, controllability, and danger of experiencing emotions, particularly negative ones	12	7-point Likert scale	1 = Totally disagree 7 = Totally agree	PCA
Emotion and Regulation Beliefs Scale (ERBS)	Veilleux et al (2015)	Beliefs about emotions and emotion regulation, including beliefs about the	21	5-point Likert scale	1= Strongly disagree 5 = Strongly agree	EFA/CFA

		controllability, utility, and acceptability of emotions and emotion regulation strategies.				
--	--	--	--	--	--	--

## Methodological Quality of Included Studies

The following section of the review outlines the COSMIN-based assessment (Mokkink et al., 2024) of each measure's psychometric properties. See Table 1.4 below for a summary.

### Reliability

#### *Internal Consistency*

Subscale internal consistency was reported in nine of the eleven studies included in the review. The reported Cronbach's alpha values ranged from 0.70 to 0.94, with all values meeting the recommended threshold of  $\alpha \geq 0.70$ . Internal consistency data were not provided for two studies, specifically the BMAS and LESS. Among the measures that were assessed, the MEQ and BES consistently showed the highest internal consistency, with Cronbach's alpha values exceeding 0.90 across the studies in which they were reported.

#### *Test-retest*

Test-retest reliability reflects the stability of a measure over time and is an important indicator of whether an instrument consistently captures a construct expected to remain relatively stable. Given that emotional schemas are typically stable traits rather than temporary states, significant changes over a short period would not be expected. Consequently, responsiveness - the ability of a measure to detect change over time - was not systematically evaluated in any of the studies, as it is of limited relevance for trait-based constructs. Only one study in the review - the validation of the DERS - assessed test-retest reliability, reporting an intra-class correlation coefficient (ICC) of 0.88. None of the remaining ten measures evaluated or reported test-retest reliability, and these were consequently assigned a 'not applicable' rating.

#### *Measurement Error*

Measurement error refers to the difference between what is being measured and the true value of the quantity being measured, and arises due to errors in the measurement process. Measurement error was not assessed or reported in any of the studies included in the review and therefore all studies were rated 'not applicable'.

## **Validity**

### *Content/Face Validity*

Seven out of the eleven studies were rated as having 'adequate' content validity, therefore demonstrating that these measures were an adequate reflection of the constructs being measured. The ERBS and DERS were the only measures to demonstrate 'very good' face validity, reflecting a notably positive subjective evaluation by experts. Two measures received a 'doubtful' rating in this domain (BMAS, MEQ).

### *Construct Validity/Hypothesis Testing*

All eleven studies included in the review assessed construct validity. Of these, three studies (PERS, BMAS, and PERCI) were rated as demonstrating 'very good' psychometric quality with respect to this domain. The remaining studies were judged to be of 'adequate' quality, indicating that while their approaches to hypothesis testing met the minimum standards, they provided less comprehensive or less robust evidence compared to the higher-rated measures.

### *Cross-Cultural Validity*

None of the measures included in this review explored cross-cultural validity, as detailed in the COSMIN guidelines. The PERS, DERS and EBQ have been translated into other languages; for example, Polish, Turkish, Iranian, Russian (PERS), French, German, Czech, Portuguese (DERS) and Japanese (EBQ). The main researcher was unable to locate documented translations of other measures. The investigations included in this review were conducted exclusively within a single linguistic and cultural context, however, ensuring consistency of findings by controlling variability which would be introduced by cross-cultural differences. The translated studies did conduct formal multi-group CFA to assess invariance across demographics, (e.g. gender), but not for cultural or linguistic groups. For example, Larionow and Mudło-Głagolska (2022) validated the PERS with 675 participants and found full stability across males and females, but this was not extended to culture. The absence of formal cross-cultural validation limits confidence in the equivalence of the construct across languages and cultural settings. Without establishing measurement invariance, it cannot be assumed that the instrument assesses the same latent constructs in the same way across groups, raising concerns about construct validity, score comparability, and the potential for cultural bias. As Byrne and van de Vijver (2010) argue, failure to test for cross-cultural measurement invariance undermines interpretability of group comparisons and may lead to erroneous conclusions about psychological differences between populations.

### *Criterion Validity*

As described by Mokkink et al. (2024), criterion validity indicates how strongly a measurement instrument relates to or predicts an external criterion. In the current review, two out of the eleven studies assessed criterion validity (EBQ and PERCI), and were rated as 'adequate'. There was no established benchmark emotional schema measure in the remaining nine studies

Table 1.4: Evaluations of included studies according to the COSMIN guidelines (Mokkink et al., 2010)

Measure	RELIABILITY			VALIDITY				RESPONSIVENESS
	Internal consistency	Test-retest	Measurement error	Content/Face Validity	Construct validity/ Hypothesis testing	Cross Cultural Validity	Criterion Validity NA	Responsiveness
Becerra et al (2019)	A	NA	NA	A	VG	NA	NA	NA
Becerra et al (2020)	A	NA	NA	A	A	NA	A	NA
Gratz et al (2004)	VG	A	NA	VG	VG	NA	NA	NA
Greenberg et al (2021)	NA	NA	NA	D	VG	NA	NA	NA
Kasper et al (2024)	VG	NA	NA	D	A	NA	NA	NA
Leahy (2002)	NA	NA	NA	A	A	NA	NA	NA
Manser et al (2012)	A	NA	NA	A	A	NA	NA	NA
Mitmansgruber et al (2009)	A	NA	NA	A	A	NA	NA	NA
Preece et al (2018)	VG	NA	NA	A	VG	NA	A	NA
Rimes & Chalder (2010)	VG	NA	NA	A	A	NA	NA	NA
Veilleux et al (2015)	A	NA	NA	VG	A	NA	NA	NA
VG - Very Good		D - Doubtful						
A - Acceptable		NA - Not applicable (i.e. not reported or assessed)						

## Discussion

This review sought to systematically identify and evaluate tools designed to assess the methodological rigor and psychometric soundness of existing measures of emotional schemas in clinical and non-clinical populations according to the COSMIN Risk of Bias criteria. Eleven studies were included in this review. No single measure currently fulfils all COSMIN criteria to a consistently high standard. Although many measures demonstrated good psychometric properties for internal consistency, structural and face validity, most measures demonstrated limitations in other domains, such as test-retest reliability, cross-cultural validity, and responsiveness.

### *Reliability*

Subscale internal consistency was reported for nine of the eleven measures included in this review, with Cronbach's alpha values ranging from 0.70 to 0.94, meeting or exceeding the commonly recommended threshold of  $\alpha \geq 0.70$ . This generally suggests that items within each subscale were sufficiently correlated to support internal coherence. However, it is important to note that alpha values calculated across all items of a scale can appear artificially high simply due to the large number of items, rather than reflecting true inter-item consistency (Clark & Watson, 1995). Internal consistency data were not reported for two measures, the BMAS and LESS, representing a gap in the psychometric evaluation of these instruments. Among the measures with reported reliability, the MEQ and BES demonstrated particularly high internal consistency, with alpha values exceeding 0.90 across studies. These findings suggest that some instruments may be more reliably structured than others, though reporting practices were inconsistent across studies.

Evidence of test-retest reliability was extremely limited. Only one study assessed and reported this statistic, focusing on the DERS, an ICC of 0.88, a value generally considered indicative of strong stability. All other instruments received a 'not applicable' rating in this category. The absence of systematic reporting on temporal stability highlights a significant gap in the literature, particularly given that test-retest reliability is essential for establishing whether an instrument can produce consistent results across repeated administrations over time. Without such evidence, the stability of many of the reviewed measures remains uncertain.

All instruments were rated 'not applicable' for measurement error, indicating that none evaluated how much observed scores might be affected by random or systematic error, which limits the strength of the

psychometric evidence. Accurately assessing measurement error usually requires latent variable modelling, such as CFA, which separates true score differences from error. Exploratory methods, including EFA and PCA, do not account for measurement error, meaning that for most of the reviewed measures, it remains unclear whether score variability reflects genuine differences in emotional schemas or error introduced during the measurement process.

Overall, while most instruments demonstrated acceptable internal consistency, evidence for other aspects of reliability - specifically test-retest reliability and measurement error - was sparse or absent. This imbalance in the psychometric evidence base indicates that although many of the measures may function well in terms of internal coherence, little is known about their stability across time or their susceptibility to error. Future research should prioritise more comprehensive reliability testing, including repeated measurement designs and formal evaluations of measurement error, to strengthen the evidence base for the consistent and accurate use of these instruments across different contexts.

### *Validity*

Overall, evidence for validity across the included measures was mixed. In terms of content, seven of the eleven studies were rated as adequate, indicating that items were generally representative of the intended constructs. The Emotion Regulation Beliefs Scale (ERBS) and the Difficulties in Emotion Regulation Scale (DERS) achieved 'very good' face validity, reflecting particularly positive expert evaluations of clarity and relevance. Two measures received a 'doubtful' rating in this domain - the Brief Mentalized Affectivity Scale (BMAS) and the Mentalizing Emotions Questionnaire (MEQ). One assesses the identifying, processing and expression of emotions (BMAS), and the other measures the capacity to reflect on emotions in oneself and others (MEQ). Both measures are grounded in theory on a conceptual level, but evidence to support content validity is inadequate. Since there is limited detail on systematic item development procedures, such as expert ratings, cognitive interviews or pilot testing, neither tool can be fully appraised in this domain in accordance with COSMIN criteria.

Structural validity was somewhat stronger. All studies assessed this domain, with four instruments - the Perth Emotional Reactivity Scale (PERS), DERS, Brief Measure of Affect Suppression (BMAS), and Perth Emotion Regulation Competency Inventory (PERCI) - rated 'very good', and supported by CFA. The remaining instruments were rated as adequate, suggesting acceptable but less rigorous evidence. While most scales displayed generally coherent factor structures, only a limited number demonstrated

clear and statistically robust evidence that their items accurately reflected the proposed underlying factors, raising questions about the validity of the remaining measures.

A notable gap concerned cross-cultural validity. None of the included studies formally tested measurement equality across cultural or linguistic groups, despite COSMIN recommendations. Although some measures have been translated (e.g., PERS into Polish, Turkish, Iranian, Russian; DERS into French, German, Czech, Portuguese; EBQ into Japanese), analyses typically focused on demographic invariance, such as gender. This limitation has important implications; without evidence of cross-cultural invariance, it cannot be assumed that these instruments assess the same constructs in the same way across different cultural or linguistic groups. As a result, comparisons of scores between populations may be biased, and the validity of conclusions drawn from such comparisons is uncertain (Byrne & van de Vijver, 2010).

It is also noteworthy that the majority of participants across studies were female (62.6%) with males representing 36.6% of the total sample. This gender imbalance may have implications for the generalisability of findings, as emotional schema expression and regulation can differ by sex, with females typically reporting greater emotional awareness, expressivity and variability than males (Brody & Hall, 2008; Nolen-Hoeksema, 2012). Consequently, some psychometric properties, particularly validity and factor structure, could be influenced by the overrepresentation of females, and future studies should aim to evaluate these measures in more gender-balanced samples to ensure applicability across sexes.

Criterion validity was the least examined. Only two instruments - the EBQ and PERCI - were assessed, with both rating as 'adequate'. The absence of this evaluation in most studies reflects the challenge of lacking a universally accepted gold standard for measuring emotional schema, which limits conclusions about external validity.

In sum, existing measures of emotional schemas show reasonable content and structural validity but lack robust evidence for cross-cultural and criterion validity. Addressing these gaps will require more comprehensive validation efforts, particularly cross-cultural measurement invariance testing and the development of agreed benchmarks, to enhance confidence in their use across populations and settings.

### *Responsiveness*

This was not systematically evaluated in any of the eleven instruments included in this review. As a result, all tools received a rating of 'not applicable' for this property, reflecting a critical gap in their capacity to assess longitudinal changes or treatment-related shifts in emotional schemas. Theoretical literature suggests that emotional schemas are trait-like in nature, reflecting stable patterns of emotional processing, but they are also modifiable through intervention, indicating a capacity for change over time (Leahy, 2002). The absence of responsiveness testing therefore limits confidence in these instruments' ability to capture clinically or developmentally meaningful changes. This limitation is particularly important for research and clinical contexts where tracking progress or intervention effects over time is essential. Without evidence of responsiveness, it is unclear whether these instruments can reliably capture meaningful changes, which restricts their utility for longitudinal studies and outcome evaluations.

### *Recommendations*

Based on the findings of this review, several recommendations can be made to strengthen the psychometric evaluation of emotional schema measures. Future research should prioritise more comprehensive reliability testing, including systematic assessment of test-retest reliability and measurement error, to establish temporal stability and accuracy. Efforts to evaluate and improve validity should extend beyond content and structural domains to include criterion validity, cross-cultural equivalence and responsiveness, with particular attention to translating and validating instruments across diverse linguistic and cultural groups. The development of consensus-based standard measures of emotional schema would provide essential reference points for evaluating criterion validity. Additionally, longitudinal studies incorporating repeated assessments are needed to determine whether these instruments can reliably detect meaningful changes over time. Collectively, these efforts would enhance confidence in the accuracy, stability and generalisability of emotional schema instruments, thereby supporting their use in both research and applied clinical settings.

### *Strengths and Limitations*

A key strength of this review is the use of COSMIN, a robust and widely recognised framework, which provided a systematic and transparent method for evaluating the methodological and psychometric quality of the included tools. Rather than simply summarising findings, using this framework makes it

more straightforward to distinguish the strengths and weaknesses of the evidence base, and identify gaps in the current literature. A further strength of the current review lay in utilising PRISMA guidelines. This ensured that the review followed clear extraction guidelines, according to the inclusion and exclusion criteria, thereby making the process transparent and replicable. Additionally, thorough search across five databases, enabled the identification of a wide range of instruments assessing emotional schema. Searching multiple databases increases the likelihood of identifying all relevant studies, and reduces potential selection bias.

Nonetheless, there were some limitations in the review. First, only studies published in English were included. While commonly thought to exclude relevant evidence, empirical research suggests that restricting reviews to English-language publications does not necessarily bias findings or substantially affect the results of systematic reviews (Morrison et al., 2012). This restriction may still limit the direct evaluation of translated measures, however, and their cross-cultural applicability. Secondly, we understand that the value of this review is largely dependent on the quality of the primary studies. For example, none of the tools included assessed measurement error, cultural validity or responsiveness, and we are therefore unable to generate assumptions with these findings. Unfortunately, this may limit the scope of the review.

#### *Implications for Future Research*

This review highlights priorities for advancing the measurement of emotional schema and related constructs. First, there is a clear need for systematic assessment of cross-cultural validity. Although some instruments (e.g., PERS, DERS, EBQ) have been translated into multiple languages, few studies have undertaken formal multi-group CFA to evaluate measurement invariance across cultural or linguistic groups. Future research should explicitly test whether these measures reflect invariance across diverse populations, ensuring that scores are directly comparable and not biased by cultural context.

Second, the lack of robust criterion validity highlights the absence of a universally accepted benchmark for emotional schema. Efforts should be directed toward establishing reference standards or composite validation frameworks against which new and existing instruments can be evaluated. This may involve the development of cross-discipline consensus on what constitutes a 'gold standard' for the assessment of emotional schema.

Third, greater emphasis should be placed on the integration of qualitative methods in measure development and validation. Incorporating stakeholder perspectives, including those of clinicians, researchers, and individuals with lived experience, could enhance both the face validity and ecological validity of instruments, ensuring that measures reflect real-world emotional experiences.

Finally, researchers should prioritise longitudinal and predictive validity studies, evaluating not only whether measures align with theoretical constructs, but also whether they can predict meaningful outcomes such as psychological adjustment, treatment response, or resilience.

By addressing these gaps, future work can strengthen the methodological and psychometric foundations of emotional schema measurement tools, thereby improving their applicability across settings, populations, and cultures.

## References

- Becerra, R., Preece, D., Campitelli, G., & Scott-Pillow, G. (2019). The assessment of emotional reactivity across negative and positive emotions: Development and validation of the Perth Emotional Reactivity Scale (PERS). *Assessment, 26*(5), 867-879.
- Becerra, R., Preece, D. A., & Gross, J. J. (2020). Assessing beliefs about emotions: Development and validation of the Emotion Beliefs Questionnaire. *PLoS one, 15*(4), e0231395.
- Brody, L. R., & Hall, J. A. (2008). *Gender and emotion in context*. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3<sup>rd</sup> ed., pp. 395-408). New York: Guildford Press.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2<sup>nd</sup> ed.) Guildford Press.
- Edwards, E. R., & Wupperman, P. (2019). Research on emotional schemas: A review of findings and challenges. *Clinical Psychologist, 23*(1), 3-14.
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of psychopathology and behavioral assessment, 26*(1), 41-54.

- Greenberg, D. M., Rudenstine, S., Alaluf, R., & Jurist, E. L. (2021). Development and validation of the Brief-Mentalized Affectivity Scale: Evidence from cross-sectional online data and an urban community-based mental health clinic. *Journal of Clinical Psychology, 77*(11), 2638-2652.
- Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological inquiry, 26*(1), 1-26.
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*(3), 271-299.
- Halle, T. G., & Darling-Churchill, K. E. (2016). Review of measures of social and emotional development. *Journal of applied developmental psychology, 45*, 8-18.
- Kasper, L. A., Hauschild, S., Berning, A., Holl, J., & Taubner, S. (2024). Development and validation of the Mentalizing Emotions Questionnaire: A self-report measure for mentalizing emotions of the self and other. *Plos One, 19*(5), e0300984.
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modelling* (4<sup>th</sup> ed.) Guildford Press.
- Larionow, P., & Mudło-Głagolska, K. (2022). Assessment of activation, intensity and duration of positive and negative emotions: psychometric properties of the Polish version of the Perth Emotional Reactivity Scale-Short Form. *Current issues in personality psychology, 12*(1), 60.
- Leahy, R. L. (2015). *Emotional Schema Therapy*. Guildford Press: New York.
- Leahy, R. L. (2002). A model of emotional schemas. *Cognitive and behavioral practice, 9*(3), 177-190.
- Manser, R., Cooper, M., & Trefusis, J. (2012). Beliefs about emotions as a metacognitive construct: Initial development of a self-report questionnaire measure and preliminary investigation in relation to emotion regulation. *Clinical Psychology & Psychotherapy, 19*(3), 235-246.
- Mitmansgruber, H., Beck, T. N., Höfer, S., & Schüßler, G. (2009). When you don't like what you feel: Experiential avoidance, mindfulness and meta-emotion in emotion regulation. *Personality and individual differences, 46*(4), 448-453.
- Mokkink, L. B., Elsmann, E. B., & Terwee, C. B. (2024). COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Quality of Life Research, 33*(11), 2929-2939.

- Mokkink, L. B., De Vet, H. C., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of life research, 27*(5), 1171-1179.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & De Vet, H. C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of life research, 19*(4), 539-549.
- Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., ... & De Vet, H. C. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC medical research methodology, 10*(1), 22.
- Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., ... Rabb, D. (2012). The effect of English-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. *International Journal of Technology Assessment in Health Care, 28*(2), 138-144.
- Myers, S. G., Solem, S., & Wells, A. (2019). The Metacognitions Questionnaire and its derivatives in children and adolescents: A systematic review of psychometric properties. *Frontiers in psychology, 10*, 1871.
- Nolen-Hoeksema, S. (2012). Emotion regulation and psychopathology: The role of gender. *Annual Review of Clinical Psychology, 8*, 161-187.
- Nook, E. C., & Somerville, L. H. (2019). Emotion concept development from childhood to adulthood. In *Emotion in the Mind and Body* (pp. 11-41). Cham: Springer International Publishing.
- Padgett, D. K. (2016). *Qualitative methods in social work research* (3rd ed.). SAGE Publications.
- Page, M. J. (2021). E McKenzie J., Bossuyt PM, Boutron I., Hoffmann TC, Mulrow CD, Shamseer L., Tetzlaff JM, Moher D. Updating guidance for reporting systematic reviews: Development of the PRISMA 2020 statement. *J. Clin. Epidemiol, 134*, 103-112.

- Peter, S., Oliver, B. R., Kabo, H. R., Raynaud, A. V., Wiggers, M., & Somerville, M. P. (2025). Measuring emotion beliefs: a systematic review. *Cognition and Emotion*, 1-31.
- Preece, D. A., Becerra, R., Robinson, K., Dandy, J., & Allan, A. (2018). Measuring emotion regulation ability across negative and positive emotions: The Perth Emotion Regulation Competency Inventory (PERCI). *Personality and Individual Differences*, 135, 229-241.
- Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of life research*, 27(5), 1147-1157.
- Rimes, K. A., & Chalder, T. (2010). The Beliefs about Emotions Scale: validity, reliability and sensitivity to change. *Journal of psychosomatic research*, 68(3), 285-292.
- Rosellini, A. J., & Brown, T. A. (2021). Developing and validating clinical questionnaires. *Annual Review of Clinical Psychology*, 17, 55-81.
- Rowan, M., & Wulff, D. (2007). Using qualitative methods to inform scale development. *The Counseling Psychologist*, 35(6), 851-879.
- Suh, J. W., Lee, H. J., Yoo, N., Min, H., Seo, D. G., & Choi, K. H. (2019). A brief version of the Leahy emotional schema scale: A validation study. *International Journal of Cognitive Therapy*, 12(1), 38-54.
- Terwee, C. B., Prinsen, C. A., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., ... & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of life research*, 27(5), 1159-1170.
- Veilleux, J. C., Salomaa, A. C., Shaver, J. A., Zielinski, M. J., & Pollert, G. A. (2015). Multidimensional assessment of beliefs about emotion: Development and validation of the emotion and regulation beliefs scale. *Assessment*, 22(1), 86-100.
- Wells, A. (2002). *Emotional disorders and metacognition: Innovative cognitive therapy*. John Wiley & Sons.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806-838.

## **Chapter 2: Major Research Project**

### **Development of a New Measure of Emotional Schemas**

Prepared in Accordance with Submission Requirements for Prepared in accordance  
with Frontiers in Psychology:

<https://www.frontiersin.org/guidelines/author-guidelines>

## Plain Language Summary

### *Background:*

Researchers believe that the way people understand and interpret their emotions (their 'emotional schemas') can strongly influence mental health. While existing questionnaires measure certain aspects of how people think about emotions, none fully capture which thought patterns matter most in turning negative feelings into mental health problems. Developing a new, more wide-ranging self-report tool could improve our understanding of emotional schemas, how they affect well-being, and how mental health problems develop and continue.

### *Aim of the study:*

To create a new measure of emotional schemas, which it is hoped will help understanding of how negative emotions develop into mental health problems.

### *What we did:*

*Stage 1:* The research team created a list of questionnaire items, which were based on existing questionnaires, research and clinical knowledge.

*Stage 2:* Key researchers in the area, clinicians, and people with experience of mental health problems, were invited to review and refine questionnaire items.

*Stage 3:* A wide and varied sample of people (232 participants) completed the questionnaire items online, in order to establish next steps for its development.

### *What we found:*

The study found three key factors, important in the evaluation of emotions (*Distress Intolerance, Fear and Stigma, Awareness and Reflection*). Our findings have formed the basis for future validation of the questionnaire, and will hopefully result in a reliable instrument which will be suitable for use in both clinical and research settings.

## Abstract

*Background:* Cognitive models propose that the way people understand and interpret their emotions (schemas) can play a fundamental role in the development and maintenance of mental health problems. Measurement tools such as The Leahy Emotional Schema Scale (LESS), Beliefs about Emotions Scale (BAES) and the Metacognitions Questionnaire (MCQ) were developed to determine how people conceptualise and manage their feelings. However, none of these instruments provide a comprehensive and reliable assessment of which emotion-related cognitions are most influential across the continuum from negative emotional experiences to the development of psychological difficulties.

*Aim:* To develop a new self-report measure of emotional schemas that overcomes the limitations of existing measures.

*Methods:* An item pool was generated and refined from existing literature and measures, and through expert consultation with people with lived experience of mental health problems, clinicians, and researchers. A three-phase cross-sectional design was employed to develop a self-report measurement tool of negative emotional schema. The Emotional Schema Self-Report of Negative Emotions (ESSeNcE) was completed online by 232 participants aged between 18 and 79 years ( $M = 44.51$ ;  $SD = 13.62$ ), including 191 females, 39 males and 2 non-binary; 170 people reported having lived experience of a mental health condition. A series of exploratory factor analyses with parallel analyses was used to determine the best-fitting latent factor structure.

*Results:* A robust three-factor solution comprising *Distress Intolerance, Fear and Stigma, Awareness and Reflection*, provided the best model fit accounting for 51.1% of the variance. Reliability indices were strong (Cronbach's alphas ranged from .78 to .95), supporting the instrument's internal consistency.

*Conclusion:* This work provides a foundational framework for future validation studies, with the goal of developing a psychometrically sound instrument, suitable for application in both clinical and research settings.

**Key words:** schemas; metacognition; emotions; cognitive models

## Introduction

A large body of research has demonstrated that the way in which we conceptualise and appraise our emotions is vital to mental wellbeing. Hayes (2004) proposed that there have been three dominant 'waves' in the development of psychological approaches to treating mood disorders, each model based on the previous one. The first wave, grounded in the methods of Skinner and Watson, focused on observation and modification of behaviour as a way of improving mental health (Skinner, 1953; Watson, 1925). Second wave therapies, exemplified by Beck's cognitive model, are aimed at helping individuals identify maladaptive behaviours and cognitive distortions, and then develop strategies to help change them, which will then hopefully lead to mood improvement. In contrast, third wave approaches, while retaining an emphasis on behaviour and cognition, differ from earlier models in their primary therapeutic focus. Rather than seeking symptom reduction or mood improvement through the direct modification of cognitive content, third wave therapies emphasise changing unhelpful ways of relating to internal experiences, such as thoughts and emotions. Central to these approaches is the cultivation of psychological flexibility, defined as the capacity to remain in contact with the present moment and to engage in behaviour that is guided by personal values, despite the presence of difficult internal experiences (Hayes et al., 1999). Accordingly, interventions such as Mindfulness-based Cognitive Therapy (MBCT) and Acceptance and Commitment Therapy (ACT) promote processes of awareness, acceptance and cognitive defusion, with the aim of enhancing adaptive functioning and valued living (Segal et al., 2002; Brown, 2011).

The notion that an individual's distress stems from the way in which they think about and respond to their thoughts and emotions is a core principle of Metacognitive Theory (Wells, 2009). This differs from other third wave models in that it emphasises the central function of an individual's interpretation of their emotions, rather than just acceptance and mindful awareness (Leahy, 2019). During his doctoral studies, Wells became interested in concepts such as excessive self-focus and the processing of information, and their role in anxiety disorders (Wells, 1987). He emphasised that cognitive processes, for example rumination and thoughts based around one's mental state, should be fundamental to the approach and treatment of psychological disorders (Capobianco & Nordahl, 2021). Emotion-Focused Therapy (Greenberg et al., 1993) also highlights the key function of emotional processing, as well as the meaning that we derive from our emotions. Mood improvement then occurs through processes such as awareness, reflection and effective emotion regulation (Greenberg, 2010).

Another third wave model targeting the emotional experience is the Emotional Schema Model, which describes schemas as plans, concepts, and strategies used when an unpleasant emotion is activated (Leahy, 2002). This meta-experiential framework suggests that people differ in how they understand and evaluate their emotions, and these interpretations are collectively known as 'emotional schemas'. Leahy (2015) proposes that there are individual variations in how we conceptualise emotions, which in turn drive the emotion regulation strategies we use. For example, individuals who hold negative perceptions of emotions, including the idea that certain emotional experiences are permanent and inescapable, or distress being a sign of weakness, are more likely to engage in problematic ways of coping, for example avoidance, catastrophising, binge eating, rumination, worry, or blaming of self or others (Leahy, 2016). This could then lead to adverse consequences for mental health, such as a chronic affective disorder, and may even be associated with poorer prognosis or treatment outcome (Rimes & Chalder, 2010).

Emotional schemas - individuals' beliefs and interpretations about their own emotions - play a central role in emotion regulation and mental health. They influence how emotions are experienced, evaluated, and managed, and maladaptive schemas have been implicated in a range of psychological difficulties, including anxiety, depression, and chronic stress (Veilleux et al., 2015). Consequently, self-report instruments have been developed to capture these patterns, with several widely used measures providing preliminary psychometric support.

The Leahy Emotional Schema Scale (LESS; Leahy, 2002) was designed to assess how individuals conceptualise their emotions and the coping strategies they employ when uncomfortable emotions are activated. Comprising 50 items across 14 theoretically driven dimensions, the LESS provides insight into beliefs related to blame, guilt, and acceptance. While this scale has demonstrated some reliability, it does not fully capture the continuum of emotional experiences from initial negative affect to potential psychological disorder, and experiencing an uncomfortable emotion does not necessarily indicate long-term mental health difficulties.

Other measures of emotional schemas and related constructs have attempted to address these gaps. The Emotion Beliefs Questionnaire (EBQ; Becerra et al., 2020) and the Beliefs about Emotions Questionnaire (BAEQ; Manser, Cooper, & Trefusis, 2012) focus on the perceived controllability and utility of emotions, highlighting how beliefs shape responses to emotional experiences. The Beliefs about Emotions Scale (BES; Rimes & Chalder, 2010) and the Emotion and Regulation Beliefs Scale

(ERBS; Mitmansgruber et al., 2009) extend this work by examining maladaptive appraisals and their association with emotion regulation difficulties. The Meta-Emotion Scale (MES; Gottman et al., 1996) additionally considers meta-emotional awareness and parental emotion socialisation, demonstrating the broader interpersonal context of emotional schemas.

Although these instruments provide valuable insights, collectively they still leave gaps in capturing the full range of emotional experiences, particularly the progression from everyday emotional responses to clinically significant difficulties. This highlights the need for the development of a more comprehensive measure that can reliably assess emotional schemas across diverse experiences and populations, forming the rationale for the present study.

Nonetheless, a valid and comprehensive measure of how individuals make sense of their emotions and the coping strategies they employ, would further existing knowledge around the development and maintenance of psychopathology, as well as aid understanding of emotions, emotion-regulation and well-being more broadly.

### *Limitations of Existing Measures*

A significant amount of research has attempted to determine the cognitive processes involved once a negative emotion has been triggered. 'Metacognition' per se is generally defined as awareness of one's own thought processes, thus making these concepts difficult to observe and even more difficult to measure (Craig et al., 2020), and pointing to the importance of self-report. Generally speaking, existing questionnaire measures draw on the following concepts: a) individuals' interpretation of their emotions or b) the regulation strategies employed to minimise emotional impact. However these tools are limited in a variety of important ways, each of which suggests the development of an alternative.

#### *i. Sample characteristics*

It is recommended that clinical questionnaires are developed and validated using samples of individuals possessing key attributes and experience of the population of interest (Rossellini & Brown, 2021). One of the major shortcomings in the development of existing measures of emotional schemas, however, is that samples are drawn from a non-representative population (e.g. undergraduate students) and therefore lacking in external validity. When samples are taken in this way from a limited source, sharing

common characteristics (e.g. age, race, socioeconomic status, and education), even large samples will not mitigate the effects of systematic variance (Worthington & Whittaker, 2006). Samples that are not representative of the population of interest may also affect the stability of the factor structure and generalisability of the resulting tool (Worthington & Whittaker, 2006). By recruiting a large, heterogeneous sample, the current research aims to develop a measure which is broadly applicable as well as clinically useful.

*ii. Lack of qualitative data*

Qualitative information collected in the process of scale development provides key information from participants and enhances the quality of research (Rowan & Wulff, 2007). Many of the current measures reviewed, however, have been developed utilising purely quantitative methods. Use of a statistical program alone will not produce an inclusive pool of dimensions or factors that comprise a valid and reliable measurement tool (Worthington & Whittaker, 2006); therefore, within the scale development process, both quantitative and qualitative data are key. Qualitative information informs the initial item pool development, whilst quantitative data assists in the refinement and piloting stages. In this mixed methods approach, qualitative information, based on empirical and clinical knowledge, as well as personal accounts and lived experience, improves the value of quantitative data, ensuring that it is substantiated from an ecological perspective (Padgett, 2016). This process informs and hones item generation and helps to ensure relevance to the construct(s).

*iii. Lack of comprehensiveness*

Another aim of this project is to address the limited dimensionality of existing tools, which do not fully capture the range of ways which individuals experience and appraise their negative emotions, and how these processes relate to psychopathology. While many existing measures focus on individuals' awareness of emotions or their ability to identify and label them, they provide limited insight into how people evaluate, interpret, and respond to these experiences. Recognising and distinguishing emotions is only the first stage of emotion processing; once a negative emotion is identified, individuals differ in how they appraise, interpret, and cope with it (Leahy, 2007). Although instruments such as the LESS, EBQ, BAEQ, BES, ERBS, and MES have advanced our understanding of emotional schemas and related beliefs, they often focus on a limited set of dimensions, populations, or contexts, and do not fully capture the progression from emotional experience to adaptive or maladaptive regulation strategies.

Developing a new measure provides the opportunity to address these gaps by more comprehensively assessing how individuals understand, evaluate, and respond to their emotions. By integrating a broader range of emotional experiences and regulatory strategies, the new instrument aims to clarify why some people develop psychological difficulties in response to negative emotions while others do not. This approach not only builds on the strengths of existing tools but also extends them, offering a more nuanced and clinically relevant understanding of emotional schemas and their role in mental health outcomes.

*iv. Limited sample sizes*

Many existing measures of emotional schemas were developed using relatively small samples, often fewer than 100 participants. While rules of thumb for sample size in scale development vary, at least 150 participants are generally recommended for exploratory factor analysis (EFA) to ensure stable and replicable factor structures (Worthington & Whittaker, 2006; Velicer & Fava, 1998). During the exploratory phase, when the underlying factor structure is not yet known, larger samples of 200-300 participants are typically advised to provide sufficient statistical power for meaningful analysis (Clark & Watson, 2019).

A further limitation is that there is currently no consensus on the factor structure of emotional schemas, with different measures proposing varying dimensions and subscales. Many instruments, such as the LESS, EBQ, and BES, capture certain beliefs about emotions but do not comprehensively evaluate the full range of appraisal, coping, and regulatory processes that may influence why some individuals develop psychological difficulties while others do not. As highlighted by Siddaway et al. (2019) in the context of self-injurious cognitions, developing a new measure provides the opportunity to systematically characterise the construct, identify a coherent factor structure, and ensure that the resulting scale is both comprehensive and psychometrically robust.

By recruiting a larger and more heterogeneous sample, the present study aims to address these gaps, providing a scale that captures the breadth of emotional schema dimensions and supports reliable, generalisable assessment across diverse populations.

## **Aims and Objectives of the Current Research**

The issues outlined above suggest that current tools may not comprehensively or reliably measure emotional schema. In response to these gaps, the present study is designed with three primary aims. First, it seeks to generate an initial pool of questionnaire items that adequately represent the breadth and facets of emotional schemas. Second, through a two-phase refinement process, it aims to evaluate the inclusiveness and clarity of these items, thereby ensuring that the resulting instrument is both comprehensive and conceptually coherent. Finally, the study aims to produce a tool comprising the most robust items, providing a foundation for future psychometric evaluation and application in clinical and/or research settings.

It is hoped that our measure will differ from Leahy's in several important ways. Whereas the LESS was developed primarily within a clinical context and conceptualises emotions broadly, the current measure focuses specifically on negative emotional experiences in a dimensional framework. Rather than restricting assessment to pathological manifestations, the proposed instrument is intended to capture the full spectrum of negative emotional schema evaluation, from normative, everyday appraisals to more maladaptive interpretations, within the general population. This approach reflects the view that psychological difficulties are present along a continuum and can be identified outside clinical settings, supporting the need for measures suitable for early identification and prevention research. In addition, the measure aims to extend existing emotion schema frameworks by incorporating physiological aspects of the emotional experience, acknowledging the somatic nature of emotions and role of beliefs about bodily responses in emotional distress. Finally, by including both maladaptive and adaptive dimensions, such as resilience and positive appraisals of negative emotions, while avoiding the assessment of specific emotional regulation strategies, the measure aims to capture more fundamental emotion-related beliefs with broader applicability than the LESS.

## **Methods**

### *Design*

A three-phase, cross-sectional design was employed to develop a self-report measurement tool of negative emotional schema (ESSeNcE). The revised research proposal is available via the link in Appendix 2.1.

## *Ethics*

Ethical approval for the study was obtained from the College of Medical, Veterinary, and Life Sciences Ethics Committee at the University of Glasgow (reference: 200230293) in April 2024 (see Appendix 2.2). Participants in Phases 2 and 3 received detailed study information and privacy notices, and were given the opportunity to ask questions before providing informed consent. They were reminded of their right to withdraw from the study at any time. All data were managed in accordance with GDPR (2018) regulations. Personal information was stored electronically in encrypted, password-protected files on secure University of Glasgow computers, accessible only to the primary researcher (CL) and supervisors (AS and HM). For analysis, all data were anonymised to maintain confidentiality. To reduce any potential distress among participants, information about available support services was provided, and participant wellbeing and safety were reviewed throughout the course of the study.

## *Scale Development*

Phase 1 involved the research team (CL, AS and HM) developing an initial item pool, informed by a literature review, existing emotional schema measures, and items adapted from tools assessing related constructs such as evaluation, assessment or appraisal of emotions. Phase 2 engaged key researchers, clinicians and academics specialising in emotions and metacognition, as well as people with lived experience of mental health problems (PWLE), to review and provide feedback on the item pool. Clinician/academics were purposively sampled from the UK, wider Europe, the USA and Australia, and defined a priori as individuals with a track record of peer-reviewed publications and/or substantial academic experience in the field. Potential contributors were identified via literature searches and approached individually via email with an invitation to review the preliminary item pool and provide structured feedback. This approach was intended to maximise content validity through input from internationally recognised specialists. Phase 3 involved recruiting a heterogeneous sample online who completed the refined item pool. Further details for each phase are provided below.

### *Phase 1: Item Generation*

A broad range of existing theory, clinical experience and knowledge was drawn upon by the team to generate a preliminary item pool. Potential items were considered from an extensive review of existing tools measuring emotional schema and appraisal of emotions in a variety of populations. This exercise produced a preliminary item pool ( $n = 235$ ), which was grouped into themes.

The list then underwent a thorough iteration procedure, whereby each item was discussed individually by the research team to ensure its relevance, clarity and alignment to the research objectives. During this stage of scale development, it is recommended to examine related constructs to make sure that the items are meaningful, relevant, and cover the intended aspects of the construct (Rossellini & Brown, 2021). There were also some items which related to feelings about thoughts, rather than emotions, which were also removed, to ensure content validity. The overall flow and structure of the questionnaire was also taken into consideration and items were refined or eliminated based on this exercise to produce a list of items ( $n = 119$ ) for the next phase of questionnaire development.

### *Phase 2: Item Refinement*

Experienced emotion and metacognition researchers, clinician/academics, and PWLE of mental health problems were asked to review the item pool. Feedback was requested on relevance (i.e. if the item was appropriate to the research question), inclusiveness (i.e. whether key concepts were missing), comprehensibility (i.e. clarity and phrasing of items and response options) and ambiguity (i.e. could the question be interpreted in multiple ways). The participants were approached as follows:

- Thirty-one clinicians and researchers, based internationally, were approached individually, via email, to ask if they would like to participate. It was felt by the research team that this purposive sampling approach, where individuals were asked directly if they would be willing to take part, might increase the likelihood of their involvement, as it felt more personal and engaging compared to mass recruitment methods. Also, this kind of targeted recruitment strategy ensured focus on the appropriate demographic, thereby increasing the relevance of the responses. The main disadvantage, however, was that this strategy was quite time-intensive, i.e. researching the appropriate audience, finding their contact details and sending out individual emails.
- Individual approaches to clinicians and researchers were combined with email outreach to seven private therapy organisations. It was felt that this would be a good way of expanding the participant pool, reaching a broader audience while maintaining a targeted focus. It was also a more efficient way of contacting a larger number of potential participants.
- PWLE of mental health problems were approached via a social media advert on Facebook.

Inclusion criteria for all participants were that they needed to be 16 years or older; could access the internet, and were able to speak and understand English.

All participants were invited to access the survey via a Qualtrics link. They were first directed to the Participant Information Sheet (PIS) with details of the study, to assist them in deciding if they wished to take part. Those who opted to participate were directed to view a privacy notice and then provide informed consent, some demographic information, e.g., their age in years, ethnicity, sex and employment status. Following this, participants reviewed the questionnaire items and were invited to provide their email address if they wished to be sent a copy of the results once the study was completed. For copies of the PIS, consent form, privacy notice, demographic information sheet, Social Media advert and email sent to clinicians and researchers, please see Appendix 2.3. Prior to sending out this link, a pilot was conducted amongst the research team as well as academic peers of the lead researcher, in order to assess end-to-end functionality of the Qualtrics link, as well as the survey's usability, logical flow and integration of mood questionnaires.

Out of the above approaches, eight clinicians and researchers, as well as eight PWLE of mental health issues responded with their feedback. In addition to online feedback, Professor Andrew Ortony, recognised globally for his empirical contribution to the field of emotions and metacognition, gave generously of his time in order to provide comprehensive input to this phase of questionnaire development. Professor Ortony has been influential in the field of emotion theory and appraisal research. One of his seminal and most highly cited works, *The Cognitive Structure of Emotions*, systematically explains how cognitive appraisals underlie different emotions and their intensities (Ortony, Glore & Collins, 2022).

This empirical grounding ensured that the list of items captured the lived experiences and perspectives of the population being studied (Creswell, 2013), and based on this expert feedback, the following revisions were made:

- Rewording or simplifying complex or ambiguous questions
- Eliminating redundant/repetitive items
- Adding additional items in accordance with feedback
- Plain English screening

Based on this feedback, items were refined through rewording, additions and deletions, resulting in a final pool of 147 items, which was then sent out for completion by a heterogeneous sample. (See Appendix 2.4).

### *Phase 3: Recruitment*

A heterogeneous sample of participants was recruited online from a diverse range of sources, with the aim of identifying items that were broadly applicable and generalisable across different populations. Recruitment was conducted via social media platforms, including Facebook, Instagram, LinkedIn, and Reddit, as well as through the research team's personal and professional networks. Study advertisements provided a brief overview of the study, eligibility criteria, and a link to the survey. Following best practice for scale development (Siddaway et al., 2019), recruiting a diverse sample enhances the generalisability of findings, ensures that the full range of the construct is represented, and provides sufficient variability to identify meaningful factor structures during exploratory analyses. For copies of the consent form, demographic information sheet, and social media advertisements, please see Appendix 2.5. In addition to the questionnaire items, participants were also asked to complete the Patient Health Questionnaire-9 (PHQ-9, Kroenke et al., 2002), Generalized Anxiety Disorder-7 (GAD-7, Spitzer et al., 2006) and the Beliefs about Emotions Scale (BAES, Rimes & Chalder, 2010). See Appendix 2.7.

### *Data Analysis*

Data were collected from 232 participants between 11<sup>th</sup> June and 3<sup>rd</sup> September 2025. In this 3<sup>rd</sup> phase, 715 initially completed the ESSeNcE survey, however 483 (67.5%) were excluded due to incomplete responses, leaving a final sample of 232. This discrepancy may reflect partial survey engagement, whereby participants accessed the link, but did not progress through all items. Such attrition is common in online data collection and may be attributable to factors including survey length, participant burden, loss of interest or disengagement during completion. Data were analysed using IBM SPSS Statistics (version 29.0.2.0), with list-wise deletion applied to missing data. Descriptive analyses were first conducted to summarise participants' demographic characteristics and psychometric variables. To examine the underlying structure of the scale and reduce the item pool, a series of exploratory factor analyses (EFA) was conducted. EFA is a data-driven technique used to identify latent factors that

account for the covariation among items, particularly useful when the factor structure of a construct is not yet established (Fabrigar et al., 1999; Worthington & Whittaker, 2006).

Parallel analysis (PA; Velicer, Eaton & Fava, 2000; Zwick & Velicer, 1986) was used to inform the number of factors in the model fitted to the data. PA compares the eigenvalues derived from the actual data with those obtained from randomly generated datasets of the same size, providing an empirical criterion for determining the number of factors to retain (O'Connor, 2000). Factors are retained when their eigenvalues exceed the corresponding values from the random data, reducing the risk of over- or under-extraction that can occur with traditional methods such as the Kaiser Criterion or scree plot inspection.

The combination of EFA and PA allows for both the identification of a conceptually meaningful factor structure and the reduction of items that do not contribute to coherent dimensions of the construct. This approach follows best practice in scale development, as exemplified by Siddaway and colleagues (2019), and ensures that the resulting scale is psychometrically robust and interpretable.

Subscales were constructed from the final factors, with reliability assessed using internal consistency and split-half methods. Validity was evaluated through convergent and discriminant analyses. Descriptive analyses of the final subscales were conducted to inform interpretation and scoring. See Appendix 2.6 for detailed data analysis plan.

## **Results**

In Phase 2 of the study, the largest proportion of participants (25%) were aged between 25 and 34 years, and the substantial majority were female (87.5%). Of the 24 participants, 47.8% identified as people with lived experience (PWLE) of mental health conditions, while 43.4% were either clinicians experienced in working with individuals with mental health conditions or researchers in the mental health field. A summary of these demographic characteristics is presented in Table 2.1.

**Table 2.1 - Demographic Characteristics of Participants in Phase 2**

<b>Variable</b>	<b>Category</b>	<b>n</b>	<b>%</b>
<b>Age</b>	Under 18	0	0
	18-24	5	20.8
	25-34	6	25.0
	35-44	3	12.5
	45-54	5	20.8
	55-64	4	16.7
	Prefer not to say	1	4.2
<b>Gender</b>	Female	21	87.5
	Male	1	4.2
	Non-binary	0	0
	Prefer not to say	2	8.3
<b>Occupational status</b>	Full-time	15	62.5
	Part-time	2	8.3
	Prefer not to say	7	29.2
<b>Identity</b>	Clinician	5	21.7
	Researcher	5	21.7
	PWLE	11	47.8
	Prefer not to say	3	8.8

In Phase 3, participants had a mean age of 44.51 years (SD = 13.62; range = 18-79 years). Of the 232 participants, 82.3% identified as female and less than 1% as non-binary. Nearly half of the sample (48.3%) was married, and 73.3% reported having lived experience of mental health issues (PWLE). The vast majority (approximately 90%) had completed some form of post-secondary education, ranging from a college diploma to a doctoral degree, while 9.9% reported having no formal qualifications. A summary of these demographic characteristics is presented in Table 2.2.

**Table 2.2 - Demographic Characteristics of Participants in Phase 3**

Variable	Category	n	%
<b>Gender</b>	Female	191	82.3
	Male	39	16.8
	Non-binary	2	0.9
<b>Relationship status</b>	Married	112	48.3
	Cohabiting	33	14.2
	In a relationship	33	14.2
	Divorced	6	2.6
	Single	48	20.7
<b>Education level</b>	None	23	9.9
	College diploma	38	16.4
	Vocational/technical	19	8.2
	Undergraduate degree	74	31.9
	Postgraduate degree	50	21.6
	PhD/Doctorate	25	10.8
<b>Identity</b>	PWLE* - Yes	170	73.3
	PWLE - No	57	24.6
	Prefer not to say	5	2.2

[Mean age 44.51 years (SD = 13.62)]

After providing informed consent, participants completed a demographic questionnaire and baseline mood measures to capture key characteristics and emotional states. Following this, they were invited to complete the full item pool, which consisted of 147 questionnaire items designed to comprehensively assess emotional schemas. The survey was structured to allow participants to provide qualitative feedback, with ample space for comments on item clarity, relevance and wording, enabling the research team to identify items that were confusing, redundant, or difficult to interpret. This feedback was used to refine the item pool and ensure that the scale was both conceptually coherent and accessible to a diverse sample.

Descriptive statistics for psychological measures are presented separately for participants with (*PWLE-Yes*) and without (*PWLE-No*) lived experience of mental health problems (see Table 2.3 below). For depressive symptoms, as measured by the PHQ-9, participants in the *PWLE-Yes* group scored between 9 and 33 (Range = 24), with a mean score of 17.46 (SD = 5.98). In contrast, participants in the *PWLE-No* group scored between 9 and 25 (Range = 16), with a lower mean score of 12.35 (SD = 3.78). Anxiety symptoms, assessed using the GAD-7, showed a similar pattern. The *PWLE-Yes* group had scores ranging from 7 to 28 (Range = 21), with a mean of 13.91 (SD = 5.36), whereas the *PWLE-No* group

scored between 7 and 19 (Range = 12), with a mean score of 9.68 (SD = 2.92). For beliefs about emotions, measured by the BAES, participants in the *PWLE-Yes* group scored between 12 and 58 (Range = 46), with a mean score of 35.69 (SD = 11.36). Participants in the *PWLE-No* group showed scores ranging from 12 to 53 (Range = 41), with a mean of 31.56 (SD = 9.85).

**Table 2.3** - Comparison of descriptive statistics for PHQ-9, GAD-7 and BAES, scores by self-reported mental health experience

Measure	Group	Range	Min	Max	Mean	SD
PHQ-9	<i>PWLE-Yes</i>	24.00	9.00	33.00	17.46	5.98
	<i>PWLE-No</i>	16.00	9.00	25.00	12.35	3.78
GAD-7	<i>PWLE-Yes</i>	21.00	7.00	28.00	13.91	5.36
	<i>PWLE-No</i>	12.00	7.00	19.00	9.68	2.92
BAES	<i>PWLE-Yes</i>	46.00	12.00	58.00	35.69	11.36
	<i>PWLE-No</i>	41.00	12.00	53.00	31.56	9.85

*Parallel and Exploratory Factor Analyses:*

A parallel analysis (PA) was first conducted using O’Connor’s (2000) syntax to determine the optimal number of factors to retain. PA compares the eigenvalues from the actual data with those obtained from randomly generated datasets, providing an empirical criterion for factor retention. The initial PA suggested that up to eight factors had eigenvalues exceeding the 95<sup>th</sup> percentile of randomly generated values, indicating the potential presence of multiple latent dimensions.

Following this, an exploratory factor analysis (EFA) was conducted on the 147-item pool using Maximum Likelihood extraction with Promax rotation to examine the underlying factor structure. The EFA was performed iteratively across five rounds. At each round, items were removed if they demonstrated low communalities (< .20), weak primary loadings (< .40), or substantial cross-loadings (> .30 on multiple factors), in order to refine the factor structure and retain conceptually coherent items. Examination of the correlation matrix indicated that some items initially exhibited limited shared variance with other items ( $r < .20$ ; Field, 2024); however, the overall Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy remained high across all iterations (final KMO = .88), exceeding the recommended minimum of .60 (Pallant, 2020). Bartlett’s Test of Sphericity was significant ( $\chi^2(10,731) = 30,094.96, p < .001$ ), confirming that the data were suitable for factor analysis.

Inspection of the scree plot, factor interpretability, and the pattern of item loadings across successive EFAs indicated that a three-factor solution provided the most conceptually coherent and parsimonious representation of the data. The final EFA retained three correlated factors, together accounting for 51.1% of the total variance based on the Extraction Sums of Squared Loadings, which falls within the mid-range of values typically considered acceptable for exploratory factor analyses in psychological research (Hair et al., 2019; Field, 2024; Pallant, 2020). Specifically, Factor 1 accounted for 37.55% of the variance, Factor 2 for 8.56%, and Factor 3 for 4.99%. After Promax rotation, the Rotation Sums of Squared Loadings demonstrated a more balanced distribution of variance across factors: Factor 1 (13.91%), Factor 2 (10.89%), and Factor 3 (2.75%).

Examination of item content and thematic coherence suggested that Factor 1 reflects *Distress Intolerance*, Factor 2 reflects *Fear and Stigma*, and Factor 3 reflects *Awareness and Reflection*. This iterative process, combining PA, EFA, and systematic item removal, ensures a conceptually meaningful and psychometrically robust factor structure, consistent with best practice in scale development (Siddaway et al., 2019).

#### *Scale refinement and subscale construction*

The following criteria were adopted to extract the greatest number of factors that would be well defined and reasonably distinct from one-another. When two items correlated strongly (.75), the item with lowest item total correlation was considered to contain redundant information and was deleted (Clark & Watson, 1995). Items with loadings .40 or which demonstrated reasonably strong loadings (.30) on more than one factor were eliminated to maximise the measurement properties and discriminant validity of individual subscales (Clark & Watson, 1995). See Table 2.4 for a summary of factors, sample items and reliability and Table 2.5 for Promax rotated loadings of items for full ESSeNcE scale.

**Table 2.4 - Summary of ESSeNcE Factors, Sample Items and Reliability**

Factor	Number of Items	Sample Item	Cronbach's Alpha
Factor 1 - <i>Distress Intolerance</i>	24	I allow my negative emotions to dictate my thoughts and behaviours	0.95
Factor 2 - <i>Fear and Stigma</i>	9	I can keep safe if I don't let myself have negative emotions	0.89
Factor 3 - <i>Awareness and Reflection</i>	7	It's important to understand why I experience negative emotions	0.78

**Table 2.5 - Promax-Rotated Loadings for ESSeNcE questionnaire items in Sample 3 (n = 232)**

	Mean (SD)	Factor Loadings		
		Distress Intolerance	Fear & Stigma	Awareness & Reflection
<i>1. Distress Intolerance</i>				
I find it difficult to control my negative emotions	5.93 (1.24)	<b>.62</b>	.05	-.10
I allow my negative emotions to dictate my thoughts and behaviours	5.60 (1.14)	<b>.55</b>	-.03	-.13
It is hard for me to wind down from negative emotions	6.25 (1.21)	<b>.64</b>	.07	-.00
My negative emotions make me feel lost	3.15 (1.52)	<b>.59</b>	.22	.12
My negative emotions are harmful for me	3.07 (1.28)	<b>.53</b>	.20	.01
My negative emotions might damage me psychologically	2.61 (1.37)	<b>.52</b>	.26	.08
My negative emotions make me feel everything is pointless	2.72 (1.44)	<b>.74</b>	.02	-.04
My negative emotions affect my self-image	3.49 (1.36)	<b>.63</b>	.04	.07
My negative emotions cause me a lot of problems in my daily life	2.57 (1.35)	<b>.75</b>	.10	.01
My negative emotions affect my ability to take care of myself	2.60 (1.44)	<b>.70</b>	.01	.02
When I experience negative emotions, I feel sure that everything I touch will go wrong	2.81 (1.41)	<b>.48</b>	.27	-.04
My negative emotions are very unhelpful to me	3.72 (1.69)	<b>.51</b>	.21	-.21
My negative emotions reflect how difficult my life is	3.46 (1.73)	<b>.43</b>	.13	.09
My negative emotions stop me from getting things done	3.84 (1.73)	<b>.76</b>	-.06	.02
My negative emotions regularly take over my attention	3.06 (1.37)	<b>.88</b>	-.16	-.01
When I'm experiencing negative emotions, it's hard for me to imagine feeling any other way	3.09 (1.39)	<b>.72</b>	.06	-.02
When I'm experiencing negative emotions, it's hard for me to think straight	3.50 (1.23)	<b>.80</b>	-.13	0.2
I often feel overwhelmed by my negative emotions	3.18 (1.44)	<b>.85</b>	-.01	.04
I'm not as able to function when I'm experiencing negative emotions	3.21 (1.29)	<b>.81</b>	-.19	-.01
I can't tolerate my negative emotions	2.58 (1.23)	<b>.47</b>	.25	-.04
My negative emotions affect my perception of reality	3.00 (1.34)	<b>.69</b>	-.10	.10
If I think too much about my negative emotions, then I feel worse	3.51 (1.52)	<b>.48</b>	.08	-.02

I often wonder why I have so many negative emotions	2.98 (1.60)	<b>.58</b>	.16	.03
My negative emotions make me feel a sense of emptiness	6.60 (2.76)	<b>.75</b>	-.11	-.01
<i>2. Fear &amp; Stigma</i>				
If I didn't control a negative emotion, and then something bad happened, it would be my fault	5.99 (1.39)	.08	<b>.59</b>	.01
Not being able to control how I express my negative emotions is a sign of weakness	5.81 (1.24)	-.12	<b>.69</b>	.02
I will be punished for not controlling certain negative emotions	5.23 (1.34)	-.14	<b>.73</b>	.18
My negative emotions are bad and need to be controlled	5.39 (1.29)	.24	<b>.52</b>	-.06
I feel ashamed of my negative emotions	2.84 (1.57)	.11	<b>.70</b>	-.04
If others find out about my negative emotions it will ruin my relationships	2.85 (1.59)	.08	<b>.74</b>	.05
I don't want to admit to having negative emotions, but I know that I have them	2.97 (1.59)	-.06	<b>.79</b>	-.04
I think that my negative emotions are strange or weird	2.38 (1.47)	.02	<b>.74</b>	-.07
I can keep safe if I don't let myself have negative emotions	2.43 (1.20)	.04	<b>.51</b>	-.07
<i>3. Awareness and Reflection</i>				
It's important to understand why I experience negative emotions	4.18 (.86)	-.05	.10	<b>.54</b>
Negative emotions can help me to avoid disastrous situations in the future	3.14 (1.16)	-.17	-.00	<b>.46</b>
I often become aware of my negative emotions in different situations	3.79 (1.19)	.14	-.01	<b>.48</b>
I usually spend time thinking about my emotions	3.50 (1.42)	.19	.09	<b>.60</b>
I think my negative emotions and state of mind deserve to be paid attention to	3.89 (1.42)	.09	.12	<b>.59</b>
I am interested in my negative emotions	3.56 (1.39)	-.07	.08	<b>.74</b>
With some distance I can understand my negative emotions in a new way	4.01 (1.32)	-.04	-.10	<b>.66</b>
Eigenvalues		15.02	3.42	1.99
% of Variance		37.55	8.56	4.99

### *Validity*

All 40 items demonstrated acceptable standardised factor loadings, each exceeding the recommended minimum of .40, thereby indicating sufficient explanatory strength for their respective latent dimensions (Hair et al., 2019). Across the three factors, loadings ranged from .43 to .88: *Distress Intolerance* (.43-.88), *Fear and Stigma* (.51-.79), and *Awareness and Reflection* (.46-.74). In evaluating convergent validity, the criterion proposed by Fornell and Larcker (1981), whereby an Average Variance Extracted (AVE) value of .50 or higher is considered indicative of strong convergence among items within a factor

(this was calculated using SPSS). Nonetheless, in exploratory contexts, slightly lower AVE values may still be deemed acceptable, as theoretical and practical considerations often take precedence over rigid numerical thresholds (Costello & Osborne, 2005).

Inter-dimension correlations were examined to assess discriminant validity among the three factors. The correlation between *Distress Intolerance* and *Fear and Stigma* was moderate to strong, and statistically significant ( $r = .705, p < .001$ ), indicating some shared variance between these two constructs. However, the correlations between *Distress Intolerance* and *Awareness and Reflection* ( $r = .109, p = .097$ ) and between *Fear and Stigma* and *Awareness and Reflection* ( $r = -.056, p = .397$ ) were weak and non-significant, suggesting that these constructs are largely independent.

According to Fornell and Larcker (1981), discriminant validity is supported when the correlations between constructs are lower than the square root of their AVE. Additionally, Kline (2011) suggests that inter-construct correlations below 0.85 indicate adequate discriminant validity. As all correlations in this study are below this threshold, the results provide evidence of acceptable discriminant validity, confirming that the three factors represent distinct conceptual dimensions.

Correlations among the three subscales, the PHQ-9, GAD-7, and BAES are presented in Table 2.6 below. Convergent validity was supported for the *Distress Intolerance* and *Fear and Stigma* subscales. The *Distress Intolerance* subscale was strongly positively correlated with the PHQ-9 ( $r = .666, p < .01$ ), GAD-7 ( $r = .656, p < .01$ ), and the BAES ( $r = .484, p < .01$ ), indicating that higher intolerance of distress is associated with greater depressive and anxiety symptoms, as well as more maladaptive beliefs about emotions. Similarly, the *Fear and Stigma* subscale showed strong positive correlations with the PHQ-9 ( $r = .610, p < .01$ ), GAD-7 ( $r = .564, p < .01$ ), and the BAES ( $r = .751, p < .01$ ), consistent with the expectation that fear and stigma regarding emotions are linked to symptoms and maladaptive beliefs.

Divergent validity was observed for the *Awareness and Reflection* subscale. This subscale showed weak and mostly non-significant correlations with PHQ-9 ( $r = -.052, ns$ ), GAD-7 ( $r = .050, ns$ ), and BAES ( $r = -.179, p < .01$ ), suggesting that greater emotional awareness and reflective capacity is largely independent of symptom severity and maladaptive beliefs.

The pattern of correlations between the subscales themselves was also informative. *Distress Intolerance* and *Fear and Stigma* were strongly positively correlated ( $r = .705, p < .01$ ), while *Awareness*

and Reflection was largely independent of both *Distress Intolerance* ( $r = .109$ , ns) and *Fear and Stigma* ( $r = -.056$ , ns), further supporting the conceptual distinctiveness of the adaptive *Awareness and Reflection* subscale.

Overall, then, these correlation patterns provide evidence of good convergent validity for the maladaptive subscales (*Distress Intolerance* and *Fear and Stigma*) with related constructs (PHQ-9, GAD-7, BAES) and good divergent validity for the adaptive *Awareness and Reflection* subscale.

**Table 2.6 - Correlations among ESSeNcE Subscales, PHQ-9, GAD-7 and BAES**

	Fear & Stigma	Awareness & Reflection	PHQ	GAD	BAES
Distress Intolerance	.705**	.109	.666**	.656**	.484**
Fear & Stigma	1	-.056	.610**	.564**	.751**
Awareness & Reflection		1	-.052	.050	-.179**
PHQ			1	.813**	.510**
GAD				1	.465**
BAES					1

Average inter-item correlations (AICs) were within the recommended range of .15-.50 for all three subscales (Clark & Watson, 1995), indicating adequate internal coherence without excessive redundancy. The *Distress Intolerance* subscale showed an AIC of .45, *Fear and Stigma* .43, and *Awareness and Reflection* .29. See Table 2.7 below.

**Table 2.7- Average Inter-Item Correlations (AICs) for the three ESSeNcE Subscales**

Subscale	Average Inter-Item Correlation (AIC)
Distress Intolerance	0.45
Fear and Stigma	0.43
Awareness and Reflection	0.29

**Note:** All AIC values fall within the recommended range of .15-.50 for broad psychological constructs (Clark & Watson, 1995).

### Reliability

Cronbach's alpha calculations showed strong internal consistency across ESSeNcE factors: *Distress Intolerance* ( $\alpha = .95$ ), *Fear and Stigma* ( $\alpha = .89$ ), *Awareness and Reflection* ( $\alpha = .78$ ). See Table 2.8 below:

**Table 2.8 - Internal Consistency of ESSeNcE Subscales**

Subscale	Number of Items	Cronbach's Alpha ( $\alpha$ )
Distress Intolerance	24	.95
Fear and Stigma	9	.89
Awareness and Reflection	7	.78

**Note:** Cronbach's alpha values indicate strong internal consistency for all ESSeNcE subscales.

Split-half reliability was assessed by dividing the 40-item scale into two 20-item halves. Cronbach's alpha for Part 1 ( $\alpha = .939$ ) and Part 2 ( $\alpha = .871$ ) indicated strong consistency. The correlation between halves was  $r = .766$ , with Spearman-Brown and Guttman Split-Half coefficients of 0.868 and 0.862, respectively, demonstrating good reliability, acceptable for exploratory scale development (Hair et al., 2010). See Table 2.9 for details:

**Table 2.9 - Reliability Statistics for the ESSeNcE Scale (Split-Half Method)**

Statistic	Value
Cronbach's Alpha - Part 1	.939
Number of Items - Part 1	20 <sup>a</sup>
Cronbach's Alpha - Part 2	.871
Number of Items - Part 2	20 <sup>b</sup>
Total Number of Items	40
Correlation Between Forms	.766
Spearman-Brown Coefficient (Equal Length)	.868

**Note:** Part 1 and Part 2 each represent half of the total 40 items used in the split-half reliability analysis. Superscripts <sup>a</sup> and <sup>b</sup> indicate the two respective halves of the scale.

### Descriptive Statistics and Correlations

Descriptive analyses (See Table 2.10) were conducted to examine the central tendency and dispersion of the three Factors. The results show that *Distress Intolerance* had a mean score of  $M = 3.61$  ( $SD =$

1.03), with a range of 4.21, suggesting that participants reported a moderate level of perceived control. *Fear and Stigma* had the highest mean (M = 3.99, SD = 1.04) and a range of 4.44, indicating a relatively high and varied level of fear across respondents. *Awareness and Reflection* recorded a mean of M = 3.72 (SD = 0.83) and a range of 4.14, reflecting moderately high awareness with slightly lower variability compared to the other constructs.

Overall, the standard deviations between 0.83 and 1.04 indicate that participant responses were reasonably dispersed, demonstrating moderate variation in perceptions across all three constructs.

**Table 2.10 - Total Descriptive Statistics for ESSeNcE Subscales**

Subscale	N	Range	Minimum	Maximum	M	SD
<i>Distress Intolerance</i>	232	4.21	1.38	5.58	3.60	1.03
<i>Fear and Stigma</i>	232	4.44	2.33	6.78	3.99	1.04
<i>Awareness and Reflection</i>	232	4.14	1.43	5.57	3.72	.83

Syntax for analyses process can be found in Appendix 2.8.

## Discussion

### *Key Findings and Factor Structure*

The present study developed and conducted an initial psychometric evaluation of the Emotional Schemas Self-Report of Negative Emotions (ESSeNcE). The scale was designed to measure individual beliefs, attitudes, and evaluations of negative emotions, incorporating both maladaptive and adaptive dimensions. A series of EFAs, supported by parallel analyses, was conducted on the initial 147-item pool, reducing it to a concise 40-item measure and yielding a clear three-factor structure - *Distress Intolerance*, *Fear and Stigma*, and *Awareness and Reflection*. This solution emerged through rigorous iterative item refinement, considering communalities, primary loadings, and cross-loadings, consistent with best practice recommendations for scale development (Clark & Watson, 1995; Costello & Osborne, 2005). The three factors collectively accounted for more than half of the total variance, which is acceptable for exploratory research in psychological measurement (Hair et al., 2019).

Dimension 1 (*Distress Intolerance*) captured difficulties in tolerating or controlling negative emotions, whereas Dimension 2 (*Fear and Stigma*) reflected self-directed shame and concerns about negative social evaluation associated with emotional expression. Dimension 3 (*Awareness and Reflection*) represented a more adaptive orientation, emphasising curiosity, insight and reflective engagement with negative emotional experiences. Inter-factor correlations demonstrated a moderate association between *Distress Intolerance* and *Fear and Stigma*, indicating that emotional intolerance tends to co-occur with shame-based avoidance and suppression. In contrast, *Awareness and Reflection* showed weak and non-significant correlations with the other two factors, supporting its role as a distinct and more helpful construct reflecting emotional insight rather than emotional avoidance.

Internal consistency estimates further supported the reliability of the subscales. Average inter-item correlations (AICs) were within the recommended range of .15-.50 for broad psychological constructs (Clark & Watson, 1995), with values of .45 for *Distress Intolerance*, .43 for *Fear and Stigma*, and .29 for *Awareness and Reflection*. These values indicate that items cohered well within each dimension without reflecting redundancy. Split-half reliability coefficients similarly demonstrated stable internal structure across the scale.

Evidence for convergent and divergent validity was also observed. *Distress Intolerance* and *Fear and Stigma* both showed strong positive correlations with the PHQ-9, GAD-7, and BAES, consistent with the theory that unhelpful emotional beliefs are associated with heightened anxiety, depression, and negative meta-emotional cognitions. In contrast, *Awareness and Reflection* demonstrated weak and largely non-significant associations with all three external measures, providing evidence of divergent validity and supporting the idea that this subscale reflects a healthy emotional stance that is largely separate from psychopathology. Taken together, these findings indicate that the ESSeNcE captures three conceptually distinct and psychometrically robust dimensions of negative emotional schemas, supporting its utility as a multidimensional measure for future research and clinical assessment.

### *Theoretical and Practical Implications*

From a theoretical perspective, the ESSeNcE model suggests that individuals' patterns of relating to their negative emotions, and consequently their questionnaire scores, should vary systematically according to their lived experience of mental health problems. Core emotion regulation theories suggest

that mental health problems involve a repeated reliance on unhelpful ways of managing emotions, such as avoidance, suppression or self-criticism. While these strategies may reduce distress in the short-term and provide temporary relief, they often prevent individuals from fully processing their emotions, thereby intensifying or maintaining emotional difficulties (Gross, 1998; Gratz & Roemer, 2004). For example, Gross's Process Model of Emotion Regulation proposes that emotions can be regulated at different stages of the emotion generation process, including through strategies such as cognitive reappraisal or emotional suppression (Gross, 1998). Similarly, ACT emphasises that experiential avoidance can contribute to the persistence of emotional distress (Hayes et al., 1996). These processes are common to multiple emotional disorders and are linked to stronger emotional reactions, lower ability to tolerate distress, poorer control over thoughts and attention, and increased feelings of internalised shame. Accordingly, individuals with lived experience of mental health difficulties would be theoretically expected to score higher on the *Distress Intolerance* and *Fear and Stigma* subscales of the ESSeNcE, reflecting more deep-seated maladaptive ways of responding to negative emotional experiences.

In contrast, the *Awareness and Reflection* dimension is theorised to capture a more adaptive mode of emotional engagement, reflecting the capacity to attend to, understand, and cognitively process emotional states. While acute emotional distress may be associated with reduced emotional clarity and reflective functioning, emotion regulation and metacognitive models suggest that lived experience, particularly when accompanied by psychological intervention, can facilitate the development of greater emotional awareness and reflective capacity over time. As a result, differences in *Awareness and Reflection* scores may reflect both weaknesses linked to vulnerability and positive learning or growth. This suggests that this dimension is different from the maladaptive subscales and can change with development and therapeutic support.

### *Comparisons with Existing Measures*

The ESSeNcE scale differs from existing measures of emotional schemas and scales measuring evaluation/appraisal in several important ways. While Leahy's Emotional Schema Scale (LESS; 2002) assesses individuals' beliefs about the nature and controllability of emotions, the ESSeNcE specifically captures self-directed stigma and evaluative responses to negative emotions. In contrast to Leahy's focus on meta-cognitive beliefs (e.g. '*Emotions are irrational*'), ESSeNcE addresses fear of judgment,

shame and the tendency to allow emotions to dominate cognition and behaviour. Also, unlike the LESS, which was developed using clinical participants and primarily focuses on emotion regulation strategies, the ESSeNcE is designed for use with a broader population and emphasises individuals' beliefs, attitudes and self-stigma related to their negative emotions rather than any focus on regulation strategies.

Other scales, such as the Beliefs About Emotions Scale (BAES; Rimes & Chalder, 2010) and the Brief Assessment of Emotional Experience (BAE; Manser et al., 2012), emphasise beliefs or awareness of emotions, but do not comprehensively integrate both maladaptive and adaptive dimensions within a single framework. Likewise, the Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004) focuses on regulatory deficits, while the Emotional Awareness Scale (Lane et al., 1990) and Emotional Processing Scale (Baker et al., 2007) target awareness and processing capacities. The Metacognitions Questionnaire (Cartwright-Hatton et al., 1997) focuses on general thinking about emotions, rather than on self-judgment or shame about them. Thus, the ESSeNcE uniquely includes maladaptive components (*Distress Intolerance and Fear and Stigma*) that reflect avoidance, shame, and negative self-judgment, as well as adaptive components (*Awareness and Reflection*) that reflect insight, understanding, and thoughtful engagement with emotions. It also focuses on people's actual emotional experiences rather than just their beliefs about emotions. By combining both maladaptive and adaptive aspects, the ESSeNcE provides a more complete and clinically useful assessment of emotional self-stigma than existing measures.

### *Strengths and Limitations*

This study has several important strengths. The ESSeNcE was developed using careful, step-by-step psychometric methods (Clark & Watson, 1995; Costello & Osborne, 2005), including parallel analysis, multiple rounds of removing weak or overlapping items, and attention to both statistical results and theory. The final three-factor solution was clear, reliable, and showed early signs of convergent and discriminant validity. By including both maladaptive and adaptive aspects, the scale offers a useful and clinically relevant way to understand emotional experiences.

However, a number of limitations should be acknowledged, particularly in relation to sampling. The sample included a high proportion of PWLE of mental health difficulties which, while offering valuable insight into informed perspectives, may have introduced bias in item responses, given that the final

measure is intended for use in the general population. This overrepresentation may therefore affect how items are interpreted by individuals without such experience. The sample was also predominantly female, reflecting the characteristics of the recruited population but limiting generalisability to male and gender-diverse groups. In addition, there was a disproportionate representation of individuals aged 24-35 and those with higher educational attainment, with 64.3% holding an undergraduate degree or above. These patterns are likely attributable to the use of convenience sampling and recruitment via online and social media platforms, which tend to engage younger, more digitally literate and research-engaged individuals. Consequently, the findings may be less applicable to older adults and those from more educationally diverse backgrounds.

It should also be noted that ethnicity data were not collected, and therefore cross-cultural validity of the measure cannot be evaluated. This is arguably an important factor, given the emphasis on cross-cultural applicability with the COSMIN framework, which was used to evaluate psychometric quality in the previous chapter. While the review identified this as a gap in the literature, this current empirical project focused on its core psychometric properties (e.g. factor structure and internal consistency), and an initial measure for further development and final validation in a more representative sample.

The study also study relied exclusively on self-report, which may introduce social desirability bias or limited self-awareness. Another limitation was that, while the EFA provided an initial factor structure, a confirmatory factor analysis (CFA) is necessary to validate the model. Temporal stability was also not assessed, so test-retest reliability remains unknown.

### *Future Directions*

Future research should aim to address these limitations through multiple avenues. First, CFA with independent samples is needed to test the robustness and generalisability of the three-factor structure. Second, comprehensive validity studies should be conducted, including convergent validity with related measures (e.g., BAE, DERS, BAEQ, Emotional Awareness Scale, Emotional Processing Scale) and criterion validity with mental health outcomes (depression, anxiety, and help-seeking behaviours). Third, demographic validation is essential to determine whether the scale functions equivalently across gender, age, differing levels of educational attainment, and ethnicity. Future research should seek to mitigate these limitations through the use of stratified and targeted recruitment strategies, as well as broader outreach beyond academic and professional networks. Fourth, longitudinal and intervention

studies could assess whether ESSeNcE scores are sensitive to change following psychological interventions such as mindfulness-based cognitive therapy, acceptance and commitment therapy, or emotion-focused therapies. Finally, item reduction strategies, such as Item Response Theory (IRT), could produce a shorter, practical scale for clinical and research use by evaluating each item's difficulty and its ability to distinguish between individuals, without compromising psychometric quality (Hambleton et al., 1991).

### *Conclusion*

The ESSeNcE scale represents a novel, psychometrically sound tool which measures the appraisal of negative emotions. Its three-factor structure captures both maladaptive and adaptive components of the emotional experience. Unlike existing measures such as the LESS, ESSeNcE emphasises self-stigma, lived emotional experience, and adaptive reflection, bridging a critical gap in the assessment of beliefs around emotions. Although further validation is required, the ESSeNcE offers a promising tool for research, clinical assessment and intervention development, enhancing understanding of the complex interplay between emotion, self-perception, and psychological well-being.

### *Reflexivity*

As a Clinical Psychologist in training, my practitioner-scientist stance and therapeutic orientation, have likely shaped the ESSeNcE in ways that both strengthen and constrain its development. Clinical experience of working with individuals in high distress may have led to an overemphasis on intolerance during item generation. This risks underrepresenting more adaptive or context-dependent functions of negative emotions. Similarly, repeated exposure to shame, fear and stigma surrounding emotions in clinical settings may have biased the framing of items within the '*Fear & Stigma*' dimension. My own use of acceptance-based approaches (i.e., ACT) could also have influenced the identification and interpretation of the '*Awareness & Reflection*' dimension. While this aligns with established therapeutic models, it raises the possibility of bias in data-driven techniques. At the stage of factor naming and interpretation, subjective judgement was unavoidable, and my clinical framework may have favoured coherence with existing models or alternative readings of the data. While steps were taken to enhance rigour, researcher bias cannot be eliminated. All of the above underscore the need for further validation across diverse samples and perspectives

## References

- Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). *Emotion-regulation strategies across psychopathology: A meta-analytic review*. *Clinical Psychology Review, 30*(2), 217-237.
- Baker, R., Thomas, S., Thomas, P. W., Gower, P., Santonastaso, M., & Whittlesea, A. (2010). The Emotional Processing Scale: scale refinement and abridgement (EPS-25). *Journal of psychosomatic research, 68*(1), 83-88.
- Bateman, A., & Fonagy, P. (2016). *Mentalization-based treatment for personality disorders: A practical guide*. Oxford University Press.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in public health, 6*, 149.
- Brown, L. A., Gaudiano, B. A., & Miller, I. W. (2011). Investigating the similarities and differences between practitioners of second-and third-wave cognitive-behavioral therapies. *Behavior modification, 35*(2), 187-200.
- Byrne, B. M. & van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of non-equivalence. *International Journal of Testing, 10*(2), 107-132.
- Capobianco, L., & Nordahl, H. (2021). A brief history of metacognitive therapy: From cognitive science to clinical practice. *Cognitive and Behavioral Practice*.
- Clark, L. A., & Watson, D. (2016). Constructing validity: Basic issues in objective scale development.
- Cartwright-Hatton, S., & Wells, A. (1997). Beliefs about worry and intrusions: The Meta-Cognitions Questionnaire and its correlates. *Journal of anxiety disorders, 11*(3), 279-296.
- Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological assessment, 31*(12), 1412.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-9.
- Craig, K., Hale, D., Grainger, C., & Stewart, M. E. (2020). Evaluating metacognitive self-reports: systematic reviews of the value of self-report in metacognitive research. *Metacognition and Learning, 15*, 155-213.
- Cresswell, J. (2013). *Qualitative inquiry & research design: Choosing among five approaches*.

- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299.
- Field, A. (2024). *Discovering statistics using IBM SPSS statistics*. Sage publications limited.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.
- Gilbert, P. (2010). *Compassion-focused therapy: Distinctive features*. Routledge.
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of psychopathology and behavioral assessment, 26*(1), 41-54.
- Greenberg, L. S. (2015). *Emotion-focused therapy: Coaching clients to work through their feelings*. American Psychological Association.
- Greenberg, L. S. (2010). Emotion-focused therapy: An overview. *Turkish Psychological Counselling and Guidance Journal, 4*(33), 1-12.
- Greenberg, L. S., Rice, L. N., & Elliott, R. (1993) *Facilitating Emotional Change: The moment-by-moment process*. New York: Guilford Press.
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of general psychology, 2*(3), 271-299.
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European business review, 31*(1), 2-24.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior therapy, 35*(4), 639-665.
- Hayes, S. C., Strosahl, K., & Wilson, K. G. (2012). *Acceptance and commitment therapy: The process and practice of mindful change* (2nd ed.). Guilford Press
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: an experiential approach to behavior change*. New York: Guilford Press
- Jorm, A. F., Kelly, C. M., & Morgan, A. J. (2007). Participant distress in psychiatric research: A systematic review. *Psychological medicine, 37*(7), 917-926.
- Kline, R. B. (2011). *Principles and practice of structural equation modelling (3<sup>rd</sup> ed.)*. New York, NY: Guildford Press.

- Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric annals*, 32(9), 509-515.
- Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A., & Zeitlin, S. B. (1990). The Levels of Emotional Awareness Scale: A cognitive-developmental measure of emotion. *Journal of personality assessment*, 55(1-2), 124-134.
- Leahy, R. L. (2019). Introduction: Emotional schemas and emotional schema therapy. *International Journal of Cognitive Therapy*, 12(1), 1-4.
- Leahy, R. L. (2016). Emotional schema therapy: A meta-experiential model. *Australian Psychologist*, 51(2), 82-88.
- Leahy, R. L. (2015). *Emotional Schema Therapy*. Guilford Press: New York
- Leahy, R. L. (2012). Introduction: Emotional schemas, emotion regulation, and psychopathology. *International Journal of Cognitive Therapy*, 5(4), 359-361.
- Leahy, R. L. (2007). Emotional schemas and resistance to change in anxiety disorders. *Cognitive and behavioral Practice*, 14(1), 36-45.
- Leahy, R. L. (2002). A model of emotional schemas. *Cognitive and behavioral practice*, 9(3), 177-190.
- Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. Guilford Press.
- Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, 4(2), 73-79.
- Neff, K.D. (2003). *The development and validation of a scale to measure self-compassion*. *Self and Identity*, 2(3), 223-250.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, 32, 396-402.
- Ortony, A., Clore, G. L., & Collins, A. (2022). *The cognitive structure of emotions*. Cambridge university press.
- Padgett, D. K. (2016). *Qualitative methods in social work research* (Vol. 36). Sage publications.
- Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta psychologica*, 104(1), 1-15.

- Rimes, K. A., & Chalder, T. (2010). The Beliefs about Emotions Scale: validity, reliability and sensitivity to change. *Journal of psychosomatic research, 68*(3), 285-292.
- Rosellini, A. J., & Brown, T. A. (2021). Developing and validating clinical questionnaires. *Annual review of clinical psychology, 17*, 55-81.
- Rowan, N., & Wulff, D. (2007). Using qualitative methods to inform scale development. *Qualitative Report, 12*(3), 450-466.
- Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2002). *Mindfulness-based cognitive therapy for depression: a new approach to preventing relapse*. New York: Guilford Press
- Shaw, R. M., Howe, J., Beazer, J., & Carr, T. (2020). Ethics and positionality in qualitative research with vulnerable and marginal groups. *Qualitative Research, 20*(3), 277-293.
- Siddaway, A. P., Wood, A. M., O'Carroll, R. E., & O'Connor, R. C. (2019). Characterizing self-injurious cognitions: Development and validation of the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS). *Psychological assessment, 31*(5), 592.
- Skinner, B. F. (1953). Some contributions of an experimental analysis of behavior to psychology as a whole. *American Psychologist, 8*, 69-78.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine, 166*(10), 1092-1097.
- Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy*, 41-71.
- Velicer, W.F., & Fava, J.L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods, 3*(2), 231-251.
- Watson, J. B. (1925). *Behaviorism*. New York, NY: Norton.
- Wells, A. (2011). *Metacognitive therapy for anxiety and depression*. Guilford press.
- Wells, A. (1987). *Self-attentional processes in anxiety: an experimental study* (Doctoral dissertation, University of Aston in Birmingham).
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The counseling psychologist, 34*(6), 806-838.
- Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological bulletin, 99*(3), 432.

## Appendices

### Appendix 1.1 - PRISMA Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
<b>TITLE</b>			
Title	1	Identify the report as a systematic review.	
<b>ABSTRACT</b>			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
<b>INTRODUCTION</b>			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	
<b>METHODS</b>			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	

Section and Topic	Item #	Checklist item	Location where item is reported
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	
<b>RESULTS</b>			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	
Study characteristics	17	Cite each included study and present its characteristics.	

Section and Topic	Item #	Checklist item	Location where item is reported
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	
<b>DISCUSSION</b>			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	
	23b	Discuss any limitations of the evidence included in the review.	
	23c	Discuss any limitations of the review processes used.	
	23d	Discuss implications of the results for practice, policy, and future research.	
<b>OTHER INFORMATION</b>			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	

Section and Topic	Item #	Checklist item	Location where item is reported
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	
Competing interests	26	Declare any competing interests of review authors.	
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	

### Appendix 1.2 Search Strategy

Database	Search Terms (linked with AND)	Filters Applied
<i>CINAHL</i>	<ol style="list-style-type: none"> <li>1. (emotion* OR affect OR mood* OR feeling* OR “model of emotional schemas”)</li> <li>2. (evaluat* OR belief* OR apprais* OR assess* OR judg* OR aware* OR think* OR attitude* OR view* OR percep* OR meta-cognit* OR metacognit* OR thought* OR cognit* OR reflect* OR process* OR “self-evaluat*” OR ruminat* OR concept* OR respon* OR attent* OR “self-focus” OR schema* OR intellig* OR focus OR meta-experien* or metaexperien*)</li> <li>3. (“measurement tool” OR “psychometric properties” OR “measurement scale” OR “questionnaire measure” OR psychometric* OR assessment OR “assessment tool” OR “questionnaire” OR checklist OR “scale development” OR “questionnaire development” OR development OR validation OR “develop* and valid*” OR “validation tool” OR model OR tool OR scale OR “assessment scale” OR “self-assessment tool” OR “self-assessment” OR “self-assessment measure” OR “self-assessment questionnaire” OR self-assessment scale”)</li> </ol>	<p><i>Language:</i> English  <i>Peer-reviewed:</i> Yes  <i>Document Type:</i> Exclude dissertations, book chapters, conference proceedings</p>
<i>Medline</i>	<ol style="list-style-type: none"> <li>1. (emotion* OR affect OR mood* OR feeling* OR “model of emotional schemas”)</li> <li>2. (evaluat* OR belief* OR apprais* OR assess* OR judg* OR aware* OR think* OR attitude* OR view* OR percep* OR meta-cognit* OR metacognit* OR thought* OR cognit* OR reflect* OR process* OR “self-evaluat*” OR ruminat* OR concept* OR respon* OR attent* OR “self-focus” OR schema* OR intellig* OR focus OR meta-experien* or metaexperien*)</li> <li>3. (“measurement tool” OR “psychometric properties” OR “measurement scale” OR “questionnaire measure” OR</li> </ol>	<p><i>Language:</i> English  <i>Peer-reviewed:</i> Yes  <i>Document Type:</i> Exclude dissertations, book chapters, conference proceedings</p>

	<p>psychometric* OR  assessment OR "assessment  tool" OR "questionnaire" OR  checklist OR "scale  development" OR  "questionnaire development"  OR development OR  validation OR "develop* and  valid*" OR "validation tool" OR  model OR tool OR scale OR  "assessment scale" OR "self-  assessment tool" OR "self-  assessment" OR "self-  assessment measure" OR  "self-assessment  questionnaire" OR self-  assessment scale")</p>	
<i>PsychArticles</i>	<ol style="list-style-type: none"> <li>1. (emotion* OR affect OR mood* OR feeling* OR "model of emotional schemas")</li> <li>2. (evaluat* OR belief* OR apprais* OR assess* OR judg* OR aware* OR think* OR attitude* OR view* OR percep* OR meta-cognit* OR metacognit* OR thought* OR cognit* OR reflect* OR process* OR "self-evaluat*" OR ruminat* OR concept* OR respon* OR attent* OR "self-focus" OR schema* OR intellig* OR focus OR meta-experien* or metaexperien*)</li> <li>3. ("measurement tool" OR "psychometric properties" OR "measurement scale" OR "questionnaire measure" OR psychometric* OR assessment OR "assessment tool" OR "questionnaire" OR checklist OR "scale development" OR "questionnaire development" OR development OR validation OR "develop* and valid*" OR "validation tool" OR model OR tool OR scale OR "assessment scale" OR "self-assessment tool" OR "self-assessment" OR "self-assessment measure" OR "self-assessment questionnaire" OR self-assessment scale")</li> </ol>	<p><i>Language:</i> English  <i>Peer-reviewed:</i> Yes  <i>Document Type:</i> Exclude dissertations, book chapters, conference proceedings</p>
<i>PsycINFO</i>	<ol style="list-style-type: none"> <li>1. (emotion* OR affect OR mood* OR feeling* OR "model of emotional schemas")</li> <li>2. (evaluat* OR belief* OR apprais* OR assess* OR judg* OR aware* OR think* OR attitude* OR view* OR</li> </ol>	<p><i>Language:</i> English  <i>Peer-reviewed:</i> Yes  <i>Document Type:</i> Exclude dissertations, book chapters, conference proceedings</p>

	<p>3. percep* OR meta-cognit* OR metacognit* OR thought* OR cognit* OR reflect* OR process* OR "self-evaluat*" OR ruminat* OR concept* OR respon* OR attent* OR "self-focus" OR schema* OR intellig* OR focus OR meta-experien* or metaexperien*) ("measurement tool" OR "psychometric properties" OR "measurement scale" OR "questionnaire measure" OR psychometric* OR assessment OR "assessment tool" OR "questionnaire" OR checklist OR "scale development" OR "questionnaire development" OR development OR validation OR "develop* and valid*" OR "validation tool" OR model OR tool OR scale OR "assessment scale" OR "self-assessment tool" OR "self-assessment" OR "self-assessment measure" OR "self-assessment questionnaire" OR self-assessment scale")</p>	
Psychology & Behavioural Science	<p>1. (emotion* OR affect OR mood* OR feeling* OR "model of emotional schemas")</p> <p>2. (evaluat* OR belief* OR apprais* OR assess* OR judg* OR aware* OR think* OR attitude* OR view* OR percep* OR meta-cognit* OR metacognit* OR thought* OR cognit* OR reflect* OR process* OR "self-evaluat*" OR ruminat* OR concept* OR respon* OR attent* OR "self-focus" OR schema* OR intellig* OR focus OR meta-experien* or metaexperien*)</p> <p>3. ("measurement tool" OR "psychometric properties" OR "measurement scale" OR "questionnaire measure" OR psychometric* OR assessment OR "assessment tool" OR "questionnaire" OR checklist OR "scale development" OR "questionnaire development" OR development OR validation OR "develop* and valid*" OR "validation tool" OR model OR tool OR scale OR "assessment scale" OR "self-</p>	<p><i>Language:</i> English  <i>Peer-reviewed:</i> Yes  <i>Document Type:</i> Exclude dissertations, book chapters, conference proceedings</p>

	assessment tool” OR “self-assessment” OR “self-assessment measure” OR “self-assessment questionnaire” OR self-assessment scale”)	
--	--	--

### Appendix 1.3 - Data Extraction Checklist

1. Study Identification	<ul style="list-style-type: none"> <li>• Author(s)</li> <li>• Publication year</li> <li>• Journal name</li> <li>• Location of study (country)</li> <li>• Setting</li> </ul>
2. Study Characteristics	<ul style="list-style-type: none"> <li>• Design</li> <li>• Eligibility criteria</li> </ul>
3. Participant Information	<ul style="list-style-type: none"> <li>• Age range</li> <li>• Clinical or non-clinical population</li> <li>• Gender</li> <li>• Sample size</li> <li>• Ethnicity</li> </ul>
4. Emotional Schema Tool Information	<ul style="list-style-type: none"> <li>• Tool Name</li> <li>• Construct(s) measured</li> <li>• Number of items</li> <li>• Scoring method</li> <li>• Response options</li> <li>• Whether self-report</li> <li>• Factor analysis method(s)</li> </ul>
5. Psychometric Properties Assessed Using COSMIN	<ul style="list-style-type: none"> <li>• Internal consistency</li> <li>• Test re-test</li> <li>• Measurement error</li> <li>• Content validity</li> <li>• Structural validity</li> <li>• Cross-cultural validity</li> <li>• Criterion validity</li> <li>• Responsiveness</li> </ul>

### Appendix 1.4 - COSMIN Risk of Bias Quality Appraisal Tool

Term			Definition
Domain	Measurement property	Aspect of a measurement property	
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same health related-patient reported outcomes (HRPRO) (internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater)
	Internal consistency		The degree of the interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is due to 'true'† differences between patients
	Measurement error		The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Validity			The degree to which an HR-PRO instrument measures the construct(s) it purports to measure
	Content validity		The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured
		Face validity	The degree to which (the items of) an HR-PRO instrument indeed looks as though they are an adequate reflection of the construct to be measured
	Construct validity		The degree to which the scores of an HR-PRO instrument are consistent with hypotheses (for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the HR-PRO instrument validly measures the construct to be measured
		Structural validity	The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the construct to be measured
		Hypotheses validity	Idem construct validity
		Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument
	Criterion validity		The degree to which the scores of an HR-PRO instrument are an adequate reflection of a 'gold standard'
Responsiveness			The ability of an HR-PRO instrument to detect change over time in the construct to be measured
	Responsiveness		Idem responsiveness
Interpretability*			Interpretability is the degree to which one can assign qualitative meaning - that is, clinical or commonly understood connotations - to an instrument's quantitative scores or change in scores.

### Appendix 1.5 - Further Information on Included Scales and Their Utility

Measure	Reference	Further information on what it Measures	Clinical/Research Utility
Perth Emotional Reactivity Scale (PERS)	Becerra et al (2019)	<p>Emotional Sensitivity</p> <ul style="list-style-type: none"> <li>How easily and frequently a person experiences emotions, with higher sensitivity indicating that emotions are triggered more readily by events.</li> </ul> <p>Emotional Intensity</p> <ul style="list-style-type: none"> <li>Strength or magnitude of emotional reactions, with higher intensity indicating more powerful emotional experiences.</li> </ul> <p>Emotional Persistence</p> <ul style="list-style-type: none"> <li>How long emotions persist, with higher persistence indicating greater difficulty 'letting go' of emotional reactions.</li> </ul>	<p>Useful in clinical assessment, formulation, and treatment planning, particularly for interventions targeting emotion regulation, mood disorders, and stress-related difficulties.</p> <p>Can guide tailored psychological interventions by highlighting patterns of emotional reactivity that maintain or exacerbate mental health problems.</p>
Emotion Beliefs Questionnaire (EBQ)	Becerra et al (2020)	<p>Assesses beliefs about whether emotions are controllable and useful, separately for positive and negative emotions.</p> <p>Identifies whether individuals see emotions as something they can influence or as harmful and to be avoided.</p>	<p>Highlights beliefs that shape emotion regulation strategies and influence psychological well-being.</p>
Difficulties in Emotional Regulation Scale (DERS)	Gratz et al (2004)	<p>Non-acceptance - difficulty accepting negative emotions.</p> <p>Goal-directed difficulties - trouble focusing or acting when distressed.</p> <p>Impulse control - problems managing behaviour under strong emotions.</p> <p>Emotional awareness - limited attention to or recognition of emotions</p> <p>Emotional clarity - difficulty understanding or identifying emotions.</p>	<p>Useful for tracking treatment progress and evaluating the effectiveness of therapeutic approaches such as CBT, DBT or ACT.</p>
Brief- Mentalized Affectivity Scale (BMAS)	Greenberg et al (2021)	<p>Measures the capacity to understand, reflect on and manage one's emotions adaptively.</p> <p>Assesses how individuals identify, process and express emotions, integrating them with reflective thinking.</p>	<p>Useful in clinical assessment and intervention planning to support emotional processing, self-reflection and overall psychological well-being.</p>

<p>Mentalizing Emotions Questionnaire (MEQ)</p>	<p>Kasper et al (2024)</p>	<p>Assesses the ability to mentalise emotions, understanding and reflecting on one's own and others' emotional states.</p> <p>Evaluates how individuals interpret emotions, link them to thoughts and intentions, and respond adaptively.</p> <p>Difficulties on the MEQ are associated with interpersonal challenges, emotion regulation problems, and psychological distress.</p>	<p>Guiding of interventions that target emotional understanding, regulation, and relational skills.</p>
<p>Leahy Emotional Schema Scale (LESS)</p>	<p>Leahy (2002)</p>	<p>Evaluates how people understand, evaluate, and respond to emotions.</p> <p>Assesses beliefs about control, duration, expression, and acceptability of emotions.</p>	<p>Clinically useful for identifying maladaptive emotional beliefs linked to distress, avoidance, and regulation difficulties.</p>
<p>Beliefs About Emotions Questionnaire (BAEQ)</p>	<p>Manser et al (2012)</p>	<p>Assesses beliefs and attitudes about emotions, including whether they are controllable, acceptable, or helpful.</p> <p>Identifies maladaptive beliefs that contribute to emotional avoidance, distress, and regulation difficulties.</p>	<p>Useful in clinical and research settings to understand how beliefs influence coping, adjustment, and mental health outcomes.</p>
<p>Meta-Emotion Scale (MES)</p>	<p>Mitmansgruber et al (2009)</p>	<p>Assesses awareness, understanding, and regulation of one's own emotions and attitudes toward emotional experiences.</p> <p>Evaluates how people perceive and manage emotions, including tendencies toward acceptance, reflection, or suppression.</p>	<p>Useful in clinical and research settings to understand emotional processing, regulation difficulties and factors affecting well-being.</p>
<p>Perth Emotion Regulation Competency Inventory (PERCI)</p>	<p>Preece et al (2018)</p>	<p>Assesses competence in emotion regulation, including awareness, understanding, and adaptive management of emotions.</p> <p>Evaluates how effectively a person can recognise, process, and respond to emotional experiences.</p>	<p>Useful for identifying strengths and difficulties, guiding interventions, and tracking change over time.</p>

The Beliefs about Emotions Scale (BES)	Rimes & Chalder (2010)	Assesses beliefs about emotions, including control, expression, and acceptability.  Identifies maladaptive beliefs linked to emotional suppression, avoidance, and distress.	Useful in understanding how beliefs affect coping, regulation, and mental health outcome.
Emotion and Regulation Beliefs Scale (ERBS)	Veilleux et al (2015)	Measures beliefs about emotions and their regulation, including controllability, usefulness, and manageability.  Identifies maladaptive beliefs that may lead to emotional avoidance and poor regulation.	Useful for understanding how beliefs influence coping and psychological well-being.

## Appendix 2.1 - Revised Research Proposal

<https://osf.io/mqvwd/files/vc5wb>

## Appendix 2.2 - Project Approval Letter



17<sup>th</sup> April 2024

Dear Prof McLeod,

**MVLS College Ethics Committee**

*Project Title: Emotional Schemas Self-Report Measure: Development and Validation The ESSence Survey*

*Project No: 200230293*

The College Ethics Committee has reviewed your application and has agreed that there is no objection on ethical grounds to the proposed study. It is happy therefore to approve the project, subject to the following conditions:

- Project end date: End May 2025
- The data should be held securely for a period of ten years after the completion of the research project, or for longer if specified by the research funder or sponsor, in accordance with the University's Code of Good Practice in Research: [https://www.gla.ac.uk/media/media\\_490311\\_en.pdf](https://www.gla.ac.uk/media/media_490311_en.pdf)
- The research should be carried out only on the sites, and/or with the groups defined in the application.
- Any proposed changes in the protocol should be submitted for reassessment, except when it is necessary to change the protocol to eliminate hazard to the subjects or where the change involves only the administrative aspects of the project. The Ethics Committee should be informed of any such changes.
- You should submit a short end of study report to the Ethics Committee within 3 months of completion.
- For projects requiring the use of an online questionnaire, the University has an Online Surveys account for research. To request access, see the University's application procedure at <https://www.gla.ac.uk/research/strategy/ourpolicies/useofonlinesurveystoolforresearch/>.

Yours sincerely



Jesse Dawson  
MD, BSc (Hons), FRCP, FESO  
Professor of Stroke Medicine  
Chair MVLS Research Ethics Committee

Institute of Cardiovascular and Medical Sciences  
College of Medical, Veterinary & Life Sciences  
Room M0.03  
Office Block  
Queen Elizabeth University Hospital  
Glasgow  
G31 4TF  
Tel - 0141 431 2888  
[jesse.dawson@glasgow.ac.uk](mailto:jesse.dawson@glasgow.ac.uk)

## Appendix 2.3 - Materials for Phase 2

Participant Information Sheet: <https://osf.io/mqvwd/files/bnpm9>

Consent Form: <https://osf.io/mqvwd/files/e7ckt>

Privacy Notice: <https://osf.io/mqvwd/files/8z52m>

Demographic information sheet: <https://osf.io/mqvwd/files/u57ah>

Email to clinicians and researchers: <https://osf.io/mqvwd/files/5u3cr>

Social Media advert: <https://osf.io/mqvwd/files/39nyr>

### Appendix 2.4 - ESSeNcE 147-item list Phase 3

Indicate your agreement or disagreement with the following statements:

(Strongly disagree; somewhat disagree; neither agree nor disagree; somewhat agree; strongly agree)

1.	I worry that if I let myself have negative emotions I might have difficulty controlling my actions
2.	I worry that I won't be able to control my negative emotions
3.	If I didn't control a negative emotion, and then something bad happened, it would be my fault
4.	I should be in control of my negative emotions all of the time
5.	I find it difficult to control my negative emotions
6.	Over time, I could learn to control my negative emotions
7.	Not being able to control how I express my negative emotions is a sign of weakness
8.	When I experience negative emotions, they won't stop
9.	I will be punished for not controlling certain negative emotions
10.	If I couldn't control my negative emotions, I wouldn't be able to go about my day as normal
11.	I'm not able to learn techniques to effectively control my negative emotions
12.	I usually have control over my negative emotions
13.	When I'm experiencing negative emotions, it's difficult to control what I say or do
14.	I have very little control over my negative emotions
15.	I ought not to allow myself to have strong negative emotions
16.	When I experience negative emotions, I have no control over what I say or do
17.	I need to have control over my negative emotions in order to prevent myself from having a breakdown
18.	My negative emotions are outside my own control
19.	It doesn't matter how hard I try, I can't change my negative emotions
20.	When I'm experiencing negative emotions, there's nothing I can do about modifying them
21.	When I have extremely intense negative emotions, I'm not really responsible for how I cope with them
22.	I allow my negative emotions to dictate my thoughts and behaviours
23.	No matter how hard I try, I can't really change the negative emotions that I have
24.	It is hard for me to wind down from negative emotions
25.	My negative emotions are bad and need to be controlled
26.	When I experience negative emotions, I often find that I lose control of myself emotionally
27.	When I experience negative emotions, I often find that I lose control of myself physically
28.	My negative emotions persist, no matter how hard I try to stop them
29.	Not understanding how to control my negative emotions makes me feel alien to others
30.	I can't ignore my negative emotions
31.	Some negative emotions are wrong to have
32.	I can't allow myself to have certain kinds of negative emotions
33.	I feel ashamed of my negative emotions
34.	I shouldn't have some of the negative emotions that I have
35.	I don't want anyone to know about some of my negative emotions
36.	If others find out about my negative emotions it will ruin my relationships
37.	If others find out about my negative emotions it will ruin my life
38.	I believe that experiencing negative emotions is a sign of weakness
39.	I don't want to admit to having negative emotions, but I know that I have them
40.	I'm embarrassed about not being able to cope when I'm experiencing negative emotions
41.	It's bad to have negative emotions
42.	My negative emotions make me feel lost
43.	It's embarrassing to have negative emotions
44.	I believe it's unacceptable for me to ever have certain negative emotions
45.	I should feel ashamed of having negative emotions
46.	I should punish myself for having negative emotions
47.	I think that my negative emotions are strange or weird
48.	It's foolish to have negative emotions
49.	Having negative emotions means I'm a failure
50.	My negative emotions are harmful for me
51.	My negative emotions are so strong they are dangerous for me

52.	My negative emotions might damage me psychologically
53.	My negative emotions make me feel everything is pointless
54.	My negative emotions aren't likely to cause me any actual harm
55.	My negative emotions affect my self-image
56.	My negative emotions cause me a lot of problems in my daily life
57.	If I experience negative emotions, then I'm a bad person
58.	My negative emotions affect my ability to take care of myself
59.	When I experience negative emotions, it can affect others around me
60.	When I experience negative emotions, I feel sure that everything I touch will go wrong
61.	I can keep safe if I don't let myself have negative emotions
62.	I see no need for negative emotions.
63.	My negative emotions cause me lots of problems
64.	My negative emotions get in the way of my daily life
65.	My negative emotions are very unhelpful to me
66.	My negative emotions prevent me from progressing in my life
67.	My negative emotions reflect how difficult my life is
68.	My life would be better if I never experienced negative emotions
69.	I don't need my negative emotions
70.	My negative emotions stop me from getting things done
71.	I have very little use for negative emotions.
72.	My negative emotions regularly take over my attention
73.	When I'm experiencing negative emotions, it's hard for me to imagine feeling any other way
74.	When I'm experiencing negative emotions, it's hard for me to think straight
75.	I experience negative emotions very strongly
76.	I often feel overwhelmed by my negative emotions
77.	I'm not as able to function when I'm experiencing negative emotions
78.	My negative emotions can completely take control of me
79.	I can't tolerate my negative emotions
80.	I can't cope with my negative emotions, even for a short while
81.	I sometimes worry that I'll have strong negative emotions that won't go away
82.	When I experience negative emotions, it feels like they consume me
83.	When I experience negative emotions, it's all I can think about it for a long time
84.	When I experience negative emotions, it means that I'm about to have a breakdown
85.	Some of my negative emotions don't make sense to me
86.	I don't understand my negative emotions
87.	Negative emotions don't appear randomly, there's always a reason for feeling that way
88.	My negative emotions affect my perception of reality
89.	It's important to understand why I experience negative emotions
90.	I'm often confused or puzzled by my negative emotions
91.	Sometimes it is quite rational for me to experience negative emotions
92.	I find it difficult to explain my negative emotions to myself and others
93.	Sometimes I don't know exactly what kind of negative emotion I'm having
94.	I don't recognise when I'm having negative emotions but others tell me I am
95.	I often think about my negative emotions and what they mean
96.	Sometimes I have a negative emotion that seems to come out of nowhere
97.	Sometimes I have strong negative emotions but I'm not sure what they mean
98.	It's difficult for me to understand my negative emotions
99.	I feel able to express my negative emotions openly
100.	I accept my negative emotions as normal
101.	I easily recognise my negative emotions as I experience them
102.	I believe that it's important to let myself have negative emotions, in order to get my feelings "out"
103.	I'm able to understand my negative emotions
104.	Experiencing negative emotions tells me something useful about myself
105.	My negative emotions are an accurate reflection of what's happening around me
106.	When I experience negative emotions, it usually means there's something to be upset about
107.	When I experience negative emotions, it's a sign that something isn't right
108.	My negative emotions help me to understand what's happening
109.	My negative emotions can spur me on to do things differently next time
110.	My negative emotions are clear to me
111.	I can usually define my negative emotions
112.	I know why my negative emotions change
113.	Negative emotions help me to get things sorted out in my mind

114.	Negative emotions help me to realise what's important to me
115.	Negative emotions help me to be creative
116.	Negative emotions help me to make decisions about what to do next
117.	Negative emotions can help me to avoid disastrous situations in the future
118.	Negative emotions inspire me to come up with new ideas
119.	Negative emotions help me to avoid problems in the future
120.	Negative emotions help me to solve problems
121.	I have all my best ideas when I'm experiencing negative emotions
122.	I use negative emotions to help myself keep trying in the face of obstacles
123.	I pay a lot of attention to my negative emotions
124.	I am constantly aware of my negative emotions
125.	I constantly examine my negative emotions
126.	If I think too much about my negative emotions, then I feel worse
127.	I'm aware of my negative emotions as I experience them
128.	I often become aware of my negative emotions in different situations
129.	I'm usually very conscious of my negative emotions
130.	I usually spend time thinking about my emotions
131.	When I'm experiencing negative emotions, I sit by myself and think about how bad I feel
132.	I often wonder why I have so many negative emotions
133.	I think my negative emotions and state of mind deserve to be paid attention to
134.	I think that there are negative emotions that I have, that I'm unaware of
135.	I am interested in my negative emotions
136.	I try to understand the different reasons for my negative emotions
137.	I think it is helpful to understand the causes of my negative emotions
138.	I am interested in understanding my negative emotions
139.	With some distance I can understand my negative emotions in a new way
140.	I think it is useful to talk about my negative emotions
141.	My negative emotions don't always seem to belong to me
142.	Some of my negative emotions seem blunt
143.	Some of my negative emotions seem dull
144.	I shut off from my negative emotions
145.	My negative emotions make me feel a sense of emptiness
146.	I'm not affected by my negative emotions
147.	My negative emotions have to be very intense to feel real to me

## Appendix 2.5 - Materials for Phase 3

Participant information sheet: <https://osf.io/mqvwd/files/u9acz>

Consent form: <https://osf.io/mqvwd/files/e7ckt>

Demographic information sheet: <https://osf.io/mqvwd/files/69wyp>

**Appendix 2.6 - Revised Analysis Plan**

<https://osf.io/mqvwd/files/k6pa4>

### Appendix 2.7 - Additional Questionnaires Phase 3

*Before proceeding to the survey, we would like to ask you to answer some questions about your wellbeing. Over the last 2 weeks, how often have you been bothered by the following problems? Please select the option that you feel best applies to you:*

PHQ-9 (Kroenke et al., 2001):

	Not at all	Several days	More than half the days	Nearly every day
Little interest or pleasure in doing things	0	1	2	3
Feeling down, depressed, or hopeless	0	1	2	3
Trouble falling or staying asleep, or sleeping too much	0	1	2	3
Feeling tired or having little energy	0	1	2	3
Poor appetite or overeating	0	1	2	3
Feeling bad about yourself – or that you are a failure or have let yourself or your family down	0	1	2	3
Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
Moving or speaking so slowly that other people could have noticed – or the opposite: being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
Thoughts that you would be better off dead or of hurting yourself in some way	0	1	2	3

GAD-7 (Spitzer et al., 2006):

	Not at all	Several days	More than half the days	Nearly every day
Feeling nervous, anxious, or on edge	0	1	2	3
Not being able to stop or control worrying	0	1	2	3
Worrying too much about different things	0	1	2	3
Trouble relaxing	0	1	2	3
Being so restless that it's hard to sit still	0	1	2	3
Becoming easily annoyed or irritable	0	1	2	3
Feeling afraid as if something awful might happen	0	1	2	3

Beliefs About Emotions Scale (BAES; Rimes & Chalder, 2010):

*Below are a number of statements about emotions. Please indicate how much you agree with each one.*

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
If I lose control of my emotions in front of others, they will think less of me.					
I should be able to control my emotions.					
It is a sign of weakness if I have miserable thoughts.					
If I let myself get depressed, it shows that I am weak.					
It is wrong to show my feelings.					
I should be able to cope with negative emotions on my own					
People will reject me if I become too emotional.					
Getting upset is a sign of failure.					
I should be able to suppress my unpleasant emotions.					

It is embarrassing to experience strong emotions.					
I should not let myself give in to negative feelings.					
If I show how I feel, people will think less of me.					

## Appendix 2.8 - Link to Syntax File

<https://osf.io/mqvwd/files/mnj6h>

### **Appendix 2.9 - Data Availability Statement**

Data for this study was not made publicly available as this was not included in the ethics proposal.

Data will be transferred to the author's University of Glasgow supervisor for use in future studies.