Personality and Individual Differences 81 (2015) 117–123

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Mindful in a random forest: Assessing the validity of mindfulness items using random forests methods ☆

Sebastian Sauer [a,b,*], Jana Lemke [b,c], Winfried Zinn [d,e], Ricardo Buettner [a], Niko Kohls [b,f]

[a] FOM University of Applied Sciences, Essen, Germany
[b] Samueli Institute, Alexandria, USA
[c] Viadrina University, Frankfurt (Oder), Germany
[d] Forschungsgruppe Metrik, Bermuthshain, Germany
[e] Applied University for Health and Sports, Berlin, Germany
[f] Coburg University of Applied Sciences, Coburg, Germany

ARTICLE INFO

Article history:
Received 30 April 2014
Received in revised form 2 September 2014
Accepted 7 September 2014
Available online 27 September 2014

Keywords:
Mindfulness
Assessment
Measurement
Psychometrics
Random forests

ABSTRACT

Whereas the number of studies supporting the efficacy of mindfulness as a health intervention is increasing, the measurement of mindfulness remains a subject of debate. Given the importance of measurement in this field, this paper aims to further our understanding of the assessment of mindfulness by employing an approach referred to as "random forests" (RF). RF is an ensemble learning method that is based on decision trees. RF is well known in biological research, for example, but is practically unknown in psychometrics. In this study, RF was used to gauge the predictive validity of the items from two mindfulness instruments concerning their ability to estimate group allocation (i.e., mindfulness practitioners vs. nonpractitioners). To allow for a better generalization of the results, we examined the research questions in two samples ($N = 76$ and $N = 202$) of different quality. We investigated two instruments: the Freiburg Mindfulness Inventory (FMI) and the Mindfulness Attention and Awareness Scale. Although results indicated that both instruments were capable of distinguishing practitioners from nonpractitioners, the predictive quality of most items on both scales was determined to be insufficient.

1. Introduction

Mind full or mindful? This pun speaks of complementary states of mind. A "full mind" can be associated with cognitive processes, subserving forms of mind-wandering such as preoccupied thinking (e.g., resentment or contentment regarding past experiences and desire or dissatisfaction regarding anticipated future events). Two characteristics of mind-wandering are that they are (a) either past- or future-centered, and correspondingly not anchored in the experience of the present moment, and (b) invoke some kind of predominantly negative emotional reaction such as displeasure or distress. By contrast, mindfulness can be conceived as a mental state in which attention is systematically focused in the present moment with a stance of equanimity such as openness, curiosity, and nonjudgmental awareness (Walach, Ferrari, Sauer, & Kohls, 2012).

The hypothesis that mindfulness may increase health and well-being has been empirically corroborated in recent years. Several (meta-analytical) reviews have documented its clinical effectiveness (e.g., Mars & Abbey, 2010). Especially for psychosomatic diagnoses (e.g., anxiety and stress), mindfulness-based interventions have been found to be effective.

Despite promising empirical observations concerning the effects of mindfulness, its measurement has been a target of criticism (Grossman, 2011). As measurement approaches for mindfulness differ considerably from a conceptual point of view (Sauer et al., 2013), it seems appropriate to state that research has thus failed to provide a theoretically accepted notion of how mindfulness should be measured.

It is therefore necessary for the advancement of this field to improve its current state of measurement. Hence, it is the aim of the present research to contribute to improving the measurement of mindfulness by scrutinizing the predictive ability of mindfulness scales to gauge whether or not a given individual practices mindfulness training or not. The rationale behind this idea is that an individual with a regular practice of mindfulness should report a higher (self-reported) mindfulness level than an individual who

☆ This article is a Special issue article – "Young researcher award 2014".

* Corresponding author at: Institute of Management & Information Systems, FOM University of Applied Sciences, Leimkugelstraße 6, 45141 Essen, Germany. Tel.: +49 89/202452 27.
E-mail address: Sebastian.sauer@fom.de (S. Sauer).

http://dx.doi.org/10.1016/j.paid.2014.09.011
0191-8869/

does not regularly practice mindfulness. Although it has occasionally been proposed that mindfulness practitioners may describe themselves as less mindful than individuals naïve to the idea of mindfulness due to a shifting of the internal reference point, the majority of published experimental studies suggest that regular mindfulness training leads to higher self-reported mindfulness levels (Bohlmeijer, Prenger, Taal, & Cuijpers, 2010; Chadwick, Hughes, Russell, Russell, & Dagnan, 2009; Gaylord et al., 2011; Mars & Abbey, 2010). This is not to say that there are no other sizeable problems associated with the quantitative assessment of mindfulness: For example, a positive relation between mindfulness levels and smoking/frequent binge-drinking behavior was found (Leigh, Bowen, & Marlatt, 2005), suggesting that mindfulness instruments may actually tap into other constructs such as conscious sensitivity or bodily awareness. Keeping this limitation in mind, the ability to estimate group allocation can naturally not be expected to be perfect.

We employed a statistical procedure referred to as "random forests" (RF). Although RF is widely used in biologically driven research (Bosch, Zisserman, & Muoz, 2007), it is still scarcely used in psychological studies (but see Strobl, Malley, & Tutz, 2009). Evidence supports the predictive strength of the method (Abu-Nimeh, Nappa, Wang, & Nair, 2007). We submitted data from two mindfulness instruments to an RF procedure in order to gauge the predictive quality of the respective instrument (i.e., predicting class membership of the binary criterion "regular mindfulness practice or not").

2. Methods

2.1. Samples

Two samples were included in this study to allow for replication of the results and in order to investigate the influence of data quality. The total sample size was $N = 278$. Sample 2 was of higher quality and Sample 1 of lower quality (see below).

2.1.1. Sample 1

Sample 1 ($N = 202$) was collected as part of an unpublished online study investigating the relation between mindfulness, health, and emotion; these data have not been published before. About two thirds of the sample ($n = 129$) reported having no prior mindfulness practice, whereas $n = 72$ individuals (36%) reported practicing mindfulness on a regular basis (1 missing value). Examples of types of mindfulness training included Buddhist meditation, Thai Chi, or Yoga exercises. The mean age was 35 years for nonpractitioners ($SD = 13$) and 39 years for practitioners ($SD = 11$); $n = 142$ (71%) persons were female, and $n = 59$ (29%) persons were male (1 missing value).

2.1.2. Sample 2

Sample 2 ($N = 76$) was described in Sauer et al. (2012). The sample consisted of 38 expert mindfulness practitioners (21 female, 17 male) and an age- and gender-matched group of 38 nonmindfulness practitioners (28 female, 10 male). Practitioners had trained in different Buddhist mindfulness traditions. The mean age was 51 years in both groups ($SD = 10$). Inclusion criteria were at least 5 years of daily meditation practice for the meditation group and no meditation experience for the control group.

2.2. Instruments

2.2.1. Freiburg Mindfulness Inventory (FMI-14)

The conceptual foundation of the FMI-14 is rooted in Buddhist Psychology (Walach, Buchheld, Buttenmüller, Kleinknecht, & Schmidt, 2006), and the instrument was designed to measure mindfulness as a stable trait. Whereas the scale was first developed as a unidimensional scale, recent research supported a two-factor solution (Kohls, Sauer, & Walach, 2009). The two factors have been labeled "Presence," indicating the awareness of stimuli in the subjective now (Sauer et al., 2012), and "Acceptance," indicating a nonjudgmental stance toward all kinds of experience (Kohls et al., 2009). One strength of the instrument is that it has been validated using not only classical psychometric methods such as exploratory and confirmatory factor analysis (Kohls et al., 2009) but also item response theory (Sauer, Walach, Offenbächer, Lynch, & Kohls, 2011a; Sauer, Ziegler, Danay, Ives, & Kohls, 2013). The instrument consists of 14 items with four answer options.

2.2.2. Mindfulness Attention and Awareness Scale (MAAS)

The MAAS is one of the most widely used instruments measuring mindfulness as a unidimensional construct with 15 items and six answer options. It is noteworthy that all items are negatively formulated to assess "mindlessness" rather than "mindfulness." Although a study suggested that mindlessness should not be conceived as the inverted construct of mindfulness (Van Dam, Earleywine, & Borders, 2010), there are several studies supporting the practical usefulness of the MAAS. We deem the MAAS an adequate candidate for our analysis in addition to the FMI, given that data support its usefulness.

2.3. Statistical method – random forests

We employed RF, a statistical method stemming from machine learning contexts to assess the predictive accuracy of the FMI-14 or the MAAS items with regard to their ability to estimate class membership as operationalized by a regular mindfulness practice or a lack thereof (i.e., mindfulness practice as a binary criterion). RF allows statistical properties to be delineated (e.g., nonlinear trends, high-degree interaction, and correlated predictors). In addition, assumptions necessary for classical multivariate analyses such as homoscedasticity (homogeneity of variance), linear associations between variables, or metric variable levels are not warranted (Breiman, 2001).

The term "random forests" is well chosen because in the respective procedure, random subsamples of decision trees are drawn, building together a "forest" of "trees" on a random basis. The RF procedure therefore represents an ensemble, averaging the results of several decision trees based on a majority vote principle. RF predictions of group or class membership are based on the part of the sample that was not used to build the RF framework. This excluded part of the sample is called the "out of bag" (OOB) sample. The advantage of the OOB procedure is that the results can be seen as stemming from a cross-validation sampling procedure, thereby increasing confidence in the results (Breiman, 2001).

2.4. Procedure

We conducted four analyses that could provide answers to the following four pivotal research questions. First, as a baseline test, we wanted to investigate if the group mean value can be seen as a suitable indicator for distinguishing between mindfulness practitioners and non-practitioners. Correspondingly, for each of the two instruments, total mean scale scores were calculated and compared. For the FMI, we additionally computed the mean scores for both factors (i.e., *Presence* and *Acceptance*). We wanted to see whether the mere aggregation of item mean scores into a total mean scale score would contain enough information to allow us to differentiate between the subgroups (practitioners vs. nonpractitioners). We expected that practitioners would exhibit higher mindfulness levels than nonpractitioners.