

Can We Agree What Skilled Mindfulness-Based Teaching Looks Like? Lessons From Studying the MBI:TAC

Global Advances in Health and Medicine

Volume 0: 1–11

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/2164956120964733

journals.sagepub.com/home/gam



Rebecca S Crane, PhD¹ , Frederick M Hecht, MD²,
Judson Brewer, MD³, Gemma M Griffith, PhD¹ ,
Wendy Hartogensis, PhD², Lynn Koerbel, MPH³,
Patricia Moran, PhD², Sophie Sansom, PhD¹, Alison Yiangou, MA⁴
and Willem Kuyken, PhD, DClinPsy⁴ 

Abstract

Background: The Mindfulness-Based Interventions: Teaching Assessment Criteria (MBI:TAC) is a widely used tool for assessing fidelity in mindfulness-based program (MBP) research and training. It also supports MBP teacher reflective and skill development. MBI:TAC assessors review MBP teaching and rate the teaching on 6 domains. The MBI:TAC yields individual domain and overall scores, using 6 levels of competence. Although the MBI:TAC is widely used in MBP research and training, research is at an early stage.

Objective: We developed and tested a method of training MBI:TAC assessors to use the tool reliably and examined interrater reliability of the tool.

Methods: A total of 31 international senior MBP teachers were recruited to join an online training to build their skills in using the MBI:TAC. The training systematically and iteratively built familiarity and skills in assessing the 6 MBI:TAC domains. Qualitative and quantitative data on trainee's experience of the training were gathered. Interrater reliability in using the tool was tested each week of the training. At the end of the training, interrater reliability was tested by asking trainees to individually assess videos that they had not previously seen. Their ratings were compared to benchmark assessments, which had been established via consensus agreement between 4 expert users of the MBI:TAC.

Results: The training was well received and appreciated, with some challenges experienced in applying the assessment methodology. Participants' ratings became progressively more in line with one another and the benchmark ratings during the training. At the end, interrater reliability was high (ranging from 0.67 to 1.0).

Conclusion: It is possible for senior MBP trainers, coming from different regions in the world, to align toward common understandings of the elements of MBP teaching competence and program integrity. An assessor training methodology was tested, and the learning from this project has led to refinements for future delivery.

Keywords

Mindfulness-Based Interventions: Teaching Assessment Criteria, fidelity, mindfulness-based teaching skill and competence, Mindfulness-Based Stress Reduction, Mindfulness-Based Cognitive Therapy

Received May 5, 2020; Revised July 1, 2020. Accepted for publication July 20, 2020

Introduction

The Mindfulness-Based Interventions: Teaching Assessment Criteria (MBI:TAC)¹ is used within mindfulness-based teacher research and training programs to enable assessment of teaching competence. It is also used informally to support skill development and reflective learning (Evans A, Griffiths GM, Sansom S, Crane RS. Using the Mindfulness-Based Interventions:

¹Centre for Mindfulness Research and Practice, Bangor University, Bangor, UK

²Osher Center for Integrative Medicine, University of California San Francisco, San Francisco, California

³Center for Mindfulness, Brown University, Providence, Rhode Island

⁴Oxford Mindfulness Centre, University of Oxford, Oxford, UK

Corresponding Author:

Rebecca Crane, Bangor University, Brigantia Building, College Rd, Bangor, LL57 2DG, UK.

Email: r.crane@bangor.ac.uk



Teaching Assessment Criteria (MBI-TAC) in supervision. *Glob Adv Ment Health*. Under review; Griffith GM, Crane RS, Fernandez E, Giommi F, Herbette G, Koerbel L. Implementing the Mindfulness-Based Interventions: Teaching Assessment Criteria (MBI-TAC) in mindfulness teacher training programs. *Glob Adv Ment Health*. Under review). Indeed, there is considerable impetus behind the uptake of the tool in both training and research settings. For example, 45 mindfulness-based program (MBP) teacher training centers are known to be using the tool both as an assessment tool and as an informal support for learning in 14 countries, and it has been translated into 7 languages.² In England, the government-funded Mindfulness-Based Cognitive Therapy (MBCT) teacher training program, designed to build capacity for delivery in the National Health Service, uses the MBI:TAC to support reflective development and assess competence.³ The internationally agreed MBCT teacher training pathway uses the MBI:TAC as part of the learning and assessment process.⁴ The MBI:TAC is being integrated into research protocols—both as a check on intervention integrity⁵ and as a tool to examine MBP teacher training and teacher factors.^{6,7} Another factor influencing uptake is that currently the MBI:TAC is the only tool that is anchored into the 2 main empirically supported MBPs (Mindfulness-Based Stress Reduction [MBSR] and MBCT) which assesses both adherence and competence.

This rapid uptake of the tool is likely being driven by a combination of expansion of MBP research and practice, alongside a commitment to ensuring integrity in practice settings. There are inevitable challenges of ensuring delivery quality both in research and in real-world practice settings. A tool such as the MBI:TAC is a unifying force in the field internationally as it communicates consensus on the common elements of MBP teaching.⁸ It is fair to say, however, that the implementation of the tool is ahead of systematic research on its reliability, validity, and its relationship to MBP's intended outcomes. Whilst there are grounds for this, in that the tool has a practical contribution to make to the pedagogy of MBP teacher training, it is also important that research examines the psychometric properties of the tool and that its implementation in practice is research-led.

Initial validation of the MBI:TAC instrument was promising. Preliminary research on the psychometric properties of the tool was conducted within the context of routine assessment of MBP teaching practice within 3 UK mindfulness Master's degree programs, which had MBP teacher training embedded into them.⁹ This preliminary study demonstrated strong interrater reliability (intraclass correlation coefficient [ICC]: $r = .81$) and good evidence of internal consistency, construct validity, and concurrent validity. The process demonstrated a proof of concept—that it is possible to define the observable aspects of MBP

teaching competence and to organize these into a reliable and valid tool. However, a second study conducted in the Netherlands with 16 trainers generated more modest ICCs (0.55 [moderate] for domain 1, 0.67 [substantial] for domain 2, 0.45 [moderate] for domain 3, 0.68 [substantial] for domain 4, 0.63 [substantial] for domain 5, and 0.58 [moderate] for domain 6).¹⁰ The same developers (RC and WK) conducted the assessor training in this Dutch study, and the learning from this process was used in this study. In particular, we hypothesized that a more systematic approach to building assessment skills over time, with tests to feedback reliability level along the way, would lead to greater assessment accuracy.

If instruments like the MBI:TAC are to be used more broadly in the training of MBP teachers and as a tool for assessing intervention fidelity, a critical question is whether new assessors can be trained in a way that results in good interrater reliability. Ensuring the quality of the implementation of the tool has the potential to influence MBP teacher training practice by linking training and research centers internationally to commonly agreed norms and standards of practice. In turn, this has the potential to positively influence the quality of MBP teaching internationally, thus supporting accessibility of the benefits of MBPs globally. Within the international mindfulness-based training context, there is increasing demand for access to quality training for prospective users of the MBI:TAC for both assessment and reflective purposes.

In this paper, we report on phase 1 of the Predictors of Outcomes in MBSR Participants from Teacher Factors (PROMPT) research study.⁶ This study was based on the premise that whilst initial validation steps for the MBI:TAC were promising, additional research was needed to inform application in MBP research and practice, including testing whether new assessors could be trained to achieve reliable results using the MBI:TAC. The study had 2 main phases—first, to train a group of assessors to use the MBI:TAC reliably (interrater reliability), and second, to examine a range of teacher factors that could influence MBP participant outcomes (reported in forthcoming paper). This paper reports the outcomes of the first phase and discusses both evidence-based practice and practice-based evidence that informs how to train MBP teacher assessors, supervisors, and trainers to use the MBI:TAC. The research reported in this paper had 2 aims: first, to develop and test a method of training MBI:TAC assessors who can use the tool reliably, and second, to examine in an independent study, interrater reliability using this training model.

Methods

We recruited 31 participants for the MBI:TAC assessor training who met the following criteria: (1) certified to teach through an established and recognized MBP teacher

training program, (2) more than 3 years of MBSR teaching experience (the focus was on MBSR as this was the anchor for the PROMPT research funding), and (3) have the interest and time available to complete assessor training and subsequent teacher ratings. We developed and delivered training sessions using a videoconference platform (Zoom). The 31 international participants were strategically selected to represent senior mindfulness-based trainers and supervisors across the world in order to both train assessors for the PROMPT research and to build capacity in the use of the MBI:TAC internationally. See Table 1 for rater characteristics.

The delivery was structured as 7×1.5 h sessions, delivered by 2 trainers (RC and WK), involving didactic presentations on the nature of integrity in MBP teaching; the background and context for the development of the MBI:TAC and its design and structure; tutorials walking trainees through observing/reviewing recordings of MBP teaching and making a rating on each MBI:TAC domain; building skills in assessing through discussions in small and the whole group; and assigning home practice of independently assessing teaching video recordings of teaching. Each week iteratively built understanding of the 6 domains of the MBI:TAC (Table 2) and practised these through focusing on different aspects of MBP curriculum elements (Table 3).

To establish interrater reliability throughout the training process, participants were asked to rate a short video clip of an MBP teaching session between each training session, and their individual MBI:TAC assessments were compared to benchmark assessments. To establish reliability after the training, we asked trainees to individually assess at least 3 videos that they had not previously seen, with teachers they did not know. Their ratings were compared to benchmark assessment points for the videos that had been established via consensus agreement between 4 expert users of the MBI:TAC (trained by the MBI:TAC developers to use the tool reliably as tested through established benchmarks).

To further evaluate the training delivery, we conducted a survey of the participants with 6 questions (see Table 4) that participants answered on a scale of 1 (not at all) to 10 (extremely). Participants were also invited to offer qualitative feedback in response to the questions: (1) Please let us know any suggestions for how we could strengthen a future MBI:TAC training, and (2) Any other comments on the training? (see Table 5).

Analysis

To assess whether the training course improved the consistency of ratings done by trainees during the 7 weeks of training, we first calculated the difference between trainee homework ratings and benchmark assessments (ratings) from each class, and for each MBI:TAC domain.

Table 1. Description of MBI:TAC Rater Training Participants (N = 31).

Characteristic	Frequency
Age, years, mean (SD) (range)	54.1 (10.7) (30–73)
Gender, n (%)	
Male	2 (6.5%)
Female	29 (93.6%)
Race/ethnicity, n (%)	
White	29 (93.5%)
Asian	1 (3.2%)
Latinx	1 (3.2%)
Education, n (%)	
Some college	1 (3.2%)
Bachelor's degree	3 (9.7%)
Master's degree	16 (51.6%)
Doctoral degree	11 (32.3%)
Marital status, n (%)	
Never married	2 (6.5%)
Married/cohabiting	23 (74.2%)
Separated/divorced	5 (16.1%)
Decline/unknown	1 (3.2%)
Spiritual affiliation, n (%)	
Buddhist	12 (38.7%)
Christian	3 (9.7%)
Jewish	1 (3.2%)
Spiritual	5 (16.1%)
None	3 (9.7%)
Decline/unknown	7 (22.6%)
Certified to teach MBSR, n (%)	
Yes	24 (77.4%)
No	7 (22.5%)
Years teaching MBI, mean (SD) (range)	12.6 (7.1) (3–33)
Years teaching MBSR, mean (SD) (range)	10.7 (5.9) (1–23)
MBSR courses/year, mean (SD) (range)	5.3 (5.6) (0–23)
Languages teaches MBSR in, n (%)	
English only	17 (54.8%)
English and 1+ other language (Spanish, Danish, German, Dutch)	4 (12.9%)
Non-English language only (eg, Danish, Spanish, French)	3 (9.7%)
Not certified in MBSR	7 (22.6%)
Years of personal meditation practice, mean (SD) (range)	21.5 (9.9) (4–40)
Number of silent meditation retreats of 3+ days attended	17.3 (13.6) (3–70)

Abbreviations: MBI, Mindfulness-Based Intervention; MBSR, Mindfulness-Based Stress Reduction.

We took the standard deviations of the differences across trainees, within class week and domain. These standard deviations (one per MBI:TAC domain and class week) were regressed on class week, with a random domain effect to allow differences by domain. Assessor (rater) reliability was calculated based on a maximum likelihood estimator of the ICC from a one-way random effects model treating the benchmark assessments as a gold standard.¹¹

Table 2. Domains of the MBI:TAC.

Domain Number	Domain Title
1	Coverage, pacing, and organization of session curriculum
2	Relational skills
3	Embodiment of mindfulness
4	Guiding mindfulness practices
5	Conveying course themes through interactive inquiry and didactic teaching
6	Holding the group learning environment

Table 3. Overview of Assessor Training.

Week	Curriculum Element	MBI:TAC Competency Domains Covered
1	Body scan	3 and 4
2	Sitting meditation	3 and 4
3	Mindful movement practice and inquiry	2, 3, 4, and 5
4	Practice + inquiry	2, 3, 4, and 6
5	Interactional components of weeks 4–5	All domains
6	Interactional components of weeks 6–7	All domains
7	Review, bringing it all together	All domains

Abbreviations: MBI:TAC, Mindfulness-Based Interventions: Teaching Assessment Criteria.

Twenty-three participants completed the anonymous feedback questionnaire. The quantitative responses were analyzed by deriving a mean, standard deviation, and variance score for each of the 6 questions. The qualitative survey data were analyzed using a content analysis approach,^{12,13} a systematic technique which allows text to be put into categories based upon rules of coding. Here, emergent coding was employed, with categories established during the content analysis as they emerged from the data. The categories thus represent text that was grouped based on having similar meanings. For parsimony, categories that contained 2 or fewer comments (for example, 1 participant reported technical difficulties) were not included as categories.

Results

Evaluation of Consistency of Homework Ratings Over Course Time

As a measure of how training improved rating consistency over time, we took the standard deviations of the

Table 4. Quantitative Feedback on the MBI:TAC Training From Trainees (N =23).

Question	Mean	Standard Deviation
1. How likely is it that you would recommend our MBI:TAC training to a friend or colleague who is an experienced mindfulness instructor?	8.39	1.24
2. How likely is it that you would recommend getting rated using the MBI:TAC to someone training to lead mindfulness interventions?	7.65	1.78
3. How likely is it that you would recommend getting rated using the MBI:TAC to someone with experience leading mindfulness interventions?	8.3	1.46
4. How would you rate your overall satisfaction with the MBI:TAC training program?	8.0	1.29
5. How would you rate your competency in using the MBI:TAC?	6.74	1.07
6. How would you rate the following aspects of the MBI:TAC training?		
Course leaders	9.48	0.83
Course content	8.17	1.13
Course length	7.61	1.79
Videoconference platform	8.96	0.95

Abbreviations: MBI:TAC, Mindfulness-Based Interventions: Teaching Assessment Criteria.

deviation between the trainee homework assessments and benchmark assessments of the same videos, by class week and MBI:TAC domain. With a random intercept allowing for differences by domain, we found that the standard deviation decreased on average across all domains by 0.10 points per week during the 7-week training (95% confidence interval: 0.05 to 0.15, $P < 0.001$). As shown in Table 2, instruction on domains was cumulative rather than sequential: each session and subsequent home practice included focus on all domains covered in prior weeks as well as any newly introduced domains. Therefore, the impact of session week on the deviations between homework rating and benchmark reflects cumulative weeks of study rather than recency of studying the domain assigned. To further assess this relationship, as a sensitivity analysis, we restricted the model to domains 3 and 4 (which are studied in all class weeks), and results were identical.

Evaluation of Interrater Reliability

Twenty-four participants completed the training and provided a set of final assessments of MBP teaching videos after completing the course. Interrater reliability coefficients were calculated for each trainee assessor (rater), comparing ratings of at least 3 videos for each

Table 5. Content Analysis of Trainee Feedback.

Category (N=Code Counts for Each Category)	Category Example Quote(s)	Subcategory (n=Code Counts for Each Subcategory)	Subcategory Example Quotes
1 A privilege to be learning new skills (N = 16)	I so appreciated the heart, time, effort and thought that went into the organization and structure of the online training format (...) Much gratitude for this opportunity for heart/mind. So glad to have had this opportunity to learn (with so many talented and experienced others) the MBI:TAC process!	1a. Enabled reflection on their own teaching (n = 6) 1b) Learned about the MBI:TAC (n = 4)	1a) I felt privileged to be a part of this training program. It raised so many interesting questions for me about teaching, competencies in my own teaching, and other factors. Thank you so much. 1b) Thank you. It was a superb entry into a more global sense of how the MBI:TAC is used and also "lands" with different cultures (including different training cultures). More to learn and explore. Deep appreciation for the opportunity to engage in this way.
2 Wanted more time for discussion together as a group (N = 13)	I think that participants would benefit from more opportunity to share reflections. The sessions need to be longer, allowing more discussion in the smaller groups particularly.	2a) Interrater reliability: needed further discussion (n = 6) 2b) More time to explore how to assign ratings to the MBI:TAC (n = 3)	2a) Where benchmarks were higher than the "consensus" by raters, and where interrater agreement was low, it would have been really helpful to look in much more detail at where, why and how those differences arose. 2b) If there were some ratings that could be done by us as individuals, then discuss with one or 2 partners in between big sessions . . . The break out sections were useful, but I would have liked more time to lean into the tool in doing a rating and then have the opportunity to discuss.
3 Skillful leadership by facilitators (N = 6)	I am very grateful to have had this opportunity and to Becca and Willem for their openness, humility, and trust. Just a huge thank you to Becca and Willem for their leadership, their responsiveness, their very beings. It was a privilege to be included in this endeavor.	n/a	n/a
4 Enjoyed the process—connection with other teacher trainers (N = 6)	I can sometimes feel isolated in my work, so connecting with a worldwide network of mindfulness teachers was very strengthening. The opportunity to have dialogue with other teachers/raters from other countries is priceless.	n/a	n/a
5 Confusion/frustration around the rationale for the	Much of the frustration that I felt coming (for me and for the group) had to do with not	5a) More people needed for interrater reliability to	5a) A team approach to setting the benchmarks will make a difference in the future.

Table 5. Continued.

Category (N=Code Counts for Each Category)	Category Example Quote(s)	Subcategory (n=Code Counts for Each Subcategory)	Subcategory Example Quotes
6 MBI:TAC base- line ratings used in the training (N = 5) Ratings on the MBI: TAC are com- plex (N = 5)	knowing more in detail the reasoning behind the benchmark ratings. I feel my ratings continue to be impacted by lack of clarity between domains/impact of one domain (embodiment) on other domains and trying to rate based on a relatively short clip not knowing what was covered previously. For example, in the “test” clip 3, although the meditation lead was competent, it didn’t seem appropriate for class 5 . . . how is that scored?	establish the baseline prior to the training (n = 3) 6a) Issues of culture and curriculum (n = 2)	6a) There is inherent complexity in the assessment process due to variations in curricula, culture, and other factors which makes consistency of evaluation very challenging. “I feel that it would be best for MBSR teachers to assess MBSR teachers, and MBCT teachers to assess MBCT teachers, as we might be more expert in the curriculum.”
7 Opportunity to learn how to offer narrative feedback, not just numerically rate the MBI: TAC (N = 4)	The first is to work with the commentary sheets in small groups. This feels like a critical piece since the ratings mean little without them and skill in formulating observations in these areas seems worthy of investigation and discussion.	7a) Wanted to practice narrative comments—not just a numerical value (n= 3)	7a) I would have liked the opportunity to practice completing the narrative feedback to teachers rather than just making a rating.
8 Wanted more specific information from training (N = 4)	It would be helpful to see a few examples with ratings prior to or at the start of the training have a range of competency level videos.	8a) More information given about the MBI:TAC domains (n = 2)	8a) I would have appreciated some more discussion, clarification, and agreement on the key features, so we could have laid more of a common ground before rating.

Abbreviations: MBCT, Mindfulness-Based Cognitive Therapy; MBI:TAC, Mindfulness-Based Interventions: Teaching Assessment Criteria; MBSR, Mindfulness-Based Stress Reduction.

MBI:TAC domain against the “gold standard” of the benchmark assessments for the same videos. Rater reliability overall was high, ranging from 0.67 to 1.0 (1.0 represents perfect agreement between participant and benchmark assessments), with a mean of 0.94. Interrater reliability was consistently high across participants for ratings of MBI:TAC domains 1 (coverage, pacing, and organization of session curriculum), 2 (relational skills), and 5 (conveying course themes through interactive inquiry and didactic teaching). Reliability remained high for a majority of participants across other MBI:TAC domains but with some outliers, and the lowest reliability was seen for domain 4 (guiding mindfulness practice).

Evaluation of Trainee Experience of the Training

The quantitative responses from the participants are laid out in Table 3.

The results of the content analysis of the qualitative trainee feedback are presented in Table 4.

Course Completion

Out of the 31 recruited participants, all completed the training. Four of these trainees were selected, based on the accuracy of their midtraining homework ratings, to join RC and WK in determining, by a process of consensus, the final benchmark ratings used as the standard for comparison for the final assignment. Twenty-four of the 27 who were asked to complete the final benchmark ratings (which were judged against the consensus standard for calculation of interrater reliability) did so. Of the 3 raters who did not complete the final benchmark ratings, 2 cited lack of time, while the other expressed intention to complete them but failed to actually do so. These 3 noncompleters did not differ from the other raters on any demographic, teaching, or meditation history characteristic. Twenty-three participants completed the optional feedback questionnaire. As is common for course evaluations, in order to encourage honest feedback, our end-of-training survey was anonymous; thus, we are unable to determine differences between the 23 completers and 8 noncompleters.

Discussion

First, we discuss the analysis of MBI:TAC interrater reliability, and then the evaluation of the trainees experience of the training process, before drawing out implications for practice.

Interrater Reliability

The consistency of assessments (ratings) improved over the 7 weeks of training, and final ratings were close to

benchmarks for the majority of participants. The level of reliability developed through this process was better than that of both the previous studies on the MBI:TAC.^{9,10} There is evidence therefore that this more systematic approach to training hones participants' skills in using the MBI:TAC reliably. Within this, there were some outliers: a small number of participants still had relatively low interrater reliability compared to benchmarks across 4 of the 6 domains at the end of the training. Our sample size does not enable any reliable inferences about the characteristics of those whose reliability was higher or lower. Thus, whilst there is evidence that suggests that the majority of participants internalized the training well, it is also important that in future trainings there is an assessment process to ensure participants reach an acceptable understanding of the MBI:TAC prior to using it in the field (and perhaps periodic checks to ensure reliability over time). A further important finding was that domain 4 (guiding mindfulness practices) showed less agreement than other domains and therefore may benefit from additional training components. This is in contrast to the Huijbers et al.'s¹⁰ study which had stronger reliability in this domain. One possible reason for this is that the teaching videos used to assess reliability in this study included both MBCT and MBSR which have similarities but also differences about how the mindfulness practices are guided. This may have confounded the process when our trainee cohort was predominantly comprised of MBSR teachers. Interestingly, embodiment was rated more consistently in this study than previous studies. This may be due to work that has taken place to make the descriptors in the embodiment domain more precise, and a greater focus within the training on these particular descriptors when rating this domain.

A study limitation was the available sample of recordings of MBP teaching. Only between 3 and 4 were available for each MBI:TAC domain. Moreover, benchmark assessments did not span the full range of possible ratings for any domain but instead ranged from 4 (competent) to 6 (advanced) (with a mean of 4.9 [$SD = 0.85$] across domains for the video clips used in the interrater reliability analysis). A more thorough evaluation of the training could be achieved with a larger set of video clips and with clips spanning the full range of possible assessment values. However, in practice, it is challenging to get clips with the appropriate permissions in place, and for ethical and practical reasons, we were not able to obtain video recordings of MBP instructors who would merit very low assessments, indicating low teacher skills. It will be important going forward to have the full range of MBI teaching (from lower levels of the competence range through to advanced) to rate in studies such as this. However, despite these limitations, we were still

able to draw useful conclusions comparing reliability across domains and across participants.

Evaluation of Participant's Experience of the Training

This discussion on the training draws both on the qualitative and quantitative feedback, along with the experiences of the trainers (RC and WK). The MBI:TAC assessor training was appreciated by participants, with an aggregated mean of 8.14 for all 9 questions (scale of 0 = not at all to 10 = extremely), and appreciative comments being the largest category in the content analysis. Participants valued the opportunity to develop personal skills, as well as the wider engagement around shared concerns about teaching standards and discussions of a potential way to create joined-up standards internationally.

The lowest score was in relation to participant's view of their competence in using the tool (mean = 6.74), which speaks to the challenges trainees experienced in engaging with the assessment process. Participants particularly struggled when their assessment points were not aligned with the benchmarks. There was a strong call for a longer training process to enable more time to reflect on the assessment process, greater examination of the rationales underpinning assessment points, reflections, and skills building. The training was the first of its kind, and many participants were only just beginning to consider assessment of teaching as a concept—so the immediate immersion in skills building with the MBI:TAC was a big step for many.

It is important to note that the questionnaire used with this particular sample at the end of the course was likely to elicit socially desirable outcomes. Future investigations of the training process should be more proactive in actively probing for critical reflections on the training and the tool, ideally conducted by an independent researcher using a topic list of possible problems and issues. This should include those who did not complete the training, who did not respond to the invitation to offer feedback, and whose assessments did not match the benchmarks.

The training included participants from America and European countries, and this triggered questions of cultural differences in teaching styles. As trainers (RC and WK), we noticed that the question of cultural differences was raised frequently by participants during the training sessions: how possible is it for an assessor to validly assess a colleague from a different culture? The MBI:TAC aims, within certain parameters, to be inclusive of different styles and approaches to the teaching. However, it is clear that we each have unconscious biases that will steer us to favor a certain style over another. Similarly, the question was raised: how possible is it for an assessor to validly assess an MBP curriculum with which they are not familiar? Within this training,

there was a mixture of MBSR and MBCT teachers participating, and a mixture of teaching clips from these programs working with different populations and within different contexts. It became clear that across the domains, there were program-specific issues that needed knowledge and experience with that program to enable accurate assessment. Some participants commented on this in the survey (see Table 4, theme 6).

During the training, the importance of the accuracy of benchmark ratings in training was highlighted. This also emerged as a theme from the survey (Table 4, theme 5). We recalibrated some of the benchmarks, and in the benchmarks used for the summary, interreliability engaged 2 further experienced MBP trainers so the benchmarks were the consensus of 4 rather than 2 assessors. These issues point to the emergent and early stage of this work. Some of the initial benchmarks were conducted within one training center and were not subject to peer review. The exposure within the training process to review from an international cohort of experienced teachers highlighted some flaws and biases in the benchmarking process. There was a particular vulnerability to bias when the assessors personally knew the teacher they were rating. In future, benchmarks that are being used to train assessors need to be subject to peer review from experienced teachers from differing training centers. Further qualitative work conducted on assessor training could usefully gather perspectives on the reasoning behind assessment points. This could help shed light on the reasons assessors divert from benchmarks and may provide valuable information for further adapting the instrument or the benchmarks. Further work could involve colleagues working with different populations, contexts, and countries with different MBP curricula, engaging collaboratively in selecting examples of teaching and developing benchmarks. For training purposes, this could include qualitative narratives giving detailed rationales anchored to the MBI:TAC key features for why particular ratings were given. Overall, there is the opportunity going forward for this process to be a vehicle for international collaboration and consensus building on the elements of MBP teaching integrity.

Retention in the training was excellent, with all participants completing the 7-week course, and 30/31 completing the homework for the last class. Of the 27 who were asked to complete final benchmark ratings, only 3 failed to do so, citing primarily lack of time. The 3 who did not complete the benchmark ratings did not differ from the other raters on any demographic, teaching, or meditation history characteristic.

Implications for Future Practice

This research informs 2 practice-related issues. First, developing greater understanding about how to

effectively train practitioners for the range of uses of the MBI:TAC. Second, developing greater clarity on good practice and governance issues needed to enable assessment of MBP teaching practice in teacher training and research contexts. The latter is addressed in a separate paper in this special issue (Crane RS, Koerbel L, Sansom S, Yiangou A. Assessing mindfulness-based teaching competency: Good practice guidance. *Glob Adv Ment Health*. Under review). These practical implementation questions also closely link with issues addressed in 2 other papers in this special issue. Griffith et al. examine how to skillfully integrate the MBI:TAC into MBP teacher training programs (with awareness of risks and cautions), and how to use it as a framework within which to offer feedback that supports learning and development (Griffith GM, Crane RS, Fernandez E, Giommi F, Herbet G, Koerbel L. Implementing the Mindfulness-Based Interventions: Teaching Assessment Criteria (MBI-TAC) in mindfulness teacher training programs. *Glob Adv Ment Health*. Under review).¹⁴ Evans et al. examine how the MBI:TAC can be used in MBP teacher supervision to support learning and reflective engagement (Evans A, Griffiths GM, Sansom S, Crane RS. Using the Mindfulness-Based Interventions: Teaching Assessment Criteria (MBI-TAC) in supervision. *Glob Adv Ment Health*. Under review). The next section draws out the key learning on how to train assessors to use the MBI:TAC reliably and with integrity.

Practitioners currently seek MBI:TAC training for 2 main purposes:

1. *A framework for reflection*: the tool is widely used as a map of the skills that the teacher cultivates during and beyond training. In a survey of international MBP trainers and supervisors (N = 79), 95% of respondents said that they use the MBI:TAC to enable trainee/supervisee reflection.² The MBI:TAC training for this function focuses on how to skillfully embed the tool as a guiding framework into personal development as well as in training and supervision, to support reflective practice and to structure feedback on teaching skills.

2. *An assessment tool*: The MBI:TAC is used for assessment:

- (i) To enable recognition and certification of their teaching competence at key points in the MBP teacher training journey—particularly at the end to enable graduation with a Certificate of Competence;
- (ii) Within research trials to enable a check on the integrity of the teaching or where fidelity is the subject of the research question.

The MBI:TAC training for this assessment function therefore focuses on how to skillfully use the tool to enable the practitioner to reliably assess teaching practice in others.

Like any tool, there are skillful and unskillful ways of using the MBI:TAC. For both the functions named above, the user of the tool needs to build appropriate skills, alongside understanding of limitations, cautions, “edges,” and good practice when implementing the MBI:TAC (see also Evans et al., under review and Griffith et al., under review). Theme 6 in the content analysis highlighted the need for training in using the MBI:TAC to support giving skilled narrative feedback. Following the PROMPT research, a collaboration between 3 university mindfulness-training centers engaged in the study has considered the outcomes from the research along with experience in the UK of delivering MBI:TAC training over the last 5 years. From these prototypes, a 3-level training process which progressively builds skills in the range of applications of the MBI:TAC has been developed and is continually evolving. Details of the training process, including learning outcomes and training methodologies, are available online.¹⁵

It has also become clear that the process of assessing MBP teaching is difficult and vulnerable making work for both assessor and assessed. Theme 4 in content analysis demonstrates the markers’ need for ongoing mentoring. Theme 2 reflects the need for further opportunity to share peer reflections, dialogue in small groups; and to do so with others working within a range of contexts and cultural settings. Ongoing collaboration between markers would also serve to moderate discrepancies in ratings, support interrater reliability, and reduce the potential for “drift” over time. Regular attendance at practice and peer reflection groups, such as those held by the “Support for Integrity in Teaching and Training” (SiTT)¹⁶ Community, is recommended good practice for MBI:TAC assessors.

Summary practice learnings that are informing next steps include:

- Some trainees are only interested in using the tool as a framework for reflection and therefore only need a short training (level 1 training);
- For those using the tool to assess, more time is needed to enable reflection on the assessments, and so to progressively build skills in the use of the tool—particularly to build reliability in assessment (level 2 training);
- On completion of the training, a check on participant reliability in using the tool is needed prior to using it in the field (level 3 training);
- Ensure that when using benchmarked assessments in the training, they include written comments offering a

rationale for each score so the trainees can see how they were arrived at. Preliminary evidence from other, as yet unpublished, data in the PROMPT trial also underlines the importance of benchmarks being agreed by 2 or more assessors;

- The research was not designed to enable us to understand whether reliability is stronger when assessors are matched to the curriculums they are experienced in delivering and to the language and culture. However, intuitively this seems to be best practice;
- Develop understanding of how the MBI:TAC can skillfully contribute to the sustainable development of MBPs in the international context through aligning adaptations of the MBI:TAC with culturally sensitive MBP adaptations and expanding capacity of assessors who can use translations of the MBI:TAC in their local language; and
- All participants appreciated the opportunity to learn together and dialogue about the issue of teaching and program quality. Therefore, it is vital to embed the principle in future trainings that the “how” of learning is as important as the content.

Conclusion

The MBP field is at an early stage in addressing the challenge of quality MBP implementation internationally. There is the risk in the transition from research to practice of an “implementation cliff” such that interventions are almost universally more effective in the context of research trials compared to “real-world” settings.^{17,18} One concern is that the promise of the research evidence does not become available to the public due to lack of internationally agreed systems for assuring acceptable levels of MBP teaching consistency and quality. This training was the first to bring together senior MBP teachers from a range of countries to examine the issue of MBP teaching standards, competence, and fidelity. The process underlined the ways that international collaboration on these issues informs a globally joined-up approach, which in turn enables the sustainable development of the field. Examination of these issues is at a germinal stage. The effort to reach interrater reliability is considerable, which poses significant challenges in implementing methods such as these. It may be that, over time, research studies such as this one provide stepping stones toward discovering other ways to assess teaching integrity which are less resource intense and more reliable. For example, we are currently developing a version of the MBI:TAC which is completed by course participants. In many ways, this study raises as many questions as it answers—questions at the interface between MBP research and practice.

The MBI:TAC is recognized by the international community of MBP trainers as accurately representing

the elements that make up the MBP teaching process. These MBP teaching characteristics laid out in the MBI:TAC also align well with the defining elements of MBP programs.⁸ These rubrics that support the building of alignment around defining anchor points for MBP programs and teaching characteristics are critical ingredients for supporting the field in the next phase of development. It is important to note, however, that the MBI:TAC only assesses one element of mindfulness-based teaching—that which is visible in the teaching space. There are many other important elements that inform integrity and quality including professional ethics and practice, theoretical knowledge, and teacher reflective skills. Therefore, its implementation needs to be part of a multifaceted approach to MBP teaching professional standards.

Given the popularity and expansion of MBP practice, it is important that there are ways of ensuring integrity and standards. Assessment of teaching quality is one part of ensuring quality in research and practice settings. In turn, the quality of the assessment process needs attention. Whatever assessment tool is employed, it is important that the user is trained to use it reliably. This research evidenced that it is possible to train a cohort of senior MBP teachers from a range of regions worldwide to align toward shared agreement about the elements of competent MBP teaching. The project informs next steps about how to deliver training for future MBI:TAC assessors. The training can be delivered online, making it accessible and enabling the outcomes of this work to contribute to MBP teaching quality globally.

Acknowledgments

The authors wish to thank the participants in the MBI:TAC training for assessors who gave their time for the training and follow-up assessment work.

Authors' Contributions

RC and WK conceived the training model. FMH and JB were awarded funding for the study and conceived and designed the study with input from RC and WK. GMG wrote the content analysis. RC wrote and edited the manuscript. All authors commented on the manuscript and approved the final version.

Declaration of Conflicting Interests


The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: RC directs a mindfulness training center which includes delivery of training to use the MBI:TAC and receives royalties for books on mindfulness. WK directs a mindfulness center and receives royalties for several books related to mindfulness.


Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Research reported in this publication was supported by the National Center for Complementary and Integrative Health of the National Institutes of Health under award number R34AT008948. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ORCID iDs

Rebecca Crane  <https://orcid.org/0000-0003-3605-0256>

Gemma M Griffith  <https://orcid.org/0000-0002-4192-8505>

Willem Kuyken  <https://orcid.org/0000-0002-8596-5252>

References

1. Crane RS, Soulsby JG, Kuyken W, Williams JMG, Eames C. The Bangor, Exeter & Oxford Mindfulness-Based Interventions: Teaching Assessment Criteria (MBI-TAC). <http://mbitac.bangor.ac.uk/>. Accessed February 21, 2020.
2. Russell S. *Investigating the Use of the Mindfulness-Based Interventions: Teaching Assessment Criteria (MBI:TAC) in Training Programmes Internationally* [master's thesis]. Bangor, UK: Bangor University; 2018.
3. Health Education England. Improving access to psychological therapies: Mindfulness-based cognitive therapy national MBCT training curriculum. <https://www.hee.nhs.uk/our-work/mental-health/improving-access-psychological-therapies>. Accessed November 15, 2019.
4. Segal Z, Williams M, Teasdale J, et al. MBCT international training pathway. <http://www.mbct.com/mbct-training-pathway.html>. Accessed November 15, 2019.
5. Crane RS, Hecht FM. Intervention integrity in mindfulness-based research. *Mindfulness*. 2018;9:1370–1380.
6. Hecht FM. Predictors of outcomes in MBSR participants from teacher factors (prompt-F), trial registration with national institutes of health. <https://clinicaltrials.gov/ct2/show/NCT03384771?term=Hecht+F&draw=2&rank=1>. Accessed February 21, 2020.
7. Kuyken W, Nuthall E, Byford S, et al. The effectiveness and cost-effectiveness of a mindfulness training programme in schools compared with normal school provision (MYRIAD): study protocol for a randomised controlled trial. *Trials*. 2017;18:194.
8. Crane RS, Brewer J, Feldman C, et al. What defines mindfulness-based programs? The warp and the weft. *Psychol Med*. 2017;47(6):990–999.
9. Crane RS, Eames C, Kuyken W, et al. Development and validation of the Mindfulness-Based Interventions: Teaching Assessment Criteria (MBI:TAC). *Assessment*. 2013;20(6):681–688.
10. Huijbers MJ, Crane RS, Kuyken W, et al. Teacher competence in mindfulness-based cognitive therapy for depression and its relation to treatment outcome. *Mindfulness*. 2017;8:960–972.
11. St. Laurent RT. Evaluating agreement with a gold standard in method comparison studies. *Biometrics*. 1998; 54(2):537–545.
12. Stemler S. An overview of content analysis. *Pract Assess Res Eval*. 2001;7(17).
13. Krippendorff K. *Content Analysis: An Introduction to its Methodology*. Newbury Park, CA: Sage; 1980.
14. Sansom S, Crane RS, Evans A, et al. A guide to skilful feedback in mindfulness-based programs training and assessment. <http://mbitac.bangor.ac.uk/documents/MBI-TAC-Feedback-Resource-2020-Final.pdf> (accessed April 17, 2020).
15. Crane RS, Koerbel L, Sansom S, Yiangou A. Training to use the Mindfulness-Based Interventions: Teaching Assessment Criteria (MBI:TAC). <http://mbitac.bangor.ac.uk/training.php.en>. Accessed November 15, 2019.
16. Sansom S. Join a peer group to reflect on your development as an MBI:TAC assessor. <http://mbitac.bangor.ac.uk/training.php.en>. Accessed February 21, 2020.
17. Weisz JR, Chorpita BF, Palinkas LA, et al. Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: a randomized effectiveness trial. *Arch Gen Psychiatry*. 2012; 69(3):274–282.
18. Dimidjian S, Segal ZV. Prospects for a clinical science of mindfulness-based intervention. *Am Psychol*. 2015; 70(7):593–620.