ORIGINAL PAPER

Validation of the German Version of the Southampton Mindfulness Questionnaire (SMQ)

Kerem Böge[1] · Carmen Schaeuffele[2] · Pamela Jacobsen[3] · Paul Chadwick[3] · Emre Ergen[1] · Inge Hahne[1] ·
Niklas Bergmann[1] · Johanna Boettcher[2,4] · Katja Wingenfeld[1] · Malek Bajbouj[1] · Eric Hahn[1]

Abstract

Objectives The aim of the study was to assess the convergent and divergent validity, reliability, utility, and treatment sensitivity of a newly translated German version of the Southampton Mindfulness Questionnaire (SMQ). The SMQ is a 16-item instrument measuring mindful awareness of distressing thoughts, images, and perceptions, developed originally within the mindfulness for psychosis field.

Methods Overall, three studies were conducted, comprising (1) a non-clinical sample of $n = 848$ (638 community sample and 210 meditators); (2) a clinical sample of $n = 213$ (106 schizophrenia and 107 depression); and (3) a clinical sample with $n = 122$ participants with emotional disorders within a randomized controlled study, of which 30 participants were also included in study 2. To assess convergent validity, participants completed the SMQ, Freiburg Mindfulness Inventory (FMI), and Comprehensive Inventory of Mindfulness Experiences (CHIME). To measure divergent validity, participants completed the Brief Symptom Inventory 18 (BSI-18), Positive and Negative Affect Schedule (PANAS), Brief Experiential Avoidance Questionnaires (BEAQ), and Anxiety Sensitivity Index 3 (ASI-3).

Results Mean internal consistency ($\alpha = 0.89$) and convergent ($r = 0.66$ to 0.73) and divergent validity ($r = -0.09$ to -0.50) were established and sensitivity to change over time following treatment ($d = 0.86$) was shown. For the clinical sample, a single-factor structure is suggested by principal component analysis.

Conclusions Results provide first evidence for the utility of the German version of the SMQ for clinical practice and research in healthy individuals, meditators, and clinical groups. Further research is needed to examine the underlying construct of mindfulness.

Keywords Validation · Southampton Mindfulness Questionnaire · SMQ · Mindfulness · Schizophrenia · Depression

Converging evidence for the effectiveness of mindfulness practices demonstrates the importance of developing novel, reliable, and valid tools to measure mindfulness states and

✉ Kerem Böge
kerem.boege@charite.de

[1] Department of Psychiatry and Psychotherapy, Campus Benjamin Franklin, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Hindenburgdamm 30, 12203 Berlin, Germany

[2] Department of Clinical Psychology and Psychotherapy, Freie Universität Berlin, Berlin, Germany

[3] Department of Psychology, University of Bath, Bath, UK

[4] Clinical Psychology and Psychotherapy, Psychologische Hochschule Berlin, Berlin, Germany

traits (Baer 2007). In several English-speaking countries, a variety of mindfulness questionnaires and tools already exist (Baer et al. 2004, 2006; Brown and Ryan 2003; Feldman et al. 2007). In Germany, however, only a handful of well-validated self-report mindfulness questionnaires (Mindful Attention and Awareness Scale (MAAS), Freiburg Mindfulness Inventory (FMI), Kentucky Inventory of Mindfulness Skills (KIMS), Five Facet Mindfulness Questionnaire (FFMQ), and Comprehensive Inventory of Mindfulness Experiences CHIME)) are available and these show various limitations (Heidenreich et al. 2006). These limitations comprise comprehension difficulties for several items that were misunderstood by mindfulness inexperienced participants due to their complex wording, as reported for the FMI (Belzer et al. 2013) and FFMQ (Christopher et al. 2014). Furthermore, the MAAS was criticized for assessing the *construct* of mindlessness instead

of mindfulness (Sauer et al. 2011). The CHIME avoids complex wording and assesses eight related domains of mindfulness. It comprises 37 items and is therefore considerably longer than other mindfulness questionnaires. Furthermore, the CHIME, similar to the other mindfulness questionnaires, does not provide a neutral answer opportunity ("undecided"). This might be overstraining, especially for clinical populations that frequently experience cognitive deficits and might not understand all items. For this reason, it is crucial to create psychometrically sound and accurate measures of mindfulness that can facilitate a deepened understanding of its underlying mechanisms as well as the psychological and therapeutic processes involved (Bishop et al. 2004; Brown and Ryan 2004; Dimidjian and Linehan 2003). This is particularly relevant for research in emerging areas, such as mindfulness for psychosis, which currently has a less well-established evidence base compared to other disorders, such as mindfulness-based cognitive therapy (MBCT) for depression.

The Southampton Mindfulness Questionnaire (SMQ) is a 16-item inventory assessing the degree to which individuals mindfully respond to distressing thoughts, images, and perceptions that constitute key psychopathological phenomena in various clinical disorders (Chadwick et al. 2008). It was specifically designed to support the outcome and research process in mindfulness for psychosis, although it is applicable transdiagnostically and in non-clinical as well as clinical populations. It consists of four related bipolar constructs converged to a unidimensional trait-like tendency to be mindful in daily life (Chadwick et al. 2005, 2008). Those components are (1) decentered awareness (Safran and Segal 1990) of cognitions vs. being lost in reacting to them, (2)non-aversion vs. experiential avoidance (Chadwick et al. 2008), (3)nonjudgment and acceptance of difficult thoughts of oneself (Kabat-Zinn 1990, p. 69), and (4) letting difficult cognitions pass without rumination and worry (Chadwick et al. 2008). By including multiple facets of mindfulness constructs, on the one hand, the SMQ covers not all but a wide range of aspects when compared to the MAAS (Grossman 2011; Höfling et al. 2011). On the other hand, the scale also retains its simplicity due to its single-factor structure (Bergomi et al. 2013).

The SMQ has thus been selected for German translation and adaptation because of its suitability for assessments in clinical practice and research by focusing on the mindful relationship with distressing thoughts and images—a central phenomenon in clinical disorders such as schizophrenia or depression (Bergomi et al. 2013; Chadwick et al. 2008). Therefore, the SMQ might be suitable to measure mindfulness as an underlying mechanism of change in psychotherapy (Bergomi et al. 2013). Chadwick et al. (2008) have shown that the scale has good convergent validity with the MAAS ($r = 0.61$), divergent validity between meditators, non-meditators, and a clinical group of people with schizophrenia as well as internal consistency ($\alpha = 0.89$). Furthermore, the SMQ

features homogenous item interpretation for meditators and non-meditators(Chadwick et al. 2008). In contrast to other available mindfulness measures, the SMQ provides a neutral answer option (3 "undecided") and has a wide distribution of response categories ranging from 0 "disagree totally" to 6 "agree totally." This allows for more diverse response patterns, which might be explicitly valuable in clinical populations. Henceforth, the translated and validated SMQ might constitute a useful transdiagnostic tool for the assessment of mindfulness in psychotic and affective disorders, which are both among the three most prevalent psychiatric conditions in German hospitals (DGPPN 2018).

The overall aim of the present study was to translate the SMQ into German, and to examine its validity, reliability, utility, and treatment sensitivity. This validation study comprises three substudies: study 1 focused on the validation of the German SMQ within a non-clinical sample, comparing meditators to non-meditators, recruited for an online study from the general population. Study 2 focused on the validation within a clinical sample of people diagnosed with either depression or schizophrenia, recruited from inpatient and community mental health services. Finally, study 3 focused on the sensitivity to change over time of the German SMQ, using data from a randomized controlled trial of an online psychological therapy for people with common mental health problems.

Study 1

The Southampton Mindfulness Questionnaire (SMQ, Chadwick et al. 2005) is a 16-item scale assessing the degree to which individuals mindfully respond to distressing thoughts and images (Chadwick et al. 2005). The questionnaire comprises eight positively and negatively worded items rated on a 7-point fully anchored Likert scale (*agree totally* (*6*) to *disagree totally* (*0*)), yielding a total score range of 0–96. The SMQ is designed to assess four related aspects of mindfulness that are constructed as bipolar items: (1) decentered awareness, (2) letting go, (3) non-judgment, and (4) non-aversion, but exploratory factor analyses recommended a onedimensional factor structure (Chadwick et al. 2005, 2008). All items are introduced with the phrase "Usually, when I have distressing thoughts or images …" and continue with a mindfulness-related response, such as "I judge the thought/image as good or bad" or "I am able just to notice them without reacting." The SMQ validated in this publication was first translated by two German native speaking psychologists from English to German. Next, two independent English native speaking psychologists translated both versions to English, employing current back translation standards (WHO 2014). Pilot testing was performed by three independent samples of people with schizophrenia, depression, and a healthy control

group. Lastly, all four versions were checked according to the World Health Organization (2014) guidelines for the "Process of translation and adaptation of instruments" to ensure the equivalence of the questionnaires. In study 1, it was aimed to determine the reliability, factor structure, and convergent validity of the SMQ in a non-clinical sample. Therefore, a large non-clinical community sample comprising regular meditators as well as non-meditators was recruited. In the original validation study, the SMQ depicted an overall single-factor structure even though the measure consists of four related constructs (Chadwick et al. 2008). Moreover, in a non-clinical community sample, the SMQ depicted good reliability and convergent validity with other mindfulness measures (Baer et al. 2006; Chadwick et al. 2008). Given the emphasis in the literature, the current trial aimed to replicate these findings in a larger sample in German language.

Method

Participants

In the study, 210 meditators (162 women, 47 men, 1 non-binary) with a mean age of 39.57 ($SD = 12.34$; range = 21–68) participated. The majority of the meditating sample was German ($n = 166$), single ($n = 124$) or married ($n = 50$), held a university degree ($n = 112$), completed an apprenticeship ($n = 26$), or reported a high-school graduation as their highest degree ($n = 36$). Most of the meditating participants were currently employed ($n = 81$), self-employed ($n = 53$), or studying ($n = 51$). Participants were included as meditators if they reported a mediation practice of at least 1 h per week over the last month in order to recruit a sample with a wide spectrum of meditation experience. On average, meditators had 60.17 ($SD = 90.18$; range 1–559) months of meditation experience and meditated on average 4.72 ($SD = 5.37$, range 1–36 h) hours per week. Out of these, 46 participants had meditation experience of 6 months or less ($M = 49.94$, $SD = 16.21$), 60 were meditating for a period of 6 months to 2 years ($M = 55.57$, $SD = 12.89$), 21 between 2 and 3 years ($M = 65.43$, $xSD = 13.35$), and 83 were meditating for more than 3 years prior to the study ($M = 61.9$, $SD = 14.12$). The community sample comprised 638 participants (549 women, 85 men, 4 non-binary) with a mean age of 30.26 ($SD = 10.81$; range 18–76). Participants from this sample were mostly German ($n = 507$), single ($n = 498$) or married ($n = 111$), and held a high-school degree ($n = 242$) or a university degree as their highest education ($n = 292$). The majority of participants in this sample were currently employed ($n = 157$) or studying ($n = 389$). Out of the undergraduates and graduates, the majority were students of humanities ($n = 360$), of which $n = 332$ were psychology, or medicine students ($n = 87$).

Procedure

The non-clinical sample was recruited via an online questionnaire created with Unipark Software Questback. Via various routes, students and alumni from universities in German-speaking countries (Germany, Austria, Switzerland) were approached via mailing lists with a focus on psychology, medicine, and other closely related fields. In addition, meditation and mindfulness centers throughout German-speaking countries were approached with the request to forward the study.

Measures

In addition to completing the newly translated German version of the SMQ (see Table 1), participants completed 2 additional mindfulness measures for the purposes of assessing convergent validity.

Freiburg Mindfulness Inventory Short Form (Buchheld et al. 2001) The construction of the Freiburg Mindfulness Inventory Short Form (FMI) was particularly inspired by the Buddhist roots of the construct and consists of 30 items assessing non-judgmental present-moment observation and openness to negative experiences (Buchheld et al. 2001). The FMI short form is a robust 14-item instrument, which is semantically independent of a Buddhist context, but still covers all aspects of mindfulness, making it more appropriate for use in the general and clinical population (Walach et al. 2006). Items are rated on a 4-point Likert-type scale (rarely (1) to almost always (4)), yielding a total range of 14–56. It has a two-dimensional structure, comprising an acceptance and a presence factor (Heidenreich et al. 2006; Kohls et al. 2009). Despite the efforts to remove the close semantic Buddhism link, a qualitative analysis showed that individuals without meditation experience systematically misunderstood three items (Belzer et al. 2011). In the current study, the Cronbach's α of the total sample was 0.88.

Comprehensive Inventory of Mindfulness Experiences (Bergomi et al. 2014) The Comprehensive Inventory of Mindfulness Experiences (CHIME) is a multidimensional questionnaire providing broad coverage of mindfulness. It comprises eight subscales: inner awareness, outer awareness, acting with awareness, openness to experiences, accepting and non-judgmental orientation, decentering and non-reactivity, insightful understanding, and relativity of thoughts. The factor structure proved to be stable over three samples and validity analyses provided good results (Bergomi et al. 2014). It comprises 37 items, scored on a 6-point Likert-type scale (almost never (1) to almost always (6)), yielding a total range of 37–222. The construction of the CHIME puts a major focus on items being as semantically precise as possible for different

Table 1 Southampton Mindfulness Questionnaire–Deutsch (SMQ-D)

SMQ-D: Southamptom Mindfulness Questionnaire – Deutsch

Stellen Sie sich eine Situation, in der Sie anstrengende Vorstellungen, Gedanken oder Bilder im Kopf haben, vor und geben Sie an, wie Sie **normalerweise** auf diese reagieren:

Meistens, wenn ich beunruhigende Vorstellungen, Bilder oder Gedanken im Kopf habe...		Trifft voll und ganz zu	Trifft zu	Trifft teilweise zu	Unsicher	Trifft wenig zu	Trifft nicht zu	Trifft überhaupt nicht zu
1	Kann ich sie wahrnehmen, ohne zu reagieren.	☐	☐	☐	☐	☐	☐	☐
2	Beschäftigen sie mich noch eine gewisse Zeit.	☐	☐	☐	☐	☐	☐	☐
3	Empfinde ich die Vorstellungen/Gedanken/Bilder als gut oder schlecht.	☐	☐	☐	☐	☐	☐	☐
4	Fühle ich mich bald danach ruhig/entspannt.	☐	☐	☐	☐	☐	☐	☐
5	Bin ich in der Lage, das Empfundene zu akzeptieren.	☐	☐	☐	☐	☐	☐	☐
6	Werde ich wütend, wenn es auftritt.	☐	☐	☐	☐	☐	☐	☐
7	Nehme ich wahr, wie kurz die Gedanken und Bilder tatsächlich sind.	☐	☐	☐	☐	☐	☐	☐
8	Beurteile ich mich abhängig von den Vorstellungen/Bildern/Gedanken selbst als gut oder schlecht.	☐	☐	☐	☐	☐	☐	☐
9	Trete ich „einen Schritt zurück" und bin mir der Gedanken, Vorstellungen oder der Bilder bewusst, ohne dass sie mich kontrollieren.	☐	☐	☐	☐	☐	☐	☐
10	Nehme ich sie wahr und lasse sie vorbeiziehen.	☐	☐	☐	☐	☐	☐	☐
11	Akzeptieren ich mich selbst, egal worüber die Gedanken oder Bilder handeln.	☐	☐	☐	☐	☐	☐	☐
12	Versuche ich, sie aus meinem Kopf zu verdrängen.	☐	☐	☐	☐	☐	☐	☐
13	Denke ich weiterhin über die Vorstellungen, Gedanken oder Bilder nach, nachdem sie vorbei sind.	☐	☐	☐	☐	☐	☐	☐
14	Empfinde ich sie so unangenehm, dass ich mich ablenken muss, um sie nicht zu beachten.	☐	☐	☐	☐	☐	☐	☐
15	Versuche ich, die Vorstellungen, Gedanken oder Bilder wahrzunehmen, ohne sie zu bewerten.	☐	☐	☐	☐	☐	☐	☐
16	Verliere ich mich in den Vorstellungen/Gedanken/Bildern.	☐	☐	☐	☐	☐	☐	☐

Table 1 (Continued)

Hintergrund	Der Grundgedanke hinter der Messung ist, dass es vier zusammen-hängende Facetten der Achtsamkeit gibt: Beobachten, Nicht-reaktivität, Bewusstsein für schwierige Erfahrungen öffnen, und Akzeptieren ohne Bewertung.

Der Grundgedanke hinter der Messung ist, dass es vier zusammen-hängende Facetten der Achtsamkeit gibt: Beobachten, Nicht-reaktivität, Bewusstsein für schwierige Erfahrungen öffnen, und Akzeptieren ohne Bewertung.

Jede Subskala besteht aus jeweils vier Items. Beobachten besteht aus drei positiv behafteten Items (Zustimmung deutet auf höhere Achtsamkeit hin) und einem negativen (Zustimmung deutet auf niedrigere Achtsamkeit hin); Bewusstsein besteht wiederum aus drei negativen Items, um diesem entgegenzuwirken. Akzeptieren und Nichtreaktanz haben beide jeweils zwei positive und zwei negative Items.

Die vier Subskalen bestehen aus den folgenden Items, mit positiver und negativer Formulierung:

	Positiv	Negativ
Beobachten	1, 7, 9	16
Nichtreaktivität	4, 10	2, 13
Bewusstsein	5	6, 12, 14
Akzeptieren	11, 15	3, 8

Auswertung

- Trifft voll und ganz zu (Punktzahl 6), trifft zu (Punktzahl 5), trifft teilweise zu (Punktzahl 4), unsicher (Punktzahl 3), trifft wenig zu (Punktzahl 2), trifft nicht zu (Punktzahl 1), trifft überhaupt nicht zu (Punktzahl 0).
- Die negativen Items werden umgekehrt ausgewertet.
- Der SMQ hat Punktzahlen von 0 bis 96 (die vier Skalen können verschiedene Ansatzpunkte für Interventionen identifizieren und sind nützlich im Unterrichten. Jedoch sollten sie nur im Kontext gesehen und nicht unabhängig betrachtet werden.

population subgroups by avoiding ambiguous words such as meditation-related idioms. The Cronbach's α for the total sample in this study was 0.92.

Data Analyses

All analyses were carried out using IBM SPSS Statistics 24 and R version 3.6.0 including the packages "semTools" (Jorgensen et al. 2020) and "lavaan" (Rosseel 2012). Outliers were identified using boxplots and stem-and-leaf plots on SMQ total scores as well as subscores, and were removed from statistical analyses if they were outside the ± 1.5 interquartile range (IQR). Individuals were outside this 1.5 IQR if their SMQ score was ≤ 5 or ≥ 91. Furthermore, one individual was removed with a score of ≥ 23 on the SMQ subdomain *letting go*. Consequently, for the analysis of the total sample, five outliers were excluded; for the analysis of the non-clinical sample, four; and for the analysis of the

clinical sample, one outlier was removed. Kolmogorov–Smirnov test of normality was applied with additional visual inspection of histogram and Q-Q plot to assess normality. Each item's psychometric performance was assessed by response frequencies and internal consistency was determined by calculation of Cronbach's alpha. Pearson's correlation coefficient r was computed for SMQ, FMI, and CHIME scores to assess convergent validity. Bartlett's sphericity test was used to determine the applicability of a factorial model on the sample. The Kaiser–Meyer–Olkin (KMO) statistic was examined to compare the correlations between variables. KMO values close to one display appropriateness while in cases for which the KMO index is closer to zero, no factor analysis should be conducted. In a next step, a single-factor solution and a four-factor solution (Observing, Letting go, Non-judgment, and Non-aversion) as proposed by Chadwick et al. (2008) were tested by confirmatory factor analysis (CFA) with diagonally weighted least squares estimation

(DWLS). Assessed model fit indices included chi-square (χ^2), root mean square error of approximation (RMSEA), confirmatory factor index (CFI), and (standardized) root mean square residual (SRM; Kline 2015). In addition, measurement invariance analysis was conducted to assess whether the meaning of the items is comprehended differently between the clinical and the non-clinical groups (Sass 2011). Lastly, a chi-square difference test was performed comparing the one- and four-factor solution in order to compare the fit of the two models. The level of significance for all analyses was set at $\alpha = 0.05$.

Results

Tables 2 and 3 show the SMQ item means, item standard deviations, and one- as well as four-factor-model loadings for the non-clinical sample (638 community sample and 210 meditators). The mean SMQ score for the non-clinical sample ($n = 848$) is 49.23 ($SD = 15.20$). Cronbach's α was 0.91, indicating excellent internal consistency. For exploratory reasons, the classification criteria for the meditating sample were set to 8 weeks of meditation experience, which corresponds to classical MBSR and MBCT therapy programs. When compared with the 4-week experience criteria applied previously, the SMQ scores for the two non-clinical samples remain fairly similar: within the community sample, the mean SMQ score changes from 46.40 to 46.49 ($SD = 14.22$) and for the meditators, the mean shifts from 57.82 to 58.52 ($SD = 14.76$).

Kolmogorov–Smirnov test of normality ($Z = 0.04$; $p = 0.008$) indicated non-normality while visual inspections indicate that the distributions are fairly normal, which suggests that the assumptions for using parametric tests seem to be met (Öztuna et al. 2006). Bartlett's test of sphericity, testing the overall significance of the correlations within the correlation matrix, was significant (χ^2 (120) = 5681.343, $p < 0.001$), indicating that applying a factorial model on the sample is appropriate. The Kaiser–Meyer–Olkin statistic indicates high strength of correlation among variables (KMO = 0.935); thus, the sample is considered suitable for subsequent factor analysis. For the one-factor model, CFA indicates χ^2 (104, $n = 848$) = 1450.52, $p < 0.001$; RMSEA = 0.124, $p < 0.001$; CFI = 0.907; and SRMR = 0.064. The four-factor solution displays χ^2 (98, $n = 848$) = 1177.79, $p < 0.001$; RMSEA = 0.114, $p < 0.001$; SRMR = 0.058; and CFI = 0.926. These results indicate a better fit of the four-factor solution, which, however, still not displays good fit (Hooper et al. 2008). The chi-square difference test indicates the two models to be significantly different ($p < 0.001$).

Correlations of SMQ scores with FMI scores ($r = 0.71$, $p < 0.01$) and CHIME scores ($r = 0.78$, $p < 0.01$) indicate strong convergent validity for the non-clinical sample (see Table 4). T testing revealed that individuals with current meditation practice of 4 weeks (t(846) = − 9.978, $p < 0.001$) score

Table 2 SMQ item mean, item standard deviation, and corrected item-total correlations for all samples as well as R-Square and unrotated factor loadings as reported by exploratory factor analysis for the clinical sample

SMQ items	Total sample			Clinical sample					Non-clinical sample			Meditators			Community		
	Item mean	SD	C. ITC	Item mean	SD	C. ITC	R^2	UFL	Item mean	SD	C. ITC	Item mean	SD	C. ITC	Item mean	SD	C. ITC
1.	3.30	1.52	0.56	3.04	1.69	0.42	0.55	0.30	3.37	1.46	0.59	3.86	1.33	0.63	3.21	1.47	0.56
2.	1.60	1.23	0.48	1.62	1.44	0.43	0.49	0.24	1.60	1.18	0.51	1.98	1.26	0.49	1.47	1.12	0.49
3.	1.95	1.34	0.44	1.86	1.37	0.29	0.33	0.11	1.98	1.33	0.48	2.44	1.53	0.48	1.82	1.22	0.43
4.	2.91	1.47	0.60	2.56	1.64	0.50	0.65	0.42	3.00	1.41	0.63	3.59	1.31	0.62	2.80	1.38	0.59
5.	3.83	1.41	0.64	3.17	1.55	0.48	0.63	0.40	3.99	1.33	0.67	4.47	1.20	0.71	3.83	1.33	0.63
6.	3.80	1.62	0.39	3.42	1.77	0.12	0.14	0.02	3.90	1.56	0.44	4.18	1.47	0.50	3.80	1.58	0.41
7.	2.94	1.50	0.46	2.76	1.60	0.31	0.47	0.22	2.99	1.48	0.49	3.61	1.39	0.58	2.78	1.45	0.40
8.	3.29	1.72	0.50	2.78	1.66	0.29	0.32	0.11	3.42	1.71	0.53	3.67	1.68	0.54	3.33	1.71	0.53
9.	3.25	1.53	0.64	2.70	1.70	0.42	0.57	0.32	3.39	1.46	0.69	3.88	1.36	0.72	3.23	1.45	0.66
10.	3.11	1.51	0.69	2.91	1.69	0.61	0.75	0.57	3.16	1.46	0.72	3.84	1.32	0.72	2.93	1.43	0.69
11.	3.69	1.57	0.63	3.13	1.71	0.50	0.65	0.42	3.84	1.50	0.65	4.22	1.37	0.68	3.71	1.52	0.64
12.	2.64	1.58	0.45	2.29	1.64	0.20	0.21	0.04	2.73	1.56	0.50	3.52	1.56	0.56	2.47	1.47	0.41
13.	2.26	1.39	0.57	2.17	1.54	0.47	0.51	0.26	2.29	1.35	0.6	2.89	1.50	0.54	2.09	1.24	0.58
14.	3.12	1.69	0.65	2.41	1.68	0.51	0.55	0.30	3.31	1.64	0.67	4.03	1.56	0.71	3.06	1.60	0.62
15.	3.12	1.51	0.58	2.76	1.61	0.34	0.49	0.24	3.21	1.47	0.63	4.07	1.28	0.67	2.92	1.42	0.57
16.	3.03	1.60	0.59	2.77	1.69	0.53	0.58	0.34	3.10	1.57	0.61	3.57	1.55	0.62	2.94	1.54	0.58

SD = standard deviation, $C.ITC$ = corrected item-total correlation, UFL = unrotated factor loadings

Table 3 Factor loadings and R-square as reported by confirmatory factor analysis for the total sample and the non-clinical sample including the factors Observing, Letting go, Non-judgment, and Non-aversion

SMQ items	Total sample							Non-clinical sample						
	Four-factor model					One-factor model		Four-factor model					One-factor model	
	Ob	LG	NJ	NA	R^2	FL	R^2	Ob	LG	NJ	NA	R^2	FL	R^2
1.	0.63				0.40	0.62	0.39	0.67				0.45	0.66	0.44
7.	0.54				0.29	0.53	0.28	0.56				0.32	0.55	0.30
9.	0.75				0.56	0.74	0.55	0.79				0.62	0.78	0.61
16.	0.65				0.42	0.64	0.40	0.67				0.45	0.66	0.44
4.		0.68			0.47	0.67	0.45		0.70			0.50	0.70	0.49
10.		0.80			0.64	0.79	0.63		0.84			0.70	0.83	0.70
2.		0.58			0.33	0.57	0.33		0.62			0.38	0.61	0.37
13.		0.64			0.41	0.64	0.40		0.68			0.46	0.68	0.46
5.			0.79		0.63	0.72	0.52			0.83		0.69	0.75	0.56
6.			0.45		0.21	0.42	0.17			0.50		0.25	0.45	0.20
12.			0.58		0.33	0.53	0.28			0.62		0.38	0.56	0.31
14.			0.78		0.60	0.71	0.50			0.81		0.66	0.73	0.54
11.				0.72	0.52	0.70	0.49				0.77	0.59	0.72	0.52
15.				0.67	0.44	0.64	0.41				0.73	0.53	0.68	0.46
3.				0.50	0.25	0.48	0.23				0.55	0.30	0.52	0.27
8.				0.57	0.32	0.55	0.30				0.61	0.38	0.58	0.34

Ob = domain "Observing"; *LG* = domain "Letting go"; *NJ* = domain "Non-judgment"; *NA* = domain "Non-aversion"; *FL* = factor loading

significantly higher on the SMQ ($M = 57.82$, $SD = 15.05$) compared to non-meditators ($M = 46.40$, $SD = 14.16$). Furthermore, for exploratory reasons, correlations between SMQ scores and the number of hours someone meditates per week ($r = 0.23$, $p < 0.001$), as well as the number of hours a person practices mindfulness ($r = 0.29$, $p < 0.001$), were assessed within the meditation sample.

Discussion

Study 1 provided further evidence for the utility of the SMQ in a non-clinical sample. It shows excellent internal consistency and reliability as well as strong convergent validity with two frequently used and well-validated mindfulness measures in clinical practice and research. Moreover, a positive relationship was assessed between SMQ scores and frequency as well

as duration of meditation practice. In addition, data indicated that the SMQ discriminates well between meditators and non-meditators, suggesting its generalizability to a wide non-clinical population. All outcomes are consistent with previous trials, but similarly could confirm neither a single-factor nor a four-factor structure. Therefore, further research is needed to examine the construct of mindfulness (Baer et al. 2006; Chadwick et al. 2008).

Study 2

The overall aim of study 2 was, analogous to study 1, to measure reliability, factor structure, and convergent validity of the SMQ, however, in a clinical sample including patients diagnosed with schizophrenia and depression. In a previous

Table 4 SMQ convergent validity. Correlation of SMQ scores with FMI and CHIME for total sample and subsamples

Questionnaire	Total sample	Clinical sample	Non-clinical sample		
				Community	Meditators
FMI	0.659**	0.383**	0.707**	0.643**	0.760**
CHIME	0.730**	0.472**	0.775**	0.726**	0.803**

**Significant at 0.01 (two-tailed)

validation trial, the SMQ showed good internal consistency and adequate concurrent validity, and discriminated well between a non-clinical sample and individuals with psychosis (Chadwick et al. 2008). Moreover, it was shown that mindfulness abilities were inversely related to intensity of delusional experiences and negative affect (Chadwick et al. 2008). Therefore, study 2 examined these psychometric properties in a larger clinical sample including in- and outpatient with schizophrenia, as well as patients with depression, thereby adding another clinically relevant group.

Method

Participants

In the clinical samples, 54 participants with schizophrenia were inpatients, and 52 were outpatients, all in treatment at a university hospital in Berlin (study site 1). Of the depressive sample, 77 were patients at the university hospital inpatient ward (study site 1). Moreover, of all participants recruited (study site 2) for study 3, 30 with a depression diagnosis were also included into the present clinical sample, resulting in a sample of 107 participants with depression (see Fig. 1). The total clinical sample therefore comprised 213 participants (126 women, 86 men, 1 non-binary), with a mean age of 43.18 ($SD = 14.57$; range 18–80). For the 106 participants with schizophrenia (65 women, 40 men, 1 diverse gender), the mean age was 42.12 ($SD = 13.19$; range 20–71), and for the 107 depressive participants (61 women, 46 men), the mean age was 44.23 ($SD = 15.82$; range 18–80). In- and outpatients did not differ in regard to any demographic variables (age, duration of illness, gender). Further analyses were conducted concerning gender, occupation, and family status. There were,

however, no statistically significant differences between any of the subsamples on the sociodemographic variables assessed. Overall, the majority of the clinical sample was German ($n = 162$), single ($n = 130$) or married ($n = 28$), and already retired ($n = 63$), currently employed ($n = 44$), or seeking work ($n = 42$). A number of $n = 33$ achieved a university degree, $n = 42$ completed an apprenticeship, $n = 42$ completed high school, and $n = 48$ achieved an intermediate school-leaving certificate.

Procedure

For the clinical sample, inclusion criteria were defined as speaking German, being able to understand and fill out the questionnaires, providing written informed consent prior participation, and a diagnosis of either F.2 spectrum (schizophrenia) or F.3 spectrum (depressive disorders) (American Psychiatric Association 2000, 2013). All diagnoses were determined according to ICD-10 or DSM-5 criteria by either a licensed psychiatrist, psychologist, and/or consulting psychiatrist through a semi-structured interview (DIPS) which is conducted routinely with every patient at the in- and outpatient facility of the university hospital.

Measures

The same questionnaires as in study 1 were assessed, including the SMQ, FMI, and CHIME.

Data Analyses

The statistical analysis of study 2 is consistent with the analysis described in study 1, besides the difference concerning
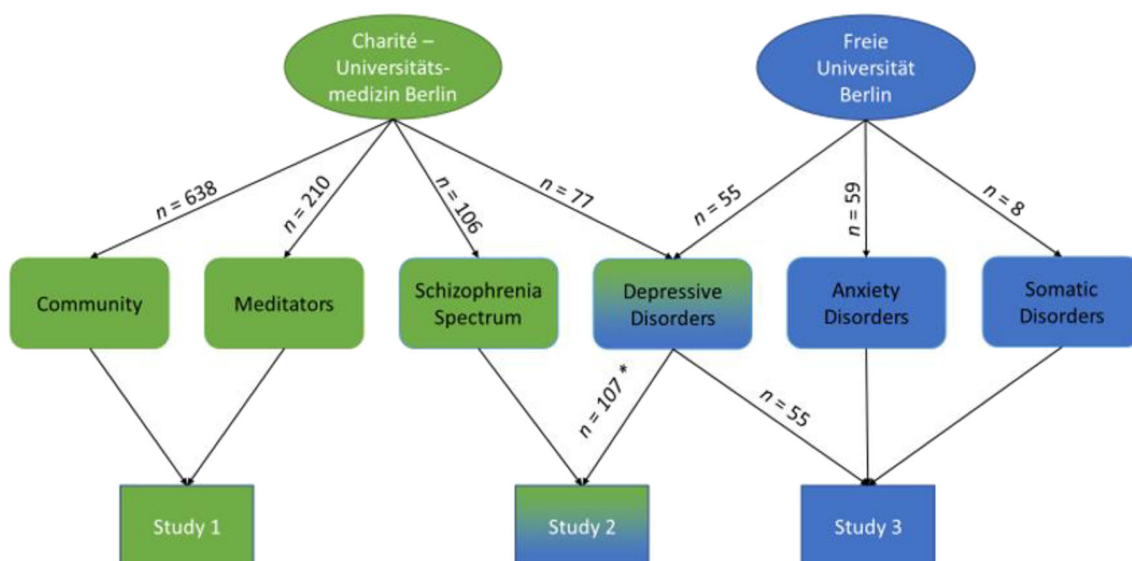

Fig. 1 Flow chart of the recruitment process. *$n = 77$ participants have been included from the Charité – Universitätsmedizin Berlin and $n = 30$ from the Freie Universität Berlin

the investigated sample population. However, no confirmatory factor analysis (CFA) was conducted, as the sample size of the clinical population in this study was too small for examining a four-factor structure (Wolf et al. 2013). Therefore, construct validity was evaluated by exploratory component factor analysis (EFA) with varimax rotation to explore the factor structure of the SMQ.

Results

Table 3 illustrates the SMQ item means, item standard deviations, corrected item-total correlations, R-Square, and rotated factor loadings for the clinical sample (106 schizophrenia and 107 depression). The mean SMQ score for the clinical sample ($n = 213$) was 42.35 ($SD = 13.03$). For the schizophrenia spectrum subsample, the mean SMQ score was 45.58 ($SD = 11.21$), whereas for the depressive disorder subsample ($n = 107$), the mean score was 39.15 ($SD = 13.93$). Furthermore, among the clinical outpatients ($n = 82$), SMQ scores had a mean of 42.71 ($SD = 12.51$), compared to inpatients ($n = 131$), who scored with a mean of 42.13 ($SD = 13.38$). The SMQ showed a Cronbach's α of 0.80, stating good internal consistency with corrected item-total correlations having a mean of 0.40 (range 0.12–0.61) for the clinical sample.

For the clinical sample, the Kolmogorov–Smirnov test of normality was insignificant ($Z = 0.04$; $p = 0.20$) and visual inspections indicated that the distributions are fairly normal. Bartlett's test of sphericity was significant ($\chi^2 (120) = 1006.557$, $p < 0.001$). The Kaiser–Meyer–Olkin statistic indicated a high correlation among variables (KMO = 0.821). EFA revealed that four factors are relevant to explain the structure of the SMQ. Factors 1, 2, 3, and 4 had eigenvalues of 4.31, 2.67, 1.29, and 1.00, respectively, and accounted for 26.93, 16.67, 8.05, and 6.29% of variance. Although four factors have eigenvalues above 1 and 2 account for 43% of the variance, an examination of the scree plot strongly suggests a two-factor structure. When attempting to interpret the two-factor solution, however, no interpretable two-factor structure could be formulated. It is a common occurrence that when dealing with clinical samples, uninterpretable factor structures are the result of factor analyses (Floyd and Widaman 1995).

Correlations between SMQ scores and FMI scores ($r = 0.38$, $p < 0.01$) as well as CHIME scores ($r = 0.47$, $p < 0.01$) were significant, indicating moderate convergent validity (see Table 4).

Discussion

The current findings are in line with the original validation study, demonstrating the SMQ's reliability and convergent validity with other mindfulness measures in clinical populations (Chadwick et al. 2008). Furthermore, these are the first

empirical data to show these outcomes in out- and inpatients with schizophrenia as well as depression, supporting clinical and research utility of the SMQ for a variety of clinical disorders. Aligned with Chadwick et al. (2008), the SMQ seems to display a one-factor structure; however, further research needs to establish the measures with other clinical groups and examine its underlying construct.

Study 3

The original validation study of the SMQ investigated the measures divergent validity in regard to positive and negative affect as well as psychotic symptoms in a sample of psychotic patients (Chadwick et al. 2008). Since mindfulness is discussed as an important process in emotional disorders and frequently targeted in psychotherapy, the aim of study 3 was to further broaden the understanding of the SMQ's relation to bordering constructs and sensitivity to change following psychological treatment. To achieve this, the SMQ's relation to symptom distress, positive and negative affect, experiential avoidance, and anxiety sensitivity was investigated using outcome data from an Internet-based treatment for emotional disorders. This investigation was part of an ongoing randomized waitlist-controlled trial on an Internet-based transdiagnostic treatment for emotional disorders (conducted at study site 2) (registered as DRKS00014820 in the German Clinical Trials Register DRKS). Participants were randomized to receive a 10-week guided transdiagnostic intervention based on the Unified Protocol (Barlow et al. 2018) or to a waitlist control. The Unified Protocol postulates mindful emotional awareness as one of the underlying mechanisms for symptom reduction. Accordingly, the third week of the Internet-based program focuses on mindfulness and encourages participants to practice mindfulness with a selection of mindfulness exercises. Other modules include motivation and goal setting (week 1), psychoeducation on emotions (week 2), cognitive restructuring (weeks 4 and 5), interoceptive and in vivo exposure (weeks 6–9), and relapse prevention (week 10). Participants work through the ten modules independently and receive weekly feedback from trained and supervised online therapists.

Participants

Of the 122 participants that took part in the study, 55 participants were diagnosed with a primary diagnosis of depression, 59 participants with a primary anxiety disorder, and 8 participants with a primary somatic symptom disorder (see Fig. 1). Participants (83 women, 39 men) were, on average, 37.32 years old ($SD = 12.49$, range 18–67). If participants fulfilled the inclusion criteria, they were randomized to the treatment group

or the waitlist control group. To examine the SMQ's sensitivity to change, the change in mindfulness between participants in the treatment group and the waitlist group was compared. The SMQ was assessed at baseline prior to randomization, after the mindfulness module/week 3, and after the 10-weekintervention/ waiting period. Of the $n = 122$ participants, all filled out the baseline assessment; $n = 51$ of the treatment group and $n = 46$ of the waitlist condition filled out the module/week 3 assessment, and $n = 36$ of the treatment and $n = 52$ of the waitlist filled out the post-intervention/waiting period assessment. Overall, $n = 35$ of the treatment group and $n = 45$ of the waitlist condition filled out the SMQ at all three time points and were included in the complete analysis. Of the $n = 35$ participants in the treatment group, $n = 15$ had a principal anxiety, $n = 18$ a principal depressive, and $n = 2$ a principal somatic symptom disorder. Of the $n = 45$ participants in the waitlist group, $n = 23$ had a principal anxiety, $n = 18$ a principal depressive, and $n = 4$ a principal somatic symptom disorder. The trial is still ongoing and $n = 8$ are currently still in treatment and thus could not be included in the analysis.

Procedure

Participants of the randomized controlled trial were recruited through mental health forums as well as on social media platforms. Diagnoses were determined with a structured diagnostic interview (Margraf et al. 2017) via telephone.

Measures

In addition to the SMQ, the following measures were assessed the following measures in the baseline assessment.

Brief Symptom Inventory 18 (Derogatis 2001; Franke et al. 2011)

The Brief Symptom Inventory 18 (BSI-18) is an 18-item short version of the Symptom-Checklist-90-R. The BSI-18 assesses symptom distress with three 6-item subscales depression, anxiety, and somatization. Items are rated on a 5-point Likert-type scale (*not at all* (*0*) to *extremely* (*4*)), yielding a total score for the General Symptom Index from 0 to 72 (Derogatis 2001). Confirmatory factor analysis supported the 3-factor structure (Franke et al. 2011; Prinz et al. 2013). The BSI-18 and the subscales show high internal consistency and overall satisfactory psychometric properties (Derogatis and Fitzpatrick 2004; Franke et al. 2011; Prinz et al. 2013). The Cronbach's α of the BSI-18 in the current study was 0.82.

Positive and Negative Affect Schedule (Krohne et al. 1996; Watson et al. 1988)

The Positive and Negative Affect Schedule (PANAS) measures positive and negative affect with two quasi-independent 10-item scales. Participants rate to what extent they experienced an affect, e.g., "active" or "afraid," on a 5-point Likert-type scale (*very slightly or not at all* (*1*) to *extremely* (*5*)), yielding a total score for positive as well as negative affect from 10 to 50. Both scales show high internal consistency as well as convergent and divergent validity (Watson et al. 1988). In the current study, for the positive affectivity scale, Cronbach's α was 0.83 and for the negative affectivity scale 0.85.

Brief Experiential Avoidance Questionnaire (Gámez et al. 2014)

The Brief Experiential Avoidance Questionnaire (BEAQ, Gámez et al. 2014), a short version of the Multidimensional Experiential Avoidance Questionnaire (Gámez et al. 2011), assesses experiential avoidance as a broad construct with 15 items. Items are rated on a 6-point Likert-type scale (*strongly disagree* (*1*) to *strongly agree* (*6*)), yielding a total score from 15 to 90. The measure covers several aspects of experiential avoidance like explicit avoidance behavior and attitudes in regard to distress. The BEAQ is internally consistent and can be distinguished from negative emotionality (Gámez et al. 2014). For study 3, the BEAQ was translated to German and validated in a German student ($N = 596$) and clinical outpatient ($N = 53$) sample. Internal consistency in both samples was good ($\alpha = 0.81$ in the student and $\alpha = 0.87$ in the clinical sample) and expected correlations to convergent and divergent measures supported its validity (please contact the second author for detailed information on the German translation of the BEAQ). In the current study, Cronbach's α for the BEAQ was 0.74.

Anxiety Sensitivity Index 3 (Kemper et al. 2009; Taylor et al. 2007)

The Anxiety Sensitivity Index 3 (ASI-3, Kemper et al. 2009; Taylor et al. 2007) assesses anxiety sensitivity—the fear of arousal-related bodily symptoms—with 18 items. Items are rated on a 4-point Likert-type scale (*very little* (*0*) to *very much* (*4*)), yielding a total score from 0 to 72. The ASI-3 has a three-factor structure with the subscales Physical, Cognitive, and Social Concern, and satisfactory psychometric qualities (Taylor et al. 2007). In the current study, the ASI-3 had an internal consistency of Cronbach's $\alpha = 0.88$.

Data Analyses

Statistical analyses were performed using IBM SPSS Statistics 25. To analyze divergent validity, correlations were calculated

between the SMQ and the BSI's total score and subscales (depression, anxiety, and somatization), the positive and negative affect scale of the PANAS, the BEAQ, and ASI-3. To explore the SMQ's sensitivity to change, we calculated a mixed ANOVA with the three assessment points of the SMQ and group as a between factor. To further examine change within the two groups, separate repeated measurement ANOVAs were calculated. The level of significance for all analyses was set at $\alpha < 0.05$.

Results

For the sample of study 3, Cronbach's α of the SMQ was 0.88. The SMQ's correlations to hypothesized divergent measures are displayed in Table 5. The SMQ correlated significantly with all measures except positive affect, ranging from -0.21 for experiential avoidance to -0.50 for negative affect.

Concerning sensitivity to change, means, standard deviations, and effect sizes for the SMQ for both groups are displayed in Table 6. The mixed ANOVA showed a statistically significant interaction between time and group, Greenhouse–Geisser $F(1.86, 144.89) = 8.34$, $p < 0.05$, partial $\eta^2 = 0.097$. Mindfulness changed more in the group receiving the online transdiagnostic treatment than in the waitlist group. In the separate analysis of the treatment and waitlist group, a significant effect of time on SMQ scores was found in the treatment group (Greenhouse–Geisser $F(1.73, 58.7) = 13.31$, $p < 0.01$, partial $\eta^2 = 0.28$), but not in the waitlist group (Greenhouse–Geisser $F(1.93, 84.99) = 2.44$, $p = 0.095$, partial $\eta^2 = 0.053$). The pre- to post-gain in SMQ scores of 10.86 points in the treatment group corresponds to an effect size of 0.71 (Cohen's d) suggesting a medium to large effect. These results indicate that the SMQ questionnaire is sensitive to change, depicting changes in the treatment group where mindfulness was actively trained and depicting no change in the passive waitlist control condition.

Discussion

Study 3 found medium to large significant correlations between the SMQ and symptom distress, negative affect, and anxiety sensitivity and a small correlation to experiential avoidance. The SMQ's correlation to positive affect was non-significant. While this is in line with other studies that found significant correlations between the SMQ and negative affect (Chadwick et al. 2008) and psychopathology (Baer et al. 2006), the medium to large correlations query future studies to examine divergent validity of the SMQ. This was the first study to explore the SMQ's sensitivity to change and found that the SMQ captured changes in mindfulness following treatment. While further research with larger sample sizes is needed, these results provide preliminary evidence for the SMQ's utility in psychotherapy process research.

Analysis of Overall Sample

In the final stage of the study, in order to examine the overall psychometric properties of the SMQ, data from the non-clinical and clinical sample were combined. As the primary analyses for the current manuscript were performed at an early stage, only data from studies 1 and 2 were included into the following analyses. Therefore, in total, 1061 participants were included. Tables 2 and 3 display the SMQ item means, item standard deviations, and one- as well as four-factor-model loadings for the total sample. Overall, the SMQ mean score for the total sample ($N = 1061$) is 47.85 ($SD = 15.04$). Cronbach's α was 0.89. Kolmogorov–Smirnov test of normality was applied ($Z = 0.04$; $p = 0.007$) indicating non-normality. However, additional visual inspection revealed a normal distribution. Bartlett's test of sphericity was significant ($\chi^2 (120) = 6415.450$, $p < 0.001$) and KMO statistic revealed high strength of correlation among variables with 0.925, indicating that the sample is suited for factor analysis. In a next step, CFA was performed indicating a higher $\chi^2 (104, n = 1061) = 1943.05$, $p < 0.001$, a higher RMSEA = 0.129, $p < 0.001$, a lower CFI = 0.885, and a higher SRMR = 0.075 for the one-factor model compared to the four-factor solution $\chi^2 (98, n = 1061) = 1729.1$, $p < 0.001$; RMSEA = 0.125, $p < 0.001$; SRMR = 0.072; and CFI = 0.898. These results suggest a better fit of the four-factor solution, which is further supported by the chi-square difference test ($p < 0.001$). However, as indicated by the $\chi^2 p$ value above 0.05, the CFI below 0.9, and the RMSEA above 0.08, the four-factor model does not display good fit, either (Hooper et al. 2008). Furthermore, measurement invariance analysis between the

Table 5 SMQ divergent validity. Correlation of SMQ scores with BSI-18 and its subdomains, the two PANAS domains, BEAQ and ASI-3

BSI-18				PANAS		BEAQ	ASI-3
Total	Depression	Anxiety	Somatization	Positive affect	Negative affect		
-0.48^{**}	-0.30^{**}	-0.48^{**}	-0.26^{**}	-0.09	-0.50^{**}	-0.21^{*}	-0.38^{**}

**Significant at 0.01 (two-tailed); *significant at 0.05 (two-tailed)

Table 6   SMQ sensitivity to change: comparison between the treatment and waitlist group

	Treatment group		Waitlist group		
	Mean	Std. deviation	Mean	Std. deviation	Cohen's d
Pre-assessment	36.0	15.87	34.82	10.96	
Module 3/week 3 assessment	35.89	12.36	32.31	11.23	0.30
10-week post-intervention/waiting period assessment	46.86	14.49	35.20	12.54	0.86

Means and standard deviations for the SMQ for the treatment and waitlist group for all three assessment points as well as between-group effect sizes (Cohen's d) for the module/week 3 and 10-week post-intervention/waiting period assessment point

clinical and the non-clinical sample showed that model fit indices differed significantly between the configural model (model 1) and a model in which factor loadings were constrained (model 2). This holds for the single-factor structure (model 1: χ^2 (208) = 1951.3; model 2: χ^2 (223) = 2160.9, $p < 0.001$) as well as the four-factor structure (model 1: χ^2 (196) = 1675.7; model 2: χ^2 (208) = 1801.9, $p < 0.01$).

In the last step, for exploratory reasons, the total SMQ scores were compared between the community sample ($n = 638$, $M = 46.40$, $SD = 14.16$), meditators ($n = 210$, $M = 57.82$, $SD = 15.05$), and the clinical sample ($n = 213$, $M = 42.35$, $SD = 13.03$). One-way ANOVA indicated significant differences in SMQ scores between subgroups ($F(2, 1058) = 71.81$, $p < 0.001$). Scheffe post hoc test showed that the mean in total SMQ scores of each subgroup was different from the other groups at 0.05 probability. Correlations between the SMQ and FMI scores ($r = 0.66$, $p < 0.01$) as well as CHIME scores ($r = 0.73$, $p < 0.01$) were significant, indicating strong convergent validity for the total sample (see Table 4).

General Discussion

As one of the few self-rating mindfulness questionnaires, the SMQ uniquely assesses individuals' mindfulness awareness of distressing thoughts, images, and perceptions which constitute major psychopathological phenomena in a variety of clinical disorders. The overall study involved three substudies to examine internal consistency and reliability, convergent validity and factor structure in meditators, non-meditators, and patients with schizophrenia and with depression as well as divergent validity and treatment sensitivity in patients with emotional disorders.

Overall, the SMQ displayed excellent internal consistency and moderate convergent validity when compared with the FMI and CHIME, two frequently used mindfulness questionnaires in clinical practice and research with a different factor structure. An explanation for these results might be that the CHIME and FMI do not provide an opportunity to respond neutrally. Consequently, a bias towards polarized response patterns might be created for clinical groups which frequently

experience cognitive impairment and therefore might find a neutral response option helpful. The inclusion of a neutral option for responding as in the SMQ decreases the occurrence of extreme response styles while instead fosters a wider spread of responses. This can be considered a systematic and important factor for clinical populations. Therefore, these results seem to support the utility of the SMQ for clinical groups, as current results indicate that the SMQ can discriminate between individuals diagnosed with schizophrenia or depression, and meditators as well as a community sample. In line with the original validation study, present outcomes support the reproducibility of the SMQ as the same relative pattern of differences between groups, as well as similar absolute values in each group were found (Chadwick et al. 2008). The similarity in SMQ scores between samples of non-meditating English speakers drawn from the UK and German speakers drawn from different Western European countries raises broader questions about the cross-cultural relevance of the scales, including similarities and differences in dispositional mindful awareness of distressing thoughts and images.

Furthermore, when the classification criteria for the meditating sample are changed to 8 weeks of prior meditation experience, which is comparable to 8-week meditation programs such as MBSR and MBCT, compared to 4 weeks of experience as initially applied in this study, the SMQ scores for the two non-clinical samples remain fairly similar. These outcomes suggest high generalizability of the results and utility for non-clinical populations such as extensively practicing meditators, as changes in SMQ mean scores were minor when narrowing the inclusion criteria for the meditating sample. Among meditators, the expectations concerning the effects of their practice may be a source of bias for the self-report assessment of mindfulness, as they might overestimate their mindfulness skills (Grossman 2008). Furthermore, as the current study did not assess the exact type of meditation the participants practice, future research should take a differentiating look on the diverse kinds of mediation practices and their influence on mindfulness. Nevertheless, in contrast to most studies assessing mindfulness, the current study provided information on the relationship between mindfulness scores and current meditation practice. Other studies on this subject

mostly conceptualize meditation experience as the number of years since the individual firstly encountered meditation, whereas current practice is often neglected as a factor (Bergomi et al. 2015). In the study, post hoc testing revealed that experienced meditators with current practice score significantly higher on the SMQ, compared to those who practiced mediation in the past, but not currently. This is in line with studies comparing mindfulness scores of meditators and non-meditators (Bergomi et al. 2015; Carmody and Baer 2008).

In the original study of Chadwick et al. (2008), an inconclusive factor structure of the SMQ was reported yet emphasizing a one-factor structure. The outcomes of the current study suggest a four-factor structure for the non-clinical population to be more appropriate. Moreover, for the total sample, the four-factor solution displayed a better model fit compared to a single-factor structure as indicated by the chi-square difference test on the one hand, but on the other hand, the model fit indices still did not indicate a good fit. Most prominently, the high RMSEA index emphasizes the lack of a good fitting model to conceptualize mindfulness and capture the underlying factor structure of the SMQ, suggesting a multifaceted understanding and conceptualization of mindfulness (Baer et al. 2006). Furthermore, as indicated by the measurement invariance testing, the SMQ items seem to be comprehended differently by the clinical and non-clinical sample. The outcomes of the current study therefore favor a four-factor over a one-factor structure of the SMQ. However, these statements have to be viewed with caution and further research with a larger clinical sample is needed to examine the construct of mindfulness. There is no consensus on the factorization of mindfulness in questionnaires so far, as some researchers argue that there are distinguishable components that are stable across populations, such as applied in the CHIME (Baer et al. 2004). Other studies conceptualized a two-factor or a multi-factor solution (for example, the CHIME, Bergomi et al. 2013).

To assess divergent validity, the relationship between the SMQ and measures of psychopathology (PANAS and BSI-18), anxiety sensitivity (ASI-3), and experiential avoidance (BEAQ) was explored. Similar to mindfulness, these constructs are discussed as mechanisms involved in the onset and maintenance of mental disorders which raises the question of their relation among each other. The current study found medium to large correlations between the SMQ and BSI-18 total score as well as its subscales anxiety and somatization, negative affect (PANAS), and anxiety sensitivity (ASI-3) and a small correlation with experiential avoidance (BEAQ). A non-significant correlation was observed between the SMQ and positive affect (PANAS). While the medium to large correlations question the SMQ's divergent validity, these findings are in line with the initial validation study (Chadwick et al. 2008) as well as previous studies that reported correlations between mindfulness and psychopathology, and anxiety sensitivity as well as experiential avoidance (Mahoney et al. 2015; Baer et al. 2006). With these constructs showing similar

conceptualizations, small correlations can be expected to some extent and further research needs to explore their distinct relation and incremental validity. Contrary to the hypothesis that the SMQ should correlate higher with measures of mindfulness than with measures of different constructs, it has been observed that the SMQ correlated lower with the FMI ($r = 0.38$) and CHIME ($r = 0.47$) in the clinical sample than with symptom distress ($r = -0.48$), negative affect ($r = -0.50$), and anxiety sensitivity ($r = -0.45$). A similar pattern was observed by Baer et al. (2006), as higher correlations were found between experiential avoidance and the SMQ as between the SMQ and FMI. In addition, in line with the current study results, Bear et al. (2006) report a similar strength in correlation between the SMQ and FMI as well as SMQ and symptom distress, albeit lower than that in the current study. One could argue that the specific aspect of mindfulness that is captured by the SMQ, the non-judging perception of negative inner states, is very closely related to symptoms of emotional disorders, which were prevalent in the current study. Emotional disorders are characterized by maladaptive reactions to frequently experienced negative emotions (Barlow et al. 2016). It could be that the SMQ, similar to the construct of experiential avoidance, taps into that aspect of emotional disorders and therefore is closer related to distressing thoughts and images captured by psychopathology measures compared to broader measures of mindfulness in clinical samples.

The study examined treatment sensitivity of the SMQ within a randomized controlled trial of an Internet-based psychotherapy treatment program with a significant mindfulness component delivered online for people with depressive, anxiety, or somatic symptom disorders. Data showed a significant improvement in mindfulness of distressing thoughts and images in favor of the treatment group. These results underpin the transdiagnostic nature of mindfulness and the SMQ's ability to capture changes in mindfulness following the treatment employed in the current study. In order to examine the question of whether improved mindfulness of distressing thoughts and images is an underlying mechanism of change in psychotherapies in general (Bergomi et al. 2013), future research needs to examine the change in the SMQ score following psychotherapy that does not directly teach mindfulness. While this delivers preliminary evidence for the SMQ's sensitivity to change, these results need to be interpreted cautiously in the light of the high percentage of missing values. To draw conclusions of the SMQ's clinical and research utility, further research is needed to assess the SMQ in a larger sample.

Limitations and Future Research Directions

The current study should be seen in the light of several limitations. First, there were high dropout rates for the sensitivity to change analysis within the randomized controlled trial. Future research should aim for a larger sample. Therewith, the current results should be replicated within different

therapeutic interventions. Second, a majority of patients' symptom severity was not measured through clinical assessments. Here, it might be useful to examine the relationship between symptom severity, disorder spectrum, and mindfulness abilities. Third, factor analysis outcomes with ordinal data, such as a Likert scale, should be treated with caution as they can lead to "over-factoring," indicating multiple underlying dimensions, even though the true structure is unidimensional. Future research therefore could employ parallel analysis for assessing dimensionality and polychoric correlations instead of Pearson correlations (Van der Eijk and Rose 2015). Lastly, future research needs to investigate test-retest reliability as well as divergent validity in extended sample sizes and also of clinical groups other than schizophrenia and depression.

In line with the initial development and validation (Chadwick et al. 2008), the results of the current study support the SMQ's value and practicality in clinical practice and research in German-speaking areas. The SMQ has particular value in process and outcome assessment in mindfulness for emotional disorders, as well as being applicable in other clinical applications.

Author Contributions KB designed and executed the study, assisted with the data analyses, and wrote the paper. CS collaborated with the design and executed the study, assisted with the data analyses, and wrote parts of the paper. PJ collaborated with the design and editing of the final manuscript. PC collaborated with the design and editing of the final manuscript. EE executed the study and wrote parts of the methods. IH executed the study, assisted with the data analyses, and edited the manuscript. NB assisted in data analyses and wrote parts of the results and the discussion. JB collaborated with the design and edited the final manuscript. KW edited the final manuscript. MB edited the final manuscript. EH designed the study and edited the final manuscript.

Compliance with Ethical Standards

Informed Consent All persons gave their informed consent before their inclusion in the study.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the ethical committee of each study site, the Helsinki declaration, and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Conflict of Interest The authors declare that they have no conflict of interest.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC.

Baer, R. A. (2007). Mindfulness, assessment, and transdiagnostic processes. *Psychological Inquiry, 18*(4), 238–242. https://doi.org/10.1080/10478400701598306.

Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-report. *Assessment, 11*(3), 191–206. https://doi.org/10.1177/1073191104268029.

Baer, R. A., 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.

Barlow, D. H., Allen, L. B., & Choate, M. L. (2016). Toward a unified treatment for emotional disorders–republished article. *Behavior Therapy, 47*(6), 838–853.

Barlow, D. H., Farchione, T. J., Sauer-Zavala, S., Latin, H. M., Ellard, K. K., Bullis, J. R., et al. (2018). *Unified protocol for transdiagnostic treatment of emotional disorders: therapist guide* (2nd ed.). New York: Oxford University Press.

Belzer, F., Schmidt, S., & Lucius-Hoene, G. (2011). *Freiburger Fragebogen zur Achtsamkeit im kognitiven Interview. Analyse der Beantwortungsprozesse in Abhängigkeit von Achtsamkeitserfahrung.* Bern, Switzerland: 3rd Conference of the Mindfulness Research Network.

Belzer, F., Schmidt, S., Lucius-Hoene, G., Schneider, J. F., Orellana-Rios, C. L., & Sauer, S. (2013). Challenging the construct validity of mindfulness assessment-a cognitive interview study of the Freiburg Mindfulness Inventory. *Mindfulness, 4*(1), 33–44. https://doi.org/10.1007/s12671-012-0165-7.

Bergomi, C., Tschacher, W., & Kupper, Z. (2013). Measuring mindfulness: first steps towards the development of a comprehensive mindfulness scale. *Mindfulness, 4*(1), 18–32. https://doi.org/10.1007/s12671-012-0102-9.

Bergomi, C., Tschacher, W., & Kupper, Z. (2014). Konstruktion und erste Validierung eines Fragebogens zur umfassenden Erfassung von Achtsamkeit. *Diagnostica*.

Bergomi, C., Tschacher, W., & Kupper, Z. (2015). Meditation practice and self-reported mindfulness: a cross-sectional investigation of meditators and non-meditators using the Comprehensive Inventory of Mindfulness Experiences (CHIME). *Mindfulness, 6*(6), 1411–1421. https://doi.org/10.1007/s12671-015-0415-6.

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., et al. (2004). Mindfulness: a proposed operational definition. *Clinical Psychology: Science and Practice, 11*(3), 230–241. https://doi.org/10.1093/clipsy/bph077.

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), 822-848. https://doi.org/10.1037/0022-3514.84.4.822.

Brown, K. W., & Ryan, R. M. (2004). Perils and promise in defining and measuring mindfulness: observations from experience. *Clinical*

Psychology: Science and Practice, 11(3), 242–248. https://doi.org/10.1093/clipsy/bph078.

Buchheld, N., Grossman, P., & Walach, H. (2001). Measuring mindfulness in insight meditation (Vipassana) and meditation-based psychotherapy : the development of the Freiburg Mindfulness Inventory (FMI). *Journal of Meditation and Meditation Research, 1*(1), 11–34. https://doi.org/10.2307/1411318.

Carmody, J., & Baer, R. A. (2008). Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. *Journal of Behavioral Medicine, 31*(1), 23–33. https://doi.org/10.1007/s10865-007-9130-7.

Chadwick, P., Taylor, K. N., & Abba, N. (2005). Mindfulness groups for people with psychosis. *Behavioural and Cognitive Psychotherapy, 33*(3), 351–359. https://doi.org/10.1017/S1352465805002158.

Chadwick, P., Hember, M., Symes, J., Peters, E., Kuipers, E., & Dagnan, D. (2008). Responding mindfully to unpleasant thoughts and images: reliability and validity of the Southampton mindfulness questionnaire (SMQ). *British Journal of Clinical Psychology, 47*(4), 451–455.

Christopher, M. S., Woodrich, L. E., & Tiernan, K. A. (2014). Using cognitive interviews to assess the cultural validity of state and trait measures of mindfulness among Zen Buddhists. *Mindfulness, 5*(2), 145–160. https://doi.org/10.1007/s12671-012-0160-z.

Derogatis, L. R. (2001). *BSI 18, Brief Symptom Inventory 18: administration, scoring and procedures manual.* NCS Pearson.

Derogatis, L. R., & Fitzpatrick, M. (2004). The SCL-90-R, the Brief Symptom Inventory (BSI), and the BSI-18. In M. E. Maruish (Ed.), *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Volume 3: Instruments for Adults.* Lawrence Erlbaum Associates Publishers.

DGPPN (2018). *Psychische Erkrankungen in Deutschland: Schwerpunkt Versorgung.* Retrieved 13 February 2020, from https://www.dgppn.de/_Resources/Persistent/f80fb3f112b4eda48f6c5f3c68d23632a03ba599/DGPPN_Dossierweb.pdf.

Dimidjian, S., & Linehan, M. M. (2003). Defining an agenda for future research on the clinical application of mindfulness practice. *Clinical Psychology: Science and Practice, 10*(2), 166–171. https://doi.org/10.1093/clipsy/bpg019.

Feldman, G., Hayes, A., Kumar, S., Greeson, J., & Laurenceau, J.-P. (2007). Mindfulness and emotion regulation: the development and initial validation of the Cognitive and Affective Mindfulness Scale-Revised (CAMS-R). *Journal of Psychopathology and Behavioral Assessment, 29*(3), 177–190. https://doi.org/10.1007/s10862-006-9035-8.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*(3), 286-299. https://doi.org/10.1037/1040-3590.7.3.286.

Franke, G. H., Ankerhold, A., Haase, M., Jäger, S., Tögel, C., Ulrich, C., & Frommer, J. (2011). Der Einsatz des Brief Symptom Inventory 18 (BSI-18) bei Psychotherapiepatienten. = the usefulness of the Brief Symptom Inventory 18 (BSI-18) in psychotherapeutic patients. *PPmP: Psychotherapie Psychosomatik Medizinische Psychologie, 61*(2), 82–86.

Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., & Watson, D. (2011). Development of a measure of experiential avoidance: the Multidimensional Experiential Avoidance Questionnaire. *Psychological Assessment, 23*(3), 692-713. https://doi.org/10.1037/a0023242.

Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., Suzuki, N., & Watson, D. (2014). The brief experiential avoidance questionnaire: development and initial validation. *Psychological Assessment, 26*(1), 35-45. https://doi.org/10.1037/a0034473.

Grossman, P. (2008). On measuring mindfulness in psychosomatic and psychological research. *Journal of Psychosomatic Research, 64*(4), 405-408. https://doi.org/10.1016/j.jpsychores.2008.02.001.

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). *Psychological Assessment, 23,* 1034-1040. https://doi.org/10.1037/a0022713.

Heidenreich, T., Ströhle, G., & Michalak, J. (2006). Achtsamkeit: Konzeptuelle Aspekte und Ergebnisse zum Freiburger Achtsamkeitsfragebogen. *Verhaltenstherapie, 16*(1), 33–40.

Höfling, V., Moosbrugger, H., Schermelleh-Engel, K., & Heidenreich, T. (2011). Mindfulness or mindlessness? *European Journal of Psychological Assessment, 27*(1), 59-64. https://doi.org/10.1027/1015-5759/a000045.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53–60.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2020). *semTools: useful tools for structural equation modeling.* R package version 0.5–2.918. Retrieved from https://CRAN.R-project.org/package=semTools.

Kabat-Zinn, J. (1990). *Full catastrophe living : Using the wisdom of your body and mind to face stress, pain, and illness.* New York: Delacorte Press.

Kemper, C. J., Ziegler, M., & Taylor, S. (2009). Überprüfung der psychometrischen Qualität der deutschen Versiondes Angstsensitivitätsindex-3. *Diagnostica, 55*(4), 223–233. https://doi.org/10.1026/0012-1924.55.4.223.

Kline, R. B. (2015). *Principles and practice of structural equation modeling.* Guilford publications.

Kohls, N., Sauer, S., & Walach, H. (2009). Facets of mindfulness – results of an online study investigating the Freiburg mindfulness inventory. *Personality and Individual Differences, 46*(2), 224–230. https://doi.org/10.1016/J.PAID.2008.10.009.

Krohne, H. W., Egloff, B., Kohlmann, C. W., & Tausch, A. (1996). Untersuchungen mit einer Deutschen version der "Positive and Negative Affect Schedule" (PANAS). *Diagnostica, 42,* 139–156.

Mahoney, C. T., Segal, D. L., & Coolidge, F. L. (2015). Anxiety sensitivity, experiential avoidance, and mindfulness among younger and older adults: age differences in risk factors for anxiety symptoms. *International Journal of Aging & Human Development, 81*(4), 217–240. https://doi.org/10.1177/0091415015621309.

Margraf, J., Cwik, J. C., Pflug, V., & Schneider, S. (2017). Strukturierte klinische Interviews zur Erfassung psychischer Störungen über die Lebensspanne. *Zeitschrift für Klinische Psychologie und Psychotherapie, 46,* 176-186. https://doi.org/10.1026/1616-3443/a000430.

Öztuna, D., Elhan, A. H., & Tüccar, E. (2006). Investigation of four different normality tests in terms of type 1 error rate and power under different distributions. *Turkish Journal of Medical Sciences, 36*(3), 171–176.

Prinz, U., Nutzinger, D. O., Schulz, H., Petermann, F., Braukhaus, C., & Andreas, S. (2013). Comparative psychometric analyses of the SCL-90-R and its short versions in patients with affective disorders. *BMC Psychiatry, 13*(1), 104. https://doi.org/10.1186/1471-244X-13-104.

Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36 URL http://www.jstatsoft.org/v48/i02/.

Safran, J. D., & Segal, Z. V. (1990). *Interpersonal process in cognitive therapy.* Basic Books.

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 29*(4), 347–363.

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

Taylor, S., Zvolensky, M. J., Cox, B. J., Deacon, B., Heimberg, R. G., Ledley, D. R., et al. (2007). Robust dimensions of anxiety sensitivity: development and initial validation of the Anxiety Sensitivity Index-3. *Psychological Assessment, 19*(2), 176-188. https://doi.org/10.1037/1040-3590.19.2.176.

Van der Eijk, C., & Rose, J. (2015). Risky business: factor analysis of survey data–assessing the probability of incorrect dimensionalisation. *PLoS One, 10*(3), e0118900.

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

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.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*(6), 913–934.

World Health Organization. (2014). *Process of translation and adaptation of instruments.* Retrieved 12 February 2020, from https://www.who.int/substance_abuse/research_tools/translation/en/.