In a similar vein, collapsing all items from the FMI and the MAAS into one set of items did not improve the accuracy of class membership prediction. This fact warrants further elaboration as the scales are conceptually different. Whereas the FMI was designed to measure "mindfulness" on the basis of both an attention and an emotion facet, the MAAS taps into "mindlessness," thereby solely building upon a lack of attention as a unidimensional construct. However, further research is warranted to determine whether this result can be seen as supporting the notion that the acceptance/emotion facet is functionally redundant.

Our results also indicate that one particular aspect of the predictive accuracy of mindfulness instruments, namely sensitivity, may be particularly vulnerable to a lower sample quality. As sensitivity figures in the low-quality sample were barely above the base rates for both the MAAS and the FMI, it can be concluded that for both scales, the algorithm had trouble correctly identifying mindfulness practitioners. We must remember that the prediction scores from the RF analysis are based on OOB data (i.e., they are from a cross-validation part of the sample and not from the part of the sample that was used to build the model). However, as the algorithm achieved high specificity in Sample 1 and high overall accuracy in Sample 2, we suggest for the time being that this may be a specific problem that is associated with sensitivity rates.

Focusing on a lower level of measurement units, our results are also somewhat discouraging: Only about half of the items of both scales were able to add to the predictive accuracy in the high-quality sample, whereas in the low-quality sample, only one or two items contributed to subgroup differentiation. This finding is disappointing as it suggests that many of the items that we investigated are in need of revision. For example, FMI Item 13, an item found to be plagued

by psychometric problems in previous analyses (Sauer, Walach, Offenbächer, Lynch, & Kohls, 2011b; Sauer, Ziegler, et al., 2013), did not contribute to predictive accuracy. Nevertheless, as several other items that were not identified as problematic by previous analyses showed similar results, no strong conclusions can be drawn. With regard to the MAAS, there were also problems with several mindfulness items. This interpretation is corroborated by Van Dam, Earleywine, and Borders' (2010) findings, which suggested that only five of the 15 items were informative. Interestingly, four of these five items were also identified by our analysis.

In sum, whereas our study seems to have revealed several shortcomings of two established mindfulness instruments—one of which we had the pleasure to contribute to substantially—a word of caution is needed. First, we used a novel approach, with which there is naturally a lack of experience and comparability. Second, our dependent variable—mindfulness practice or lack thereof—is *ipso facto* not the only aspect that mindfulness instruments should be able to differentiate. Although we decided to employ this binary selection criterion for pragmatic reasons, it may well be the case that, for other dependent variables (e.g., psychological health, well-being), the predictive accuracy may be different. Nevertheless, even when applying a seemingly straightforward binary criterion, problems naturally associated with the self-attribution of inner states can arise: Although response shift and other form of biases may actually impair sensitivity indices, it may also be argued that not all mindfulness practitioners have higher (self-reported) mindfulness levels than nonpractitioners. We opine that it would therefore appear naïve to assume clear-cut, homogeneous results—even when a powerful procedure such as random forests is applied. Nonetheless, it may well be the case that response shift and other forms of biases relevant for assessing self-attributed mindfulness levels may

occur as a consequence of a regular mindfulness practice. Our data may suggest that some mindfulness practitioners may actually tend to report lower self-attributed mindfulness scores, probably as a consequence of their sensitivity to states of mindlessness. We believe that there are presently only two remedies to that problem: (a) Researchers should strive to implement well-controlled experimental settings, and (b) researchers should endeavor to employ supplementary measurement approaches in addition to the quantitative measurement paradigm such as bistable images (Sauer et al., 2012).

In summary, the results shed new light on the quality of the instruments. It appears that many of the MAAS and FMI items are unable to reliably predict class membership. A substantial need for revision may be deduced. However, both instruments showed good predictive accuracy overall—the probability of correctly classifying an individual as a nonpractitioner was twice as high as chance would suggest.

References

Abu-Nimeh, S., Nappa, D., Wang, X., & Nair, S. (2007). A comparison of machine learning techniques for phishing detection. Proc eCrime Res Summit 2007.

Bohlmeijer, E., Prenger, R., Taal, E., & Cuijpers, P. (2010). The effects of mindfulness-based stress reduction therapy on mental health of adults with a chronic medical disease: A meta-analysis. J Psychosom Res., *68*(6), 539–544.

Bosch, A., Zisserman, A., & Muoz, X. (2007). Image classification using random forests and ferns. In Computer Vision, 2007 ICCV . Rio de Janeiro.

Breiman, L. (2001). Random forests. Mach Learn., *45*(1), 5–32.

Chadwick, P., Hughes, S., Russell, D., Russell, I., & Dagnan, D. (2009). Mindfulness groups for distressing voices and paranoia: a replication and randomized feasibility trial. Behav Cogn Psychother., *37*, 403–412. doi:10.1017/S1352465809990166

Gaylord, S. A., Palsson, O. S., Garland, E. L., Faurot, K. R., Coble, R. S., Mann, J. D., … Whitehead, W. E. (2011). Mindfulness Training Reduces the Severity of Irritable Bowel Syndrome in Women: Results of a Randomized Controlled Trial. Am J Gastroenterol., *106*(9), 1678–1688.

Grossman, P. (2011). Defining Mindfulness by How Poorly I Think I Pay Attention During Everyday Awareness and Other Intractable Problems for Psychology's (Re)Invention of Mindfulness: Comment on Brown et al. (2011). Psychol Assess., *23*(4), 1034–1040.

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. J Comput Graph Stat., *15*(3), 651–674.

Kohls, N., Sauer, S., & Walach, H. (2009). Facets of mindfulness–Results of an online study investigating the Freiburg Mindfulness Inventory. Pers Individ Dif., *46*(2), 224–230.

Leigh, J., Bowen, S., & Marlatt, G. A. (2005). Spirituality, mindfulness and substance abuse. Addict Behav., *30*(7), 1335–1341. doi:10.1016/j.addbeh.2005.01.010

Mars, T. S., & Abbey, H. (2010). Mindfulness meditation practise as a healthcare intervention: A systematic review. Int J Osteopath Med., *13*(2), 56–66. doi:10.1016/j.ijosm.2009.07.005

R-Core-Team. (2008). R: A language and environment for statistical computing. R Found Stat Comput., *3*(10).

Sauer, S., Lemke, J., Wittmann, M., Kohls, N., Mochty, U., & Walach, H. (2012). How long is now for mindfulness meditators? Pers Individ Dif., *52*(6), 750–754.

Sauer, S., Walach, H., Offenbächer, M., Lynch, S., & Kohls, N. (2011a). Measuring Mindfulness: A Rasch Analysis of the Freiburg Mindfulness Inventory. Religions., *2*(4), 693–706.

Sauer, S., Walach, H., Offenbächer, M., Lynch, S., & Kohls, N. (2011b). Measuring mindfulness: a Rasch analysis of the Freiburg mindfulness inventory. Religions., *2*(4), 693–706.

Sauer, S., Walach, H., Schmidt, S., Hinterberger, T., Horan, M., & Kohls, N. (2011). Implicit and explicit emotional behavior and mindfulness. Conscious Cogn., *20*(4), 1558–1569.

Sauer, S., Walach, H., Schmidt, S., Hinterberger, T., Lynch, S., Büssing, A., & Kohls, N. (2013). Assessment of mindfulness: review on state of the art. Mindfuln., *4*(1), 3–17.

Sauer, S., Ziegler, M., Danay, E., Ives, J., & Kohls, N. (2013). Specific Objectivity of Mindfulness—A Rasch Analysis of the Freiburg Mindfulness Inventory. Mindfuln., *4*(1), 45–54.

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics., *8*(1), 25. doi:10.1186/1471-2105-8-25

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests. Psychol Methods., *14*(4), 323–348.

Van Dam, N. T., Earleywine, M., & Borders, A. (2010). Measuring mindfulness? An Item Response Theory analysis of the Mindful Attention Awareness Scale. Pers Individ Dif., *49*(7), 805–810. doi:10.1016/j.paid.2010.07.020

Walach, H., Buchheld, N., Buttenmüller, V., Kleinknecht, N., & Schmidt, S. (2006). Measuring mindfulness—the Freiburg Mindfulness Inventory (FMI). Pers Individ Dif., *40*(8), 1543–1555.

Walach, H., Ferrari, M.-L. G., Sauer, S., & Kohls, N. (2012). Mind-Body Practices in Integrative Medicine. Religions., *3*(1), 50–81. doi:10.3390/rel3010050