



Making it Short: Shortening the Comprehensive Inventory of Mindfulness Experiences Using Ant Colony Optimization

Johannes A. Karl^{1,2} · Leticia Ribeiro³ · Claudia Bergomi⁴ · Ronald Fischer^{2,5} · Simon Dunne¹ · Oleg N. Medvedev⁶

Accepted: 6 January 2024 / Published online: 6 February 2024
© The Author(s) 2024

Abstract

Objectives The Comprehensive Inventory of Mindfulness Experiences (CHIME) comprises 37 items over eight domains and was validated in German and English languages using Rasch methodology. However, the length of the scale might limit its application due to the response burden it places on participants, especially in studies that examine additional constructs. This study aimed to develop a short but comprehensive CHIME version using cutting-edge methodology.

Method Ant colony optimization was applied to optimize and reduce item content of individual CHIME facets and create 24-item (CHIME-S) and 16-item (CHIME-XS) versions. We used data collected in New Zealand ($n=512$), and in the USA, which included two independent samples ($n=605$ and $n=210$). The construct validity of the newly developed CHIME-S and CHIME-XS was established by correlating its scores with measures of mindfulness, affect, and distress.

Results Overall, the ant-colony algorithm identified a stable solution which included 3 best fitting items per facet CHIME short form (CHIME-S) and the 2 best-fitting items CHIME-XS (16 items). This solution was successfully replicated across all samples and the scale demonstrated acceptable confirmatory factor analysis fit and good internal consistency, and correlated with measures of mindfulness, distress, and affect in expected directions.

Conclusions The CHIME-S and CHIME-XS are reliable and valid short-scale versions that can be used for assessment of total mindfulness and its facets in a comprehensive way. Future research efforts are invited to validate the CHIME-S and CHIME-XS across different cultures and sample populations, aiming at establishing robustness of the measure.

Keywords Mindfulness · Comprehensive Inventory of Mindfulness Experiences · CHIME · Ant colony optimization · Validation · Assessment · Psychometrics

Mindfulness is commonly defined as the awareness that arises from intentionally paying attention to the present moment in a non-judgmental manner (Kabat-Zinn, 2003).

Claudia Bergomi is an independent researcher.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12671-024-02302-z>.

✉ Johannes A. Karl
johannes.karl@dcu.ie
Leticia Ribeiro
lribeiro@iastate.edu
Claudia Bergomi
claudiab@protonmail.ch
Ronald Fischer
ronald.fischer@idor.org
Simon Dunne
simon.dunne@dcu.ie

It has been shown to alleviate various psychological and physical symptoms associated with depression, anxiety, stress, and other psychopathologies (Krägeloh et al., 2019). Furthermore, mindfulness-based interventions (MBIs) have demonstrated positive effects on well-being (Bennett

Oleg N. Medvedev
oleg.medvedev@waikato.ac.nz

- ¹ School of Psychology, Dublin City University, Dublin, 9, Ireland
- ² School of Psychology, Victoria University of Wellington, Wellington, New Zealand
- ³ Iowa State University, Student Counselling Center, Ames, IA, USA
- ⁴ Bern, Switzerland
- ⁵ Cognitive Neuroscience and Neuroinformatics Unit, Institute D'Or for Research and Teaching, Rua Diniz Cordeiro, 30 - Botafogo, Rio de Janeiro - RJ 22281-100, Brazil
- ⁶ School of Psychology, University of Waikato, Hamilton, New Zealand

& Dorjee, 2016) and emotional regulation (de Vibe et al., 2018). Given the growing importance of MBIs, researchers and clinicians require accurate and comprehensive measures of mindfulness to discriminate between individual levels of mindfulness. Scale development and validation are essential steps in creating reliable and valid tools for assessment.

Comprising 37 items, the CHIME offers a multifaceted understanding through its eight subscales. The Awareness of Internal Experiences subscale gauges an individual's attentiveness to their inner emotional, cognitive, and bodily sensations, reflecting the introspective facet of mindfulness. The Awareness of External Experiences component focuses on an individual's level of awareness and attention to external stimuli, capturing mindfulness' extrospective dimension. Third, the Acting with Awareness subscale assesses the conscious presence and attention in one's actions, which underscores the importance of being present in the moment and engaging with tasks deliberately. Fourth, the Accepting Non-judgmental Attitude dimension emphasizes the unconditional acceptance of experiences without evaluating or labeling them, epitomizing the non-critical nature of mindfulness. Fifth, the Nonreactive Decentering component pertains to the ability to observe one's thoughts and feelings without getting entangled in them, signifying the detached observation characteristic of mindfulness. Sixth, the Openness to Experience subscale measures an individual's willingness to engage with and accept a wide range of experiences, which reflects the open-hearted quality of mindfulness. Seventh, the Awareness of Thoughts' Relativity subscale captures the recognition that thoughts are transient and not absolute truths, underlining the discerning aspect of mindfulness. Finally, the Insightful Understanding component delves into the deeper realizations and insights derived from mindful practices, an important aspect that taps into the profound transformative potential of mindfulness. Each of these subscales, as detailed by Bergomi et al. (2014), intricately captures distinct yet interrelated facets of mindfulness, ensuring a comprehensive assessment of an individual's mindfulness experiences.

We recognize and appreciate the rich historical and spiritual roots of mindfulness, especially its foundational significance in Buddhism. It is crucial to acknowledge that while mindfulness has been embraced within psychology both as a therapeutic intervention and a means to enhance well-being, this adaptation often leans towards a secular perspective, distanced from its spiritual origins. The intent behind such secular applications is to make mindfulness accessible and beneficial to a broader audience, irrespective of their spiritual or religious affiliations. However, we are cognizant of the ongoing discourse regarding the authenticity and completeness of defining mindfulness purely in secular, psychological terms, as highlighted by studies like Van Gordon et al. (2015). While acknowledging this limitation,

it is noteworthy that among the CHIME domains, "Insightful Understanding" does come closer to capturing the spiritual essence of mindfulness, setting it apart from many purely secular scales. We believe that the inclusiveness of such a domain provides a bridge, however modest, to the profound spiritual depths of mindfulness, while still catering to its broader, secular applications. While the CHIME is an effective measure rooted in theoretical frameworks, its length may limit its applicability in large-scale studies involving numerous variables, where shorter scales are necessary for valid and/or complete responses. Therefore, developing a concise yet comprehensive version of CHIME is important.

Currently, the Five Facet Mindfulness Questionnaire (FFMQ; Baer et al., 2006) stands as the predominant multifaceted tool for evaluating mindfulness. The FFMQ's development involved using factor analysis on the Kentucky Inventory of Mindfulness Skills (KIMS; Baer et al., 2004), Mindful Attention and Awareness Scale (MAAS; Brown & Ryan, 2003), and other existing mindfulness measures. It encompasses five subscales: Describing, Observing, Non-Judging, Non-Reacting to Inner Experience, and Acting with Awareness. However, a study by Bergomi et al. (2013) demonstrated that the FFMQ, along with other mindfulness assessments, did not sufficiently cover all relevant aspects of mindfulness. To tackle this limitation and devise a more comprehensive tool, the CHIME was created (Bergomi et al., 2014). The CHIME integrates all the mindfulness elements emphasized by Bergomi et al. (2013) and is grounded in pertinent theoretical frameworks (Krägeloh et al., 2019). It consists of eight subscales, as described above.

The CHIME scale possesses two primary advantages over other mindfulness measures. First, it was designed with a solid theoretical foundation in traditional mindfulness conceptualizations (Bergomi et al., 2014; Krägeloh et al., 2019). This differs from the FFMQ, which was devised through factor analysis of existing mindfulness measures (Baer et al., 2006). As a result, the majority of FFMQ items originate from the MAAS and KIMS, inheriting their inherent methodological flaws. Specifically, the MAAS has faced substantial critique in the mindfulness literature for measuring mindlessness/inattention rather than mindfulness (Bergomi et al., 2013). Furthermore, the KIMS was developed based on the mindfulness conceptualization found in Dialectical Behavior Therapy (DBT; Linehan, 1993), an intervention primarily used to address symptoms of borderline personality disorder. In response to these concerns, the CHIME was designed using a mindfulness conceptualization derived from Eastern spiritual traditions.

Secondly, the CHIME encompasses a broader array of characteristics of the practice of mindfulness that are frequently misrepresented in other measures, such as awareness of internal experiences, openness to experience, awareness of thoughts' relativity, and insightful

understanding (Bergomi et al., 2014). Most mindfulness measures focus on broad concepts of practice, including awareness, attention, and non-judgmental attitude factors. The CHIME offers a distinctive approach to assessing mindfulness by delving deeper into nuances that are often not explicitly highlighted in other instruments. While many mindfulness measures predominantly focus on overarching constructs such as awareness, attention, and non-judgmental attitudes, the CHIME introduces facets that provide a richer understanding of the practice. For instance, the CHIME uniquely emphasizes components like *Awareness of thoughts' relativity* and *Insightful understanding*, both of which resonate with traditional mindfulness teachings from Eastern traditions.

Drawing from foundational texts and traditional interpretations (Harvey, 2013), mindfulness is not just about being aware or non-judgmental; it is also a tool for understanding the true nature of reality, ultimately aiming to alleviate suffering and enhance well-being. It is from this perspective that the inclusion of “wisdom factors” becomes pivotal. These wisdom elements, which encompass insights into the impermanent and interconnected nature of existence, are integral to a comprehensive understanding of mindfulness. CHIME's *Insightful understanding* and *Awareness of thoughts' relativity* can be viewed as operationalizations of these wisdom factors, bridging conventional measures and traditional conceptualizations.

Initially created in German, the CHIME has undergone validation through both conventional techniques and Rasch analysis (Bergomi et al., 2014; Medvedev et al., 2019), verifying its outstanding psychometric qualities. The CHIME boasts high internal consistency (α ranging from 0.70 to 0.90) and temporal reliability (test–retest reliability across 7 to 9 weeks, with r ranging from 0.70 to 0.90). Moreover, CHIME items displayed measurement invariance over time, signifying no significant differences in participants' comprehension of the items at different time points (Krägeloh et al., 2018). The scale's external validity was established through a strong positive correlation with the FFMQ ($r=0.85$) and moderate negative correlations with depression ($r=-0.46$), anxiety ($r=-0.39$), and stress ($r=-0.40$) (Bergomi et al., 2014). Moreover, a Dutch version of the CHIME, along with a shortened Dutch edition, has been developed and validated using classical test theory approaches (Cladder-Micus et al., 2019). The English version of the CHIME, recently validated using Rasch methodology by Wilkinson et al. (2023), provides a solid foundation for further adaptation. However, to maximize the measure's use in both research and clinical settings around the world, a shorter version of the CHIME in English is necessary. Building strong and reliable instruments relies on applying appropriate and rigorous psychometric methods. To achieve that goal, we used ant colony optimization with confirmatory factor analysis in our

validation of the shorter version of the CHIME (Dorigo & Stützle, 2004; Olaru et al., 2019).

It is indisputable that the field of mindfulness measurement has grown substantially over the years, with several short forms emerging to suit diverse research and clinical needs. For instance, the shortened versions of the FFMQ (Bohlmeijer et al., 2011) and the already concise MAAS offer insights into different facets of mindfulness. Moreover, the study that selected a 24-item short form from an initial pool of 173 items using an ant colony optimization approach (Altgassen et al., 2023), encompassing various mindfulness questionnaires including CHIME, presents a commendable approach to synthesizing mindfulness metrics across theoretical approaches. Nevertheless, while these measures offer significant advantages, a CHIME short version would distinguish itself through its comprehensive coverage and clear theoretical alignment. The CHIME encapsulates all the components of mindfulness as illuminated by Bergomi et al. (2013), deeply rooted within pertinent theoretical frameworks (Krägeloh et al., 2019). Comprising eight diverse subscales, the CHIME offers a holistic evaluation of mindfulness, ranging from awareness of internal and external experiences, acting with awareness, accepting nonjudgmental attitude, nonreactive decentering, openness to experience, awareness of thoughts' relativity, and extending to insightful understanding. Since its inception, the CHIME has consistently demonstrated exemplary psychometric properties (Bergomi et al., 2014; Wilkinson et al., 2023).

While the existing version of the CHIME shows good psychometric properties and is clearly aligned with a broader theory, its use might be limited by its length. To address this, we are using an ant colony optimization approach on the English version of the CHIME to create two short forms of the CHIME with 24 (CHIME-S) and 16 (CHIME-XS) items respectively, comparing their psychometric properties against the full CHIME and non-English short versions of the CHIME. Ant colony optimization is an advanced metaheuristic machine-learning technique inspired by the foraging behavior of ants (Dorigo et al., 1996). Collective foraging of ants involves multiple ants exploring an open space for food. Successful tracks are marked via pheromones and as an increasing number of ants is using the most efficient path, the pheromone levels become stronger until only the shortest or most efficient route is being used. Inspired by this biological model, the method uses simulated agents (artificial ants) to search for optimal model solutions in a network graph and over multiple interactions, the shortest paths encountered are being given greater weights (similar to pheromone tracks by ants). This method has been widely applied for solving complex computational problems and can be applied to the question of optimizing and reducing scales by combining the ant colony optimization algorithm with confirmatory factor analysis to search for the optimal

solution within defined parameters (Dorigo & Stützle, 2004; Olaru et al., 2019). In essence, the algorithm iteratively improves the quality of selected items, ultimately retaining those that best represent the underlying construct (in this case, mindfulness) and simultaneously optimize associations with third variables. Ant colony optimization offers an innovative and efficient approach to simplifying complex measures while maintaining validity and reliability (Blum & Roli, 2003; Olaru et al., 2019).

One of the main benefits of using ant colony optimization in scale development is its ability to overcome the limitations of traditional methods, which often rely on subjective decisions around a limited set of criteria such as factor loadings of individual items or the absence of cross loadings and may lead to suboptimal solutions (Schroeders et al., 2016). For instance, conventional methods like factor loadings are prone to researcher bias, as decisions regarding item retention or deletion are often based on arbitrary cutoff values or guidelines. To provide an example from the mindfulness literature in the development of the FFMQ, differing cutoffs are applied for factor loadings depending on sub-scale examined (0.40 for Non-Reactivity and 0.50 for Observing). Furthermore, traditional factor analytic approaches often only a single exploratory run with one factor solution is evaluated and, subsequently, items for validation are selected. This may result in the exclusion of potentially important items or the retention of less relevant ones as the factor structure and solution is likely to change with the inclusion and exclusion of different items, thereby compromising the content validity and psychometric properties of the shortened scale (Matsunaga, 2010). Beyond internal validity, construct validity (e.g., correlation with criterion variables) is a crucial element of scale validation and is often of primary concern in applied settings.

Ant colony optimization provides an objective, data-driven approach to item selection, ensuring that the resulting scale is both psychometrically sound and theoretically grounded. The underlying approach is inherently robust and adaptive because it was developed to efficiently explore complex search spaces and identify optimal solutions that might be overlooked by traditional methods (Dorigo et al., 1996). By harnessing the collective intelligence of the simulated ant colony, the algorithm balances exploration and exploitation, effectively avoiding local optima and converging towards a global optimum solution. Consequently, the use of ant colony optimization in scale development can lead to the identification of more accurate, reliable, and parsimonious measurement instruments that are better suited to assess the intended constructs such as mindfulness (Olaru & Danner, 2021; Olaru et al., 2019; Schroeders et al., 2016).

The primary aim of the present study was to develop a shorter version of the CHIME using ant colony optimization, ensuring the resulting measure is comprehensive and

psychometrically sound. In the current study, we utilized three samples from anglophone populations to comprehensively examine the robustness of the derived CHIME versions across student and general populations. We used internal validity criteria (model fit) as optimization criteria and construct validity criteria (correlation with other established mindfulness and subjective well-being measures) as additional validation criteria. By harnessing the advantages of this cutting-edge methodology, we expected to create an efficient and practical tool that can be widely used in research and clinical settings, contributing to the field of mindfulness assessment and promoting the understanding and practice of mindfulness in various contexts.

Method

Participants

We collected responses from 512 undergraduate students in psychology at a New Zealand University using an online questionnaire through Qualtrics between January 2020 and December 2020, which took 15 min to complete on average. Participants consented to take part in the study in exchange for research credit and could participate only once. The study obtained approval from the authors' institutional ethics review board, and all participants who took part in the study gave their informed consent for their involvement in the research. The participants were on average 19.14 years old ($SD = 3.34$) and majority female (74.66%).

Data for the first US sample was collected in March 2018, via an online questionnaire hosted on Qualtrics Research Services, ensuring an equal distribution between male and female respondents. By referring to the time taken by five volunteers, we gauged that an average response time was around 15 min. This online collection process spanned about 10 days. For their efforts, respondents received US\$5 upon completion. To ensure no repeated submissions, we logged each respondent's IP address. The wide distribution of IP addresses indicated a comprehensive national reach, assuring a well-represented sample. The study secured approval from the institution's ethics review board, and all respondents willingly gave their consent before participating. The participants were on average 41.68 years old ($SD = 13.04$) and were nearly balanced in gender (51.07% female).

The second US dataset was collected from students taking part in an introductory university course in exchange for research credit. The studies obtained approval from the authors' institutional ethics review board, and all participants gave their informed consent prior to participating. The participants were on average 18.64 years old ($SD = 1.25$), majority female (75.24%), and identified largely as White (88.10%) followed by Asian (4.76%) and Black (4.29%).

Measures

Comprehensive Inventory of Mindfulness Experiences

We administered the 37-item CHIME (Wilkinson et al., 2023) which allows participants to rate themselves on a 6-point Likert scale with greater scores indicating higher mindfulness (1 = *Almost never* to 6 = *Almost always*). The original 37 items are distributed across eight facets: *Awareness of internal experiences* (“When my mood changes, I notice it right away.”), *Awareness of external experiences* (“I notice details in nature, such as colors, shapes, and textures.”), *Acting with Awareness* (“I break or spill things because I am not paying attention or I am thinking of something else.”), *Accepting nonjudgmental attitude* (“In the ups and downs of life, I am kind to myself.”), *Nonreactive decentering* (“When I have distressing thoughts or images, I am able to feel calm soon afterward.”), *Openness to experience* (“I try to stay busy to avoid specific thoughts or feelings from coming to mind.”), *Awareness of thoughts’ relativity* (“It is clear to me that my evaluations of situations and people can change easily.”), and *Insightful understanding* (“In everyday life, I notice when my negative attitudes toward a situation make things worse.”).

Convergent Validity Measure Collected in US Sample 2

Five Facet Mindfulness Questionnaire (FFMQ)

The FFMQ (Baer et al., 2006) contains 39 items, measuring five separate facets: Describing, Observing, Non-Judging, Non-Reacting to Inner Experience, and Acting with Awareness. Answers are scored on a 5-point Likert scale with greater sum scores indicating greater mindfulness, from *never* = 1 to *always true* = 5. The measure contains 19 items that are negatively worded (3, 5, 8, 10, 12, 13, 14, 16, 17, 18, 22, 23, 25, 28, 30, 34, 35, 38, and 39), which require reverse coding before the total and subscale scores can be calculated. All subscales showed good ω_{Total} reliability: Describing (0.91, 95% CI[0.89, 0.93], $M = 26.90$, $SD = 5.65$, $\text{Min} = 12$, $\text{Max} = 40$), Observing (0.79, 95% CI[0.75, 0.84], $M = 25.29$, $SD = 5.14$, $\text{Min} = 8$, $\text{Max} = 37$), Non-Judging (0.93, 95% CI[0.91, 0.94], $M = 25.78$, $SD = 6.64$, $\text{Min} = 8$, $\text{Max} = 40$), Non-Reacting to Inner Experience (0.81, 95% CI[0.77, 0.85], $M = 20.97$, $SD = 4.12$, $\text{Min} = 7$, $\text{Max} = 31$), and Acting with Awareness (0.88, 95% CI[0.86, 0.91], $M = 26.11$, $SD = 5.49$, $\text{Min} = 9$, $\text{Max} = 40$).

Depression Anxiety Stress Scales (DASS)

The DASS (Lovibond & Lovibond, 1995) is a self-administered questionnaire evaluating the negative emotional

aspects of depression, anxiety, and stress. Comprising 42 items, the DASS features 14 items per subscale. Items are scored on a 4-point Likert scale from *Did not apply to me at all* = 0 to *Applied to me very much or most of the time* = 3. A shorter 21-item version, the DASS-21 (Antony et al., 1998), is also accessible. Both variants demonstrate strong internal consistencies and reliable psychometric properties. For this study, the DASS-21 was employed resulting in a total score range for each scale from 0 to 21. All subscales showed good ω_{Total} reliability: Stress (0.88, 95% CI[0.86, 0.91], $M = 4.14$, $SD = 3.78$, $\text{Min} = 0$, $\text{Max} = 19$), Anxiety (0.88, 95% CI[0.86, 0.91], $M = 2.96$, $SD = 3.27$, $\text{Min} = 0$, $\text{Max} = 16$), Depression (0.92, 95% CI[0.90, 0.94], $M = 3.18$, $SD = 3.80$, $\text{Min} = 0$, $\text{Max} = 18$).

Positive and Negative Affect Schedule (PANAS)

The PANAS (Watson et al., 1988) is a concise list of adjectives describing various emotions and feelings. Participants are asked to rate the extent to which they experienced these emotions/feelings over the past week using a 5-point Likert scale, with responses ranging from *Very slightly or not at all* = 1 to *Extremely* = 5. After completing the questionnaire, scores from positive emotion adjectives are combined to form the Positive Activation scale, while scores from negative emotion adjectives create the Negative Activation scale. The ω_{Total} reliability was high for both scales: Positive Activation (0.89, 95% CI[0.87, 0.91], $M = 21.93$, $SD = 6.88$, $\text{Min} = 10$, $\text{Max} = 42$), Negative Activation (0.92, 95% CI[0.90, 0.94], $M = 17.06$, $SD = 6.81$, $\text{Min} = 10$, $\text{Max} = 46$).

Beck Anxiety Inventory (BAI)

The BAI (Beck et al., 1988) is a self-report instrument designed to assess the severity of anxiety symptoms. The BAI comprises 21 items, each describing a common symptom of anxiety. Participants are asked to rate how much they have been bothered by each symptom over the past week on a 4-point Likert scale, ranging from *Not at all* = 0 to *Severely, I could barely stand it* = 3. The total score, ranging from 0 to 63, is obtained by summing the individual item scores, with higher scores indicating greater anxiety severity. The BAI has demonstrated good internal consistency, test–retest temporal reliability, and construct validity in various populations and showed good ω_{Total} reliability in the current sample (0.95, 95% CI[0.94, 0.96], $M = 13.06$, $SD = 10.17$, $\text{Min} = 0$, $\text{Max} = 59$).

Penn State Worry Questionnaire (PSWQ)

The PSWQ (Meyer et al., 1990) is a self-administered measure developed to evaluate the tendency to worry excessively. The PSWQ contains 16 items, and respondents are instructed

to rate the extent to which each statement is characteristic of their usual worry style on a 5-point Likert scale, ranging from *Not at all typical of me* = 1 to *Very typical of me* = 5. The total score is calculated, ranging from 16 to 80, by summing the item ratings, with higher scores indicating a higher propensity for worry. The PSWQ has shown strong internal consistency, test–retest temporal reliability, and convergent and discriminant validity in multiple studies and showed good ω_{Total} reliability in the current sample (0.95, 95% CI[0.95, 0.96], $M = 54.09$, $SD = 14.46$, $\text{Min} = 19$, $\text{Max} = 80$).

Center for Epidemiological Studies-Depression (CES-D)

The CES-D (Lewinsohn et al., 1997) is a widely used tool for measuring depressive symptoms in the general population. The CES-D assesses the frequency of depressive symptoms experienced over the past week and consists of 20 items that are designed to represent major symptoms of depression measured on a 4-point Likert scale ranging from 0 = *Rarely or none of the time (less than 1 day)* to 3 = *Most or all of the time (5–7 days)*. The scale is scored on a range of 0 to 60, with higher scores indicating a greater likelihood of depression. The CES-D also includes four subscales, each of which measures a different aspect of depression. The Positive Activation subscale measures the frequency of positive emotions such as happiness and joy, while the negative activation subscale measures the frequency of negative emotions such as sadness and guilt. The somatic symptoms and retarded activity subscales measure the extent to which depression interferes with physical functioning and daily activities, while the interpersonal difficulties subscale measures the extent to which depression affects social relationships. In the current sample, the total score had good ω_{Total} reliability (0.92, 95% CI[0.91, 0.94], $M = 14.85$, $SD = 9.48$, $\text{Min} = 0$, $\text{Max} = 44$).

Data Analyses

We initially used the New Zealand dataset, running an ant colony optimization using 20 ants (which represents distinct short-form configurations of the CHIME), 5% evaporation, a stability of 2000 runs, and a maximum of 20,000 runs (based on recent recommendations Leite et al., 2008; Raborn & Leite, 2018), to estimate the best stable CHIME version specified to provide a final model with 3 items per facet as this allows for a just-identified model should a researcher be interested in only administering a single factor. For researchers interested in applying the whole scale, we derived an extra short version of the CHIME with only 2 items per factor. We used the CFI (Comparative Fit Index), SRMR (Standardized Root Mean Square Residual), and RMSEA (Root Mean Square Error of Approximation) fit based on a Confirmatory Factor Analysis model

as criteria for optimization. We modelled the facets of the CHIME as correlated facets without a higher order factor, using ordinal Weighted Least Squares Means and Variances (WLSMV) estimation. We defined good fit as $\text{CFI} > 0.95$, $\text{RMSEA} < 0.06$, and $\text{SRMR} < 0.08$ (Hu & Bentler, 1999). To examine differences between scales, we examined the Δ_{CFI} between two scales with 0.01 defined as a substantial difference in fit (Fischer & Karl, 2019). We have used 3 decimal points to report the relevant fit indices in the results section for the purpose of precision. Subsequently, we fitted the resulting short model and the extra short model to our first US dataset using a WLSMV estimated Confirmatory Factor Analysis method and comparing it against the original CHIME longform and the alternative short form of the CHIME developed by Cladder-Micus et al. (2019) in Dutch. We repeated this analysis in the second independent US dataset, in which then also examined the correlation of the short vs long forms of the CHIME as a criterion for external validity. We used R4.30 for all analyses.

Results

NZ Sample

The ant colony algorithm converged after 164 runs representing 3280 individual ant runs for the 24-item short form and after 170 runs representing 3400 individual ant runs for the 16-item extra short form at which point the model showed no more improvement in mean γ (standardized latent variable loadings). We show the improvement in total pheromone level across the runs in Fig. S1 in the supplementary material and the final selected items for the short and extra short CHIME in Table 1. Examining the fit of all scale forms, we found good fit for all scale forms (Table 2), but found that both the CHIME-S ($\Delta_{\text{CFI-Original}} = 0.05$, $\Delta_{\text{CFI-Dutch-Short}} = 0.05$) and CHIME-XS ($\Delta_{\text{CFI-Original}} = 0.03$, $\Delta_{\text{CFI-Dutch-Short}} = 0.03$) showed improved measurement characteristics above the original CHIME and the previously short version derived in Dutch. Similarly, both the CHIME-S and CHIME-XS showed comparable reliability to the original and previous short form, with the notable exception of the CHIME-XS Openness to experience facet which showed low reliability (Table 3). This facet showed the lowest reliability across all scale forms, which might be exacerbated by the 2-item solution. Similarly, the CHIME-S and CHIME-XS showed substantial correlations with the CHIME long form (as can be seen in Table 4). Finally, the CHIME-S and CHIME-XS showed a very high similarity of their facet intercorrelations compared to the facet intercorrelation in the original scale (CHIME-S $r_{\text{Mantel}} = 0.94$, $p < 0.001$; CHIME-XS $r_{\text{Mantel}} = 0.92$, $p < 0.001$) indicating a high comparability of facet intercorrelations.

Table 1 Items selected for the short forms (CHIME-S and CHIME-XS) by the ant colony algorithm

Awareness of internal experiences

***When my mood changes, I notice it right away.**

When I am sitting or lying down, I notice the sensations in my body.

***†When I talk to other people, I notice what feelings I am experiencing.**

I clearly notice changes in my body, such as quicker or slower breathing.

***†I am usually aware of how I am feeling at any given time.**

Awareness of external experiences

***I notice details in nature, such as colors, shapes, and textures.**

When I ride in a car or train, I am aware of the surroundings, such as the landscape.

***†I pay attention to sensations, such as the wind in my hair or sunshine on my face.**

***†I notice sounds in my environment, such as birds chirping or cars passing.**

Acting with awareness

I break or spill things because I am not paying attention or I am thinking of something else.

***It is easy for me to stay focused on what I am doing.**

***†In everyday life, I get distracted by many memories, images, or daydreams.**

***†When I read, I have to reread paragraphs because I was thinking of something else.**

Accepting nonjudgmental attitude

***†In the ups and downs of life, I am kind to myself.**

I am hard on myself when I make a mistake.

***I see my mistakes and difficulties without judging myself.**

***†Even when I make a big mistake, I treat myself with kindness and understanding.**

I resent my own mistakes and weaknesses.

Nonreactive decentering

When I have distressing thoughts or images, I am able to feel calm soon afterward.

When I have distressing thoughts or images, I am able to notice them without having to react.

In difficult situations, I can pause for a moment without reacting immediately.

***†When caught in thoughts and emotions, I am able to “step back” and quickly notice the thought or image without getting taken over by it.**

***†I am able to notice my thoughts and feelings without getting tangled up in them.**

***I notice my thoughts and feelings and can also “step back” and observe them from a distance.**

Openness to experience

†I try to stay busy to avoid specific thoughts or feelings from coming to mind.

***I try to distract myself when I feel unpleasant emotions.**

***†I do not like it when I am angry or fearful and I try to get rid of these feelings.**

***When I am in pain, I try to avoid this sensation as much as possible.**

Awareness of thoughts’ relativity

It is clear to me that my evaluations of situations and people can change easily.

***†In everyday life, I realize my thoughts are not always facts.**

***†In everyday life, I am aware that my view on things is not always based on facts.**

***I am aware that even my strongly held opinions may change over time.**

Insightful understanding

In everyday life, I notice when my negative attitudes toward a situation make things worse.

***†I am able to smile when I notice myself seeing things as more complicated than they actually are.**

***When I have needlessly given myself a hard time, I can see it with humor.**

***†I am able to smile to myself when I notice I have made a big deal out of a small problem.**

I am able to notice when I needlessly make life more difficult for myself.

Items in bold and marked with * were selected for the short-form; items marked with † were selected for the extra short form

Table 2 Comparison between all confirmatory factor analysis models in all samples

Scale	Parameters	χ^2	df	CFI	RMSEA	RMSEA lower	RMSEA upper	SRMR	$\hat{\gamma}$
New Zealand									
Original	102	1442.36	601	0.93	0.07	0.06	0.07	0.07	0.91
CHIME-S	76	441.98	224	0.98	0.04	0.04	0.05	0.05	0.96
CHIME-XS	60	288.1	76	0.96	0.06	0.06	0.07	0.06	0.95
CHIME-S-Dutch	76	690.97	224	0.93	0.07	0.06	0.07	0.07	0.93
USA Sample 1									
Original	102	2848.39	601	0.94	0.08	0.08	0.09	0.10	0.83
CHIME-S	76	963.04	224	0.97	0.06	0.05	0.06	0.07	0.91
CHIME-XS	60	152.62	76	1.00	0.03	0.02	0.03	0.03	0.98
CHIME-S-Dutch	76	1515.12	224	0.94	0.09	0.08	0.09	0.09	0.85
USA Sample 2									
Original	102	961.74	601	0.89	0.08	0.07	0.09	0.10	0.91
CHIME-S	76	351.74	224	0.98	0.06	0.04	0.07	0.07	0.95
CHIME-XS	60	161.03	76	0.91	0.07	0.05	0.08	0.07	0.95
CHIME-S-Dutch	76	405.89	224	0.87	0.07	0.06	0.08	0.09	0.93

All models were fitted with a WLSMV estimator to account for the ordinality of the scale. The CHIME-S-Dutch is based on the 24-item model developed by Cladder-Micus et al., (2019) in a Dutch population

Table 3 Reliability of scale forms across samples

Facet	Short	Extra-short	Dutch-Short	Original
NZ				
<i>Awareness of internal experiences</i>	0.70[0.65, 0.74]	0.62	0.61[0.55, 0.67]	0.73[0.69, 0.77]
<i>Awareness of external experiences</i>	0.82[0.79, 0.85]	0.73	0.82[0.79, 0.85]	0.83[0.80, 0.85]
<i>Acting with Awareness</i>	0.72[0.68, 0.76]	0.66	0.72[0.68, 0.76]	0.72[0.68, 0.76]
<i>Accepting nonjudgmental attitude</i>	0.84[0.82, 0.87]	0.78	0.74[0.70, 0.78]	0.86[0.84, 0.88]
<i>Nonreactive decentering</i>	0.86[0.84, 0.88]	0.78	0.80[0.78, 0.83]	0.87[0.86, 0.89]
<i>Openness to experience</i>	0.59[0.52, 0.65]	0.29	0.48[0.40, 0.56]	0.63[0.58, 0.68]
<i>Awareness of thoughts' relativity</i>	0.79[0.76, 0.82]	0.75	0.79[0.76, 0.82]	0.76[0.72, 0.79]
<i>Insightful understanding</i>	0.74[0.70, 0.78]	0.69	0.74[0.70, 0.78]	0.73[0.70, 0.77]
US1				
<i>Awareness of internal experiences</i>	0.81 [0.78, 0.84]	0.746	0.77 [0.73, 0.80]	0.85 [0.83, 0.87]
<i>Awareness of external experiences</i>	0.90 [0.88, 0.91]	0.827	0.90 [0.88, 0.91]	0.90 [0.89, 0.92]
<i>Acting with awareness</i>	0.63 [0.58, 0.69]	0.761	0.63 [0.58, 0.69]	0.72 [0.69, 0.76]
<i>Accepting nonjudgmental attitude</i>	0.85 [0.83, 0.87]	0.795	0.63 [0.57, 0.69]	0.72 [0.69, 0.75]
<i>Nonreactive decentering</i>	0.88 [0.87, 0.90]	0.846	0.89 [0.87, 0.90]	0.93 [0.92, 0.94]
<i>Openness to experience</i>	0.74 [0.71, 0.78]	0.608	0.71 [0.67, 0.75]	0.80 [0.77, 0.82]
<i>Awareness of thoughts' relativity</i>	0.82 [0.80, 0.85]	0.745	0.82 [0.80, 0.85]	0.83 [0.80, 0.85]
<i>Insightful understanding</i>	0.85 [0.83, 0.87]	0.782	0.85 [0.83, 0.87]	0.85 [0.84, 0.87]
US2				
<i>Awareness of internal experiences</i>	0.71 [0.64, 0.78]	0.58	0.63 [0.55, 0.72]	0.72 [0.66, 0.78]
<i>Awareness of external experiences</i>	0.84 [0.80, 0.88]	0.77	0.84 [0.80, 0.88]	0.83 [0.80, 0.87]
<i>acting with awareness</i>	0.77 [0.71, 0.82]	0.67	0.77 [0.71, 0.82]	0.74 [0.69, 0.80]
<i>Accepting nonjudgmental attitude</i>	0.80 [0.76, 0.85]	0.68	0.73 [0.67, 0.79]	0.82 [0.79, 0.86]
<i>Nonreactive decentering</i>	0.80 [0.75, 0.84]	0.58	0.69 [0.62, 0.76]	0.86 [0.84, 0.89]
<i>Openness to experience</i>	0.76 [0.70, 0.81]	0.46	0.66 [0.58, 0.74]	0.77 [0.72, 0.82]
<i>Awareness of thoughts' relativity</i>	0.56 [0.46, 0.66]	0.47	0.56 [0.46, 0.66]	0.65 [0.57, 0.73]
<i>Insightful understanding</i>	0.79 [0.75, 0.84]	0.70	0.79 [0.75, 0.84]	0.72 [0.66, 0.78]

All reliability estimates are based on $\omega_{\text{Total Ordinal}}$ with the exception of the extra short form which represents the Spearman-Brown coefficient

Table 4 Pearson's correlation between original CHIME and the shortened versions in all samples

Facet	CHIME-Dutch	CHIME-S	CHIME-XS
<i>NZ</i>			
<i>Awareness of internal experiences</i>	0.91	0.87	0.83
<i>Awareness of external experiences</i>	0.97	0.97	0.92
<i>Acting with Awareness</i>	0.93	0.93	0.86
<i>Accepting nonjudgmental attitude</i>	0.95	0.93	0.89
<i>Nonreactive decentering</i>	0.94	0.92	0.89
<i>Openness to experience</i>	0.95	0.93	0.85
<i>Awareness of thoughts' relativity</i>	0.95	0.95	0.88
<i>Insightful understanding</i>	0.92	0.92	0.87
<i>US1</i>			
<i>Awareness of internal experiences</i>	0.95	0.92	0.87
<i>Awareness of external experiences</i>	0.98	0.98	0.95
<i>Acting with Awareness</i>	0.94	0.94	0.89
<i>Accepting nonjudgmental attitude</i>	0.95	0.83	0.81
<i>Nonreactive decentering</i>	0.96	0.95	0.92
<i>Openness to experience</i>	0.97	0.96	0.92
<i>Awareness of thoughts' relativity</i>	0.97	0.97	0.91
<i>Insightful understanding</i>	0.94	0.94	0.90
<i>US2</i>			
<i>Awareness of internal experiences</i>	0.90	0.86	0.82
<i>Awareness of external experiences</i>	0.97	0.97	0.93
<i>Acting with Awareness</i>	0.94	0.94	0.89
<i>Accepting nonjudgmental attitude</i>	0.95	0.92	0.86
<i>Nonreactive decentering</i>	0.93	0.90	0.86
<i>Openness to experience</i>	0.97	0.94	0.89
<i>Awareness of thoughts' relativity</i>	0.94	0.94	0.83
<i>Insightful understanding</i>	0.93	0.93	0.86

All correlation coefficients were significant at $p < 0.001$

US Sample 1

To test the out-of-sample applicability of this solution, we fitted the long and ant colony shortened form in our US Sample 1. Overall, we found a pattern similar to our initial results in our New Zealand sample with both the CHIME-S ($\Delta_{\text{CFI-Original}} = 0.03$, $\Delta_{\text{CFI-Dutch-Short}} = 0.03$) and CHIME-XS ($\Delta_{\text{CFI-Original}} = 0.06$, $\Delta_{\text{CFI-Dutch-Short}} = 0.06$) showing improved measurement characteristics above the original CHIME and the previously short version derived in Dutch (Table 2), while showing comparable reliability (Table 3) and robust correlations with the long form of the CHIME (Table 4). Notably, while the CHIME-XS Openness to experience facet, it was above commonly acceptable cutoffs indicating that the reliability of this facet might be sample dependent. As in the New Zealand sample, the CHIME-S and CHIME-XS showed a very high similarity of their facet intercorrelations compared to the facet intercorrelation in the original scale (CHIME-S $r_{\text{Mantel}} = 0.92$, $p < 0.001$; CHIME-XS $r_{\text{Mantel}} = 0.91$, $p < 0.001$). Overall,

this shows that both solutions derived by the ant colony optimization algorithm perform robustly across samples.

US Sample 2

Replicating our results from the US Sample 1 with the US Sample 2, we found a similar pattern, with both the CHIME-S ($\Delta_{\text{CFI-Original}} = 0.08$, $\Delta_{\text{CFI-Dutch-Short}} = 0.10$) and CHIME-XS ($\Delta_{\text{CFI-Original}} = 0.02$, $\Delta_{\text{CFI-Dutch-Short}} = 0.04$) showing improved measurement characteristics above the original CHIME and the previously short version derived in Dutch (Table 2), while showing comparable reliability (Table 3) and robust correlations with the long form of the CHIME (Table 4).

We further examined post hoc differences in construct validity correlations, between the short and long forms of the CHIME, the FFMQ, the DASS, the PANAS, the BAI, the PSWQ, and the CESD. Overall, we found a highly similar pattern of correlations between the long and short forms and the respective construct validity variables (Table 5). To compute the statistical similarity of the matrices, we computed

Table 5 Correlation of the original CHIME, the CHIME-S, the CHIME-XS, and the Dutch-CHIME-S with validation variables

Observing	Describing	Acting with awareness	Non-judging	Non-reacting	Stress	Anxiety	Depression	Positive activation	Negative activation	Anxiety-BAI	Worry	Depression-CESD
Internal experience												
0.56***	0.34***	0.23***	0.07	0.11	-0.05	0.00	-0.02	0.12	-0.01	0.20**	0.06	0.04
0.37***	0.42***	0.26***	0.15*	0.13	-0.08	-0.05	-0.04	0.08	-0.04	0.06	-0.04	-0.03
0.38***	0.43***	0.27***	0.18**	0.20**	-0.12	-0.06	-0.10	0.09	-0.08	0.02	-0.05	-0.08
0.43***	0.39***	0.28***	0.15*	0.10	-0.07	-0.03	-0.05	0.09	-0.03	0.16*	0.05	0.00
External experience												
0.63***	0.28***	0.10	0.11	0.01	-0.04	0.00	-0.06	0.11	-0.06	0.17*	0.02	0.07
0.63***	0.26***	0.06	0.09	0.04	-0.04	0.00	-0.06	0.11	-0.08	0.16*	0.02	0.06
0.66***	0.26***	0.06	0.08	-0.01	-0.04	0.01	-0.04	0.13	-0.07	0.18**	0.05	0.10
0.63***	0.26***	0.06	0.09	0.04	-0.04	0.00	-0.06	0.11	-0.08	0.16*	0.02	0.06
Acting with awareness												
-0.12	0.30***	0.70***	0.33***	0.14*	-0.23***	-0.22**	-0.24***	0.11	-0.29***	-0.32***	-0.29***	-0.37***
-0.12	0.27***	0.71***	0.27***	0.13	-0.22**	-0.19**	-0.25***	0.15*	-0.27***	-0.29***	-0.29***	-0.35***
-0.17*	0.23***	0.64***	0.22**	0.12	-0.20**	-0.19**	-0.25***	0.15*	-0.28***	-0.31***	-0.33***	-0.35***
-0.12	0.27***	0.71***	0.27***	0.13	-0.22**	-0.19**	-0.25***	0.15*	-0.27***	-0.29***	-0.29***	-0.35***
Non-judging												
-0.14	0.31***	0.28***	0.60***	0.29***	-0.29***	-0.27***	-0.44***	0.13	-0.32***	-0.49***	-0.60***	-0.60***
-0.07	0.33***	0.28***	0.55***	0.35***	-0.28***	-0.24***	-0.42***	0.14*	-0.30***	-0.45***	-0.50***	-0.56***
-0.05	0.37***	0.31***	0.54***	0.32***	-0.28***	-0.26***	-0.41***	0.12	-0.29***	-0.43***	-0.41***	-0.56***
-0.08	0.34***	0.31***	0.60***	0.32***	-0.31***	-0.25***	-0.43***	0.15*	-0.32***	-0.46***	-0.56***	-0.59***
Non-reacting												
0.02	0.34***	0.28***	0.41***	0.68***	-0.29***	-0.27***	-0.41***	0.23***	-0.31***	-0.37***	-0.50***	-0.49***
0.00	0.34***	0.32***	0.43***	0.54***	-0.27***	-0.24***	-0.42***	0.20**	-0.29***	-0.31***	-0.41***	-0.46***
-0.02	0.31***	0.32***	0.43***	0.50***	-0.27***	-0.23***	-0.41***	0.20**	-0.27***	-0.32***	-0.44***	-0.49***
0.02	0.35***	0.26***	0.41***	0.63***	-0.30***	-0.26***	-0.41***	0.23***	-0.28***	-0.36***	-0.50***	-0.52***
Openness to experience												
-0.20**	0.10	0.08	0.29***	0.04	-0.13	-0.10	-0.17*	0.05	-0.19**	-0.29***	-0.40***	-0.30***
-0.16*	0.05	0.03	0.19**	0.00	-0.08	-0.03	-0.08	0.06	-0.12	-0.18**	-0.27***	-0.18*
-0.25***	0.06	0.09	0.34***	0.03	-0.10	-0.13	-0.19**	0.02	-0.19**	-0.31***	-0.42***	-0.34***
-0.19**	0.12	0.10	0.33***	0.04	-0.14*	-0.11	-0.20**	0.07	-0.18**	-0.31***	-0.42***	-0.33***
Awareness of thoughts/relativity												
0.22**	0.07	0.11	-0.05	0.09	-0.02	-0.01	-0.05	0.12	0.00	0.02	0.05	-0.05
0.22**	0.05	0.07	-0.08	0.12	-0.02	-0.03	-0.04	0.14*	0.00	0.01	0.07	-0.06
0.17*	-0.03	-0.07	-0.12	0.13	-0.01	-0.02	-0.04	0.14*	0.01	0.00	0.09	-0.06
0.22**	0.05	0.07	-0.08	0.12	-0.02	-0.03	-0.04	0.14*	0.00	0.01	0.07	-0.06
Insightful understanding												

Table 5 (continued)

Observing	Describing	Acting with awareness	Non-judging	Non-reacting	Stress	Anxiety	Depression	Positive activation	Negative activation	Anxiety-BAI	Worry	Depression-CESD
0.12	0.31***	0.26***	0.16*	0.21**	-0.17*	-0.10	-0.24***	0.23***	-0.19**	-0.17*	-0.24***	-0.27***
0.05	0.27***	0.26***	0.20**	0.17*	-0.19**	-0.13	-0.27***	0.20**	-0.22**	-0.24***	-0.29***	-0.31***
0.09	0.28***	0.27***	0.18**	0.13	-0.18**	-0.13	-0.29***	0.20**	-0.23***	-0.23***	-0.24***	-0.30***
0.05	0.27***	0.26***	0.20**	0.17*	-0.19**	-0.13	-0.27***	0.20**	-0.22**	-0.24***	-0.29***	-0.31***

Correlation coefficients within each cell represent (from top to bottom within each panel) the original configuration, the CHIME-S, the CHIME-XS, and the Dutch CHIME-S

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

the asymmetric correlation matrix between the construct validity variables and the long or short form of the CHIME respectively. Subsequently, we calculated a Mantel test to determine the overall similarity of the matrices based on 9999 permutations of the test. Overall, we found a very high similarity (CHIME-S $r_{\text{Mantel}} = 0.94$, $p < 0.001$; CHIME-XS $r_{\text{Mantel}} = 0.97$, $p < 0.001$), indicating the similarity in the patterns of relationships between the long and short forms. In summary, the results demonstrated that the shortened versions of the CHIME not only possess comparable or superior measurement properties to the long-form in most samples but also exhibit similar patterns in their relationships within the instrument and with construct validity variables.

Discussion

The present study aimed to create a comprehensive 24-item short-form (CHIME-S) and a 16-item extra short-form (CHIME-XS), of the CHIME by employing cutting-edge ant colony optimization methodology. The shortened scales exhibited comparable or improved confirmatory factor analysis fit indices compared to the full version of the CHIME and existing non-English short forms of the CHIME (Cladder-Micus et al., 2019), good internal reliability, and expected correlations with other measures of mindfulness, distress, and activation. The CHIME-S and CHIME-XS preserve the original CHIME's comprehensive nature while making it more accessible and applicable for large-scale studies, where shorter scales are preferred to ensure response validity and completion. This shorter version maintains the theoretical foundations of the original scale while offering an efficient and reliable assessment tool for researchers and clinicians alike. (The full scale with scoring instructions can be found in the supplementary material.)

When comparing the CHIME-S and CHIME-XS with other mindfulness measures, such as the FFMQ (Baer et al., 2006), our findings suggest that the CHIME-S and CHIME-XS provide a more comprehensive assessment of mindfulness in line with traditional conceptualizations. The development of the CHIME-S and CHIME-XS supports the growing body of research emphasizing the importance of incorporating a broader range of mindfulness characteristics (Bergomi et al., 2013; Wilkinson et al., 2023), such as Insightful Understanding and Awareness of Thoughts' Relativity, to better understand and measure the construct. The CHIME-S and CHIME-XS might enable researchers to utilize the CHIME's broad conceptualization of mindfulness in a wider range of studies and measure these often disregarded mindfulness concepts. The shorter version of the CHIME offers a more feasible way to explore the effects of mindfulness on cognition beyond the changes in attentional control. More complex skills, such as the ability to

understand and incorporate concepts such as impermanence in one's life, will now be possible to capture by an objective measure, making room for a more in-depth understanding of the mechanisms of change related to mindfulness practice. Researchers and practitioners alike are encouraged to consider the CHIME-S and CHIME-XS as an alternative to the FFMQ. We also encourage further validation of the measure across different cultural contexts and populations, in order to strengthen its robustness and generalizability.

Limitations and Directions of Future Research

While our study offers valuable insights into the development and validation of the short CHIME version, there are several limitations to consider. Primarily, our sample selection predominantly comprised students, which might limit the generalizability of our findings to broader populations. The student demographic might have specific characteristics that differ from the general public or specific clinical groups; and thus, future studies should consider a more diverse sampling strategy. Another significant limitation is the absence of objective mindfulness measures to further validate the scale. Relying solely on self-report measures can introduce biases and may not capture the full depth and nuances of an individual's mindfulness practices. The study by Altgassen et al. (2023) is particularly illuminating in this respect, showcasing how a mindfulness measure can be derived across existing mindfulness scales. Nevertheless, this approach favors simplicity above clear alignment with theory and might led to the derivation of factors which are empirically stable but encompass a large amount of non-specific variance (Alexandrova & Haybron, 2016). Our CHIME short forms retain a balance between an empirically derived short form which nevertheless can be clearly mapped onto a wider theory of mindfulness.

The successful replication of the CHIME-S's and CHIME-XS's psychometric properties across three independent samples including both student and non-student populations indicates its potential for generalizability across different populations. Moving forward, it is essential to validate the scale in other cultural contexts, transcending Western samples, and diverse sample populations to establish its robustness as a measure of mindfulness which has been a consistent issue for mindfulness measures (Karl et al., 2020, 2022). Similarly, while the current study indicates the cross-sectional robustness of the shortened CHIME, this does not indicate temporal stability, which should be explored in future studies. Additionally, future studies could explore the clinical utility of the CHIME-S in assessing changes in mindfulness following mindfulness-based interventions and investigating the relationship between mindfulness and various psychological outcomes.

In conclusion, the CHIME-S and CHIME-XS are reliable and valid short-form scales for assessing mindfulness and its facets in a comprehensive manner. The development of this shortened scale will facilitate its use in large-scale studies and enable researchers and clinicians to assess mindfulness more efficiently. We also demonstrated the application of a novel optimization technique which shows significant promise for further scale development work. Future research should focus on validating the CHIME-S and CHIME-XS across different cultural contexts and sample populations, as well as exploring its clinical utility in various therapeutic settings. The CHIME-S and CHIME-XS administration format and scoring instructions are included here (Supplementary Tables S1 and S2).

Acknowledgements The authors are grateful to all participants who participated in this study.

Author Contribution JAK: conceptualization, methodology, data analysis, writing—original draft preparation and editing. LR: data collection and curation, writing—reviewing and editing. CB: conceptualization, writing—reviewing and editing. RF: methodology, writing—reviewing and editing. SD: writing—reviewing and editing. ONM: conceptualization, methodology, writing—original draft preparation and editing.

Funding Open Access funding provided by the IReL Consortium.

Data Availability The data is available on the OSF <https://osf.io/8uyac/>.

Declarations

Ethics Approval The ethical approval for this study was granted by the Pacific University Institutional Research Board (protocol number 131–17), and the Victoria University of Wellington Ethics Committee (#0000027072), which follow internationally recognized ethical standards.

Informed Consent All participants provided their informed consent prior to participation in the study.

Use of Artificial Intelligence We did not use any artificial intelligence in this article.

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alexandrova, A., & Haybron, D. M. (2016). Is construct validation valid? *Philosophy of Science*, 83(5), 1098–1109. <https://doi.org/10.1086/687941>
- Altgassen, E., Geiger, M., & Wilhelm, O. (2023). Do you mind a closer look? A jingle-jangle fallacy perspective on mindfulness. *European Journal of Personality*. <https://doi.org/10.1177/08902070231174575>
- Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. *Psychological Assessment*, 10(2), 176–181. <https://doi.org/10.1037/1040-3590.10.2.176>
- Azizi, Z., Jandaghi, G., Firoozi, M., Zia-Tohidi, A., & Ebnerasouli, S. (2022). The relative importance of mindfulness facets and their interactions: Relations to psychological symptoms in chronic pain. *Mindfulness*, 13(4), 1054–1068. <https://doi.org/10.1007/s12671-022-01862-2>
- Baer, R., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-report: The Kentucky Inventory of Mindfulness Skills. *Assessment*, 11(3), 191–206. <https://doi.org/10.1177/1073191104268029>
- Baer, R., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, 13(1), 27–45. <https://doi.org/10.1177/1073191105283504>
- Bartos, L. J., Posadas, M. P., Wrapson, W., & Krägeloh, C. (2023). Increased effect sizes in a mindfulness- and yoga-based intervention after adjusting for response shift with then-test. *Mindfulness*, 14(4), 953–969. <https://doi.org/10.1007/s12671-023-02102-x>
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 56(6), 893–897. <https://doi.org/10.1037/0022-006X.56.6.893>
- Bennett, K., & Dorjee, D. (2016). The impact of a mindfulness-based stress reduction course (MBSR) on well-being and academic attainment of sixth-form students. *Mindfulness*, 7(1), 105–114. <https://doi.org/10.1007/s12671-015-0430-7>
- Bergomi, C., Tschacher, W., & Kupper, Z. (2013). The assessment of mindfulness with self-report measures: Existing scales and open issues. *Mindfulness*, 4(3), 191–202. <https://doi.org/10.1007/s12671-012-0110-9>
- Bergomi, C., Tschacher, W., & Kupper, Z. (2014). Konstruktion und erste Validierung eines Fragebogens zur umfassenden Erfassung von Achtsamkeit [Construction and initial validation of a questionnaire for the comprehensive investigation of mindfulness]. *Diagnostica*, 60, 111–125. <https://doi.org/10.1026/0012-1924/a000109>
- Blum, C., & Roli, A. (2003). Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys*, 35(3), 268–308. <https://doi.org/10.1145/937503.937505>
- Bohlmeijer, E., ten Klooster, P. M., Fledderus, M., Veehof, M., & Baer, R. (2011). Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. *Assessment*, 18(3), 308–320. <https://doi.org/10.1177/1073191111408231>
- Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, 84(4), 822–884. <https://doi.org/10.1037/0022-3514.84.4.822>
- Cladder-Micus, M. B., Verweij, H., van Ravesteijn, H., Van der Gucht, K., Raes, F., & Speckens, A. E. (2019). Validation of the Dutch Comprehensive Inventory of Mindfulness Experiences (CHIME) and development of a short form (CHIME-SF). *Mindfulness*, 10(9), 1893–1904. <https://doi.org/10.1007/s12671-019-01125-7>
- de Vibe, M., Solhaug, I., Rosenvinge, J. H., Tyssen, R., Hanley, A., & Garland, E. (2018). Six-year positive effects of a mindfulness-based intervention on mindfulness, coping and well-being in medical and psychology students: Results from a randomized controlled trial. *PLoS ONE*, 13(4), e0196053. <https://doi.org/10.1371/journal.pone.0196053>
- Dorigo, M., & Stützle, T. (2004). *Ant colony optimization*. MIT Press.
- Dorigo, M., Maniezzo, V., & Colomi, A. (1996). Ant system: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(1), 29–41. <https://doi.org/10.1109/3477.484436>
- Fischer, R., & Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Frontiers in Psychology*, 10, 1507. <https://doi.org/10.3389/fpsyg.2019.01507>
- Harvey, P. (2013). Dukkha, Non-Self, and the Teaching on the Four “Noble Truths”. In S. M. Emmanuel, (Ed.), *A companion to Buddhist philosophy*. Wiley. <https://doi.org/10.1002/9781118324004.ch2>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: Past, present, and future. *Clinical Psychology: Science and Practice*, 10(2), 144–156. <https://doi.org/10.1093/clipsy.bpg016>
- Karl, J. A., Johnson, F. N., Bucci, L., & Fischer, R. (2022). In search of mindfulness: A review and reconsideration of cultural dynamics from a cognitive perspective. *Journal of the Royal Society of New Zealand*, 52(2), 168–191. <https://doi.org/10.1080/03036758.2021.1915804>
- Karl, J. A., Méndez Prado, S. M., Gračanin, A., Verhaeghen, P., Ramos, A., Mandal, S. P., Michalak, J., Zhang, C.-Q., Schmidt, C., Tran, U. S., Druica, E., Solem, S., Astani, A., Liu, X., Luciano, J. V., Tkaličić, M., Lilja, J. L., Dundas, I., Wong, S. Y. S. Y., ..., Fischer, R. (2020). The cross-cultural validity of the Five-Facet Mindfulness Questionnaire across 16 countries. *Mindfulness*, 11(5), 1226–1237. <https://doi.org/10.1007/s12671-020-01333-6>
- Krägeloh, C. U., Bergomi, C., Siegert, R. J., & Medvedev, O. N. (2018). Response shift after a mindfulness-based intervention: Measurement invariance testing of the Comprehensive Inventory of Mindfulness Experiences. *Mindfulness*, 9(1), 212–220. <https://doi.org/10.1007/s12671-017-0764-4>
- Krägeloh, C. U., Henning, M. A., Medvedev, O. N., Feng, X. J., Moir, F., Billington, R., & Siegert, R. J. (2019). *Mindfulness-based intervention research: Characteristics, approaches, and developments*. Routledge.
- Lecuona, O., García-Rubio, C., de Rivas, S., Moreno-Jiménez, J. E., & Rodríguez-Carvajal, R. (2022). Unraveling heterogeneities in Mindfulness Profiles: A review and latent profile analysis of the Five Facet Mindfulness Questionnaire short-form (FFMQ-SF) in the Spanish population. *Mindfulness*, 13(8), 2031–2046. <https://doi.org/10.1007/s12671-022-01939-y>
- Leite, W. L., Huang, I.-C., & Marcoulides, G. A. (2008). Item selection for the development of short forms of scales using an ant colony optimization algorithm. *Multivariate Behavioral Research*, 43(3), 411–431. <https://doi.org/10.1080/00273170802285743>
- Lewinsohn, P. M., Seeley, J. R., Roberts, R. E., & Allen, N. B. (1997). Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychology and Aging*, 12(2), 277–287. <https://doi.org/10.1037/0882-7974.12.2.277>
- Linehan, M. M. (1993). *Skills training manual for treating borderline personality disorder*. Guilford Press.
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress

- Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335–343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)
- Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research*, 3(1), 97–110. <https://doi.org/10.21500/20112084.854>
- McBride, E. E., Chin, G. R., Clauser, K. S., & Greeson, J. M. (2022). Perceived stress mediates the relationship between trait mindfulness and physical symptoms of stress: A replication study using structural equation modeling. *Mindfulness*, 13(8), 1923–1930. <https://doi.org/10.1007/s12671-022-01928-1>
- Medvedev, O. N., Bergomi, C., Röthlin, P., & Krägeloh, C. U. (2019). Assessing the psychometric properties of the Comprehensive Inventory of Mindfulness Experiences (CHIME) using Rasch analysis. *European Journal of Psychological Assessment*, 35(5), 650–657. <https://doi.org/10.1027/1015-5759/a000453>
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, 28(6), 487–495. [https://doi.org/10.1016/0005-7967\(90\)90135-6](https://doi.org/10.1016/0005-7967(90)90135-6)
- Okafor, G. N., Ford, B. Q., Antonoplis, S., Reina, A. M., Lutfiali, S., & Shallcross, A. J. (2023). Measuring mindfulness in Black Americans: A psychometric validation of the Five Facet Mindfulness Questionnaire. *Mindfulness*, 14(3), 565–581. <https://doi.org/10.1007/s12671-023-02072-0>
- Olaru, G., & Danner, D. (2021). Developing cross-cultural short scales using ant colony optimization. *Assessment*, 28(1), 199–210. <https://doi.org/10.1177/1073191120918026>
- Olaru, G., Schroeders, U., Hartung, J., & Wilhelm, O. (2019). Ant colony optimization and local weighted structural equation modeling. A tutorial on novel item and person sampling procedures for personality research. *European Journal of Personality*, 33(3), 400–419. <https://doi.org/10.1002/per.2195>
- Raborn, A. W., & Leite, W. L. (2018). ShortForm: An R package to select scale short forms with the ant colony optimization algorithm. *Applied Psychological Measurement*, 42(6), 516–517. <https://doi.org/10.1177/0146621617752993>
- Schroeders, U., Wilhelm, O., & Olaru, G. (2016). Meta-heuristics in short scale construction: Ant colony optimization and genetic algorithm. *PLoS ONE*, 11(11), e0167110. <https://doi.org/10.1371/journal.pone.0167110>
- Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. *Structural Equation Modeling*, 5(4), 411–419. <https://doi.org/10.1080/10705519809540115>
- Van Gordon, W., Shonin, E., Griffiths, M. D., & Singh, N. N. (2015). There is only one mindfulness: Why science and Buddhism need to work together. *Mindfulness*, 6(1), 49–56. <https://doi.org/10.1007/s12671-014-0379-y>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Wilkinson, S., Ribeiro, L., Krägeloh, C. U., Bergomi, C., Parsons, M., Siegling, A., Tschacher, W., Kupper, Z., & Medvedev, O. N. (2023). Validation of the Comprehensive Inventory of Mindfulness Experiences (CHIME) in English using Rasch methodology. *Mindfulness*, 14(5), 1204–1218. <https://doi.org/10.1007/s12671-023-02099-3>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.