Psy-Flex: A contextually sensitive measure of psychological flexibility

Andrew T. Gloster [a,*], Victoria J. Block [a], Jens Klotsche [b], Jeanette Villanueva [a], Marcia T. B. Rinner [a], Charles Benoy [c], Marc Walter [c], Maria Karekla [d], Klaus Bader [c]

[a] *University of Basel, Department of Psychology, Division of Clinical Psychology & Intervention Science, Missionsstrasse 62-64, 4055, Basel, Switzerland*
[b] *German Rheumatism Research Center Berlin, Epidemiology Unit and Charité Universitaetsmedizin Berlin, Institute for Social Medicine, Epidemiology and Health Economics, Charitéplatz 1, 10117, Berlin, Germany*
[c] *University of Basel, Psychiatric Hospital, Centre for Psychosomatics and Psychotherapy, Wilhelm Klein-Strasse 27, 4002, Basel, Switzerland*
[d] *Department of Psychology, University of Cyprus, P.O. Box 20537, CY 1678, Nicosia, Cyprus*

ARTICLE INFO

ABSTRACT

Psychological flexibility is increasingly studied in various contexts including clinical studies. The construct has been hypothesized to be a major determinant of mental health. Existing measures lack context-sensitivity (e.g, implicitly measuring it as a trait) and/or reference to a limited time frame. We developed a short self-report measure that covers all facets of the construct and is context-sensitive. Data was collected from four separate samples (n = 744), including a community (n = 346), non-clinical (n = 188), and two clinical (n = 163 in- and outpatients and n = 47 inpatients) samples from a psychiatric hospital. Confirmatory Factor Analysis confirmed a one-factor structure with excellent reliability (Raykov's r = 0.91). Correlations with related (such as other measures of psychological flexibility and symptomatology) and unrelated constructs (such as age and sex) were all consistent with predictions. The Psy-Flex differentiated clinical and non-clinical samples and predicted unique variance in well-being. The Psy-Flex is an easily administrable questionnaire, useful in research and clinical settings. Its context-sensitive nature makes it applicable to repeated administrations aimed at capturing change.

1. Introduction

Psychological flexibility refers to the ability to be aware of and open to internal and external stimuli as they occur while choosing to act in ways that are consistent with what a person deeply cares about (Hayes, Strohsahl & Wilson, 2012a). Theoretical conceptualizations generally refer to six interrelated skills that comprise psychological flexibility and that are targeted in treatments that aim to promote it (i.e., Acceptance and Commitment Therapy; ACT): acceptance (open and flexible embrace of moment-to-moment experience); defusion (creating contexts in which the function of unhelpful cognitions is diminished); present moment awareness (non-judgmental awareness of ongoing psychological and environmental events); stable self-awareness (relating from a perspective of I-here-now); values (chosen qualities of purposive action); and committed action (engagement in activities that are a manifestation of chosen values) (Hayes, 2006; Hayes, Strosahl, & Wilson, 2012). Psychological flexibility is seen as an important determinant of

mental health and well-being and psychological inflexibility is associated with numerous indices of dysfunction and psychopathology (Gloster, Meyer, & Lieb, 2017; Hayes, Luoma, Bond, Masuda, & Lillis, 2006; Kashdan & Rottenberg, 2010).

The conceptualization of psychological flexibility within ACT theory as a set of interconnected skills suggests that it is comprised of trainable and malleable behaviors. This has been demonstrated in multiple clinical and non-clinical intervention studies (Gloster, Walder, Levin, Twohig, & Karekla, 2020; Hofer et al., 2018; Levin, Hildebrandt, Lillis, & Hayes, 2012). It also suggests that the skills are flexibly applied and utilized in particular situations as needed in a particular moment. Given this conceptualization, the utility of such skills is determined upon the fit between the skill and the situations one finds themselves in. Thus, it is not automatically advantageous for an individual to be high or low on any particular skill per se, but rather to be able to apply the skill as needed in any given context. Simultaneously, psychological flexibility, like any collection of skills, profits from regular practice. That is, the way

---

an individual utilizes skills tends to occur more often and be similar over time, at least within similar contexts. Clinically, the regular practice aspect of psychological flexibility suggests that if someone is able to exhibit a skill in a challenging situation (e.g., urges to consume, depressive thoughts, anxiety provoking stimuli, etc.), then there is an increased probability (but not a guarantee) that the skill is used again the next time it is needed.

1.1. Measurement considerations

Psychological flexibility is increasingly studied in various contexts including clinical studies. Given that the construct has been hypothesized to be a major determinant of mental health, any assessment measure should ideally be able to capture this in clinical and non-clinical samples. Multiple questionnaires of psychological flexibility have been developed. See Table 1 for an overview of the existing questionnaires and their characteristics.

When people respond to items in questionnaires, they are influenced by internal and external contextual effects (Sudman, Bradburn, & Schwarz, 1996). Context effects are caused by words that orient the respondent to contemplate some experiences during a certain time period and not others. For example, items can be framed such that they measure a tendency to do something (e.g., "mostly true") or they can specify the conditions under which it is most important or necessary to engage in specific behavior (i.e., situational context). Items that do not specify situational context cues implicitly ask people to consider how they generally responded across all experiences they have had. This makes the assumption that these skills are fixed-like traits. The importance of defining the situational context in which psychological flexibility is used lies in the assumption that it is not necessary to maximize or even utilize a psychological flexibility skill in all situations. Rather, there are natural fluctuations in one's ability to employ these skills and in the importance of doing so. This is analogous to wearing the right clothes for a given weather. Context-sensitivity is also important because the underlying theory of psychological flexibility, namely contextual behaviorism, stipulates that behavior always happens in context (Hayes, Barnes-Holmes, & Wilson, 2012). Thus, the context-sensitivity of a measure informs about the degree to which a person can appropriately utilize and benefit from a skill: negating such an item might entail acknowledging that whereas some skill level is present, the respondent is unable to apply it when needed. This is often the case with patients presenting for treatment (Levin et al., 2012). Research has begun to document that compared to measures that are more trait-like, contextually sensitive measures demonstrate more treatment sensitivity (Benoy, Knitter, Schumann, et al., 2019) and better incremental validity (Ong, Lee, Levin, & Twohig, 2019).

Time is another important contextual variable. Existing measures of psychological flexibility or experiential avoidance generally do not stipulate a time frame to which a respondent is asked to relate when answering the items. This, too, contributes to the trait-like conceptualization of these measures. Taken literally, the absence of a time frame implies the respondent needs to consider all events across their life and then choose a response category (e.g., "often") that reflects this. Given that this is impossible to do, respondents engage in heuristics that lead to bias and inaccuracies (Menon, 1994; Walentynowicz, Schneider, & Stone, 2018). If one aim of a measure is to be treatment sensitive, then setting a concrete timeframe can support the respondent in focusing on recent memories, therefore providing more accurate responses (Rinner et al., 2019). To date, only two questionnaires measuring psychological flexibly included time frames in their instructions: the Multidimensional Psychological Flexibility Inventory (MPFI; Rolffs, Rogge, & Wilson, 2018) and the Openness and Engagement State Questionnaire (OESQ; Benoy, Knitter, Knellwolf, et al., 2019). The OESQ (Benoy, Knitter, Knellwolf, et al., 2019) outperformed the Attention and Action Questionnaire (AAQ-II; Bond et al., 2011) with respect to treatment sensitivity (i.e., treatment related change of a measure) (Benoy, Knitter, Schumann, et al., 2019). That is, a 2.08 times greater effect size was observed among the same patients by simply adding symptom-specific language (i.e., anxiety, worry, depression) and stipulating the time frame of the items (i.e., in the last week) (Benoy, Knitter, Schumann, et al., 2019). The lack of time specification and implicitly formulating psychological flexibility as trait may be appropriate for some situations but is of clear importance when estimating the effect size of an intervention due to the increase in treatment sensitivity (Benoy, Knitter, Schumann, et al., 2019).

Existing questionnaires differ further with respect to the explicit measurement of all six skills that comprise psychological flexibility. For example, the AAQ-II (Bond et al., 2011) measures psychological inflexibility and experiential avoidance, although it has also been suggested that the AAQ-II more closely measures distress than abilities of acceptance (Wolgast, 2014). Similarly, the Multidimensional Experiential Avoidance Questionnaire (MEAQ; Gámez, Chmielewski, Kotov, Ruggero, & Watson, 2011) and the Brief Experiential Avoidance Questionnaire (BEAQ; Gámez et al., 2014) both focus on experiential avoidance, which is a part of psychological inflexibility. The Everyday Psychological Inflexibility Checklist (EPIC; Thompson, Bond, & Lloyd, 2019) focuses on avoidance and behavioral rigidity. These questionnaires leave some aspects of psychological flexibility unassessed, even though information about where a skill is lacking could be important to therapy conceptualization. For example, actively working with values can be an important precursor to decreasing suffering in patients (Gloster, Klotsche, et al., 2017). This finding encourages the inclusion of *values* and *committed action* into measures to ensure that all relevant information is present before starting with an intervention. To date, two scales assess all skills of psychological flexibility. One, the MPFI (Rolffs et al., 2018), additionally assesses the adaptive and problematic

Table 1
Characteristics of questionnaires that measure psychological flexibility.

Name	Abbreviation	Citation	Time frame	Covers all skills	Context specific	No. of Items
Acceptance and Action Questionnaire-II	AAQ-II	Bond et al. (2011)	–	No	No	7
Comprehensive assessment of Acceptance and Commitment Therapy processes	CompACT	Francis et al. (2016)	–	Yes	Few	23
Comprehensive assessment of Acceptance and Commitment Therapy processes – Abbreviated version	CompACT-8	Dawson and Golijani-Moghaddam (2020)	–	Yes	Few	8
Multidimensional Psychological Flexibility Inventory	MPFI	Rolffs et al. (2018)	2 Weeks	Yes	Few	60
Multidimensional Experiential Avoidance Questionnaire	MEAQ	Gámez et al. (2011)	–	No	No	62
Brief Experiential Avoidance Questionnaire	BEAQ	Gámez et al. (2014)	–	No	No	15
Everyday Psychological Inflexibility Checklist	EPIC	Thompson et al. (2019)	–	No	Few	7
Open and Engaged State Questionnaire	OESQ	Benoy, Knitter, Knellwolf, et al. (2019)	1 Week	No	Yes	4
Psy-Flex	Psy-Flex	This manuscript.	1 Week	Yes	Yes	6

Note. Time frame refers to the time frame the respondent was asked to consider when answering the questions. Covers all skills reflects if the questionnaire has items representing all skills of psychological flexibility.

expression of each skill (e.g., fusion and defusion). The CompACT (Francis, Dawson, & Golijani-Moghaddam, 2016) also measures all six skills of psychological flexibility, albeit grouped into openness, awareness and active engagement.

Lastly, another point to consider when constructing a measure is the time it takes to fill it out. Especially for settings in which repeated assessments are necessary, the respondents are often asked to fill out the same questionnaire numerous times. Measures such as the AAQ-II (Bond et al., 2011), OESQ (Benoy, Knitter, Knellwolf, et al., 2019), EPIC (Thompson et al., 2019), and the CompACT-8 (Dawson & Golijani-Moghaddam, 2020) all contain fewer than 10 items, thereby keeping the response burden low. Generally, the aim of a short questionnaire is that it takes no longer than a couple of minutes and so that they can easily be integrated into clinical and research routines. The importance of brevity is documented by the fact that respondents' answers become less reliable as a function of length, but not repetition (Eisele et al., 2020).

1.2. The present study

The aim of this study was to develop and evaluate a brief and clinically useful self-report measure called Psy-Flex that assesses psychological flexibility. We aimed to include all six skills assumed to comprise psychological flexibility according to ACT theory so it can be used by clinicians for treatment planning. We also aimed to create a clinically useful measure that a) is sensitive to change by formulating items in a context-sensitive manner and explicitly including a time reference in the items; and b) differentiates clinical and non-clinical samples. The resultant measure was investigated in a two-step process. We, first, examined the general psychometric properties of the Psy-Flex. Second, we tested specific hypotheses regarding the clinical utility of the measure. Data for our examinations came from four independent samples comprised of both predominately healthy participants (i.e., community samples) and help-seeking individuals (i.e., individuals presenting for psychotherapy). Following the psychometric evaluation of factorial validity and reliability, we tested the following hypotheses. First, we hypothesized that participants' responses on the Psy-Flex would correlate positively with measures of mental health and well-being and negatively with measures of psychopathology (convergent validity). Second, we hypothesized that they would not correlate with non-psychological variables (divergent validity). Third, we hypothesized that participants' responses would differentiate clinical from non-clinical samples (discriminant validity). Finally, we hypothesized that the scores on the Psy-Flex would explain unique variance and therefore moderate the relationship between measures of pathology and well-being (incremental validity).

2. Method

2.1. Scale development

The Psy-Flex expanded on previously developed questionnaires that assess psychological flexibility with enhanced contextual cues. This included specifying specific emotions (e.g., anxiety) instead of generic terms (e.g., feelings) as well as stipulating time scales (e.g., last seven days vs. no specification of time). These contextual cues resulted in greater treatment sensitivity than measures without these elements (Benoy, Knitter, Schumann, et al., 2019). Unfortunately, at that time no existing scale measured all six skills of the psychological flexibility model. As such, we set out to utilize the advantages of contextual cues while also measuring all six skills.

Using the previous development as a starting point, several ACT experts with many years' experience in practicing, supervising, teaching, and publishing about ACT, developed the questionnaire in the following systematic manner. We first conducted a literature review on current conceptualizations of psychological flexibility including

theoretical and empirical studies (e.g., Hayes, Strosahl, & Wilson, 2012; Kashdan & Rottenberg, 2010; Wilson, Hayes, Biglan, & Embry, 2014). Then, existing questionnaires were collected that measure either psychological flexibility, experiential avoidance, or components thereof (e.g., Acceptance and Action Questionnaire, AAQ-II; Valued Living Questionnaire, VLQ; Cognitive Fusion Questionnaire, CFQ-7; Fear Questionnaire, FQ; Five Facets of Mindfulness Questionnaire, FFMQ; etc.). The experts then conducted a close examination of all the items used in prior questionnaires to get detailed knowledge about how psychological flexibility skills had been measured to date. Experts (experienced psychotherapists and ACT researchers) independently selected items that best captured each of the psychological flexibility skills, with the goal of including one item for each skill. The items in this pool were then shaped to (1) accentuate the contextual nature we were aiming for. This involved stipulating the conditions under which the skills are most important for a person. For example, mindfulness may be generally good, but the skill is especially useful if it can be deployed by an individual in turbulent times and not simply when quietly practicing in a distraction-free room. We also shaped the items (2) so that they specified the time frame for each question. That is, we formulated the items so that they could at least theoretically change across a day and formulated the instructions so that the time frame was clear to participants (i.e., "The questions refer to your experiences in the last seven days."). This process was accompanied by on-going, in-depth theoretical discussions of the items. This resulted in changes of wording until the experts were satisfied that the items adequately measured each skill. This list of items was then given to experts in the fields of psychology and psychiatry (ACT clinicians, psychiatric nurses trained in ACT) for critical examination and tested on a small sample of mental health patients of varying levels of education and therapy experience. Feedback from this round was integrated to ensure that the items were easy to understand. They were again adapted until a final form of 6 items, one per psychological flexibility skill, was arrived at. In a next step, this newly developed questionnaire (see Appendix) was then used in several studies to check its construct validity and reliability.

The six items measure psychological flexibility in a brief and context-sensitive manner. Each item refers to one of the core skills that ACT focuses on when developing psychological flexibility and well-being with a patient. Wording is constructed to be contextually sensitive to highlight the situations in which using each particular skill becomes useful. Items are rated on a scale from 5 ("very often") to 1 ("very rarely") and then summed. The score is then inter-preted such that higher scores represent higher psychological flexibility. The items are formulated such that in a particular situation they assess the presence of a skill rather than its absence. Therefore, the Psy-Flex avoids measuring the absence of the targeted skills inherent in psychological flexibility, a claim that other questionnaires of this construct have failed to fulfill (Gloster, Klotsche, Chaker, Hummel, & Hoyer, 2011; Wolgast, 2014).

2.2. Participants and procedure

The data included in this study stemmed from four distinct samples encompassing different target populations of both clinical and non-clinical participants. The combined sample with complete Psy-Flex scores resulted in a total of n = 744 participants with a mean age of 33.16 (SD = 12.60, 18–91 years). Of these, 534 participants (Samples 1 and 2) comprise the non-clinical portion of the sample, as they were all recruited from the general population, and 210 participants comprise the clinical portion of the sample (Samples 3 and 4), as all participants presented with a mental health diagnosis. Detailed descriptive statistics of the four samples can be found in Table 2. For all samples, participants were informed about the study's purpose and gave informed consent before filling out any questionnaires. The studies were all approved by Ethics Committees.

Table 2

Characteristics of four samples.

	Sample 1 (community)	Sample 2 (couples)	Sample 3 (patient)	Sample 4 (patient)
	N = 346	N = 188	N = 47	N = 163
Age				
Mean (SD)	31.42 (12.05)	32.82 (13.43)	38.96 (13.60)	35.60 (11.63)
Sex				
Females N (%)	228 (65.90)	97 (51.60)	24 (51.06)	83 (50.92)
Relationship Status				
Single	124 (35.94)	0 (0.00)	23 (48.94)	84 (57.53)
Partnership	221 (64.06)	188 (100.00)	24 (51.16)	62 (42.47)
Income				
Low	114 (33.24)	111 (68.10)	18 (60.00)	59 (50.00)
Middle	159 (46.36)	44 (26.99)	8 (26.67)	37 (31.36)
High	70 (20.41)	8 (4.91)	4 (13.33)	22 (18.64)
Enough Income				
Mean	1.96 (0.75)	2.12 (0.91)	2.02 (0.81)	1.97 (0.81)

Note. Sample 1 = community; Sample 2 = couples; Sample 3 = inpatients; Sample 4 = in- and outpatients; Enough Income was defined as sufficient funds to be able to live and accumulate savings from 1 (very easily) to 4 (not at all), lower scores meaning more agreement with the statement.

2.3. Sample 1 (community members)

The data from the first sample was collected with an online survey. Participants were recruited via pamphlets, an online study platform, and social media. In total, the link was clicked 410 times. In 69.3% of the cases, the questionnaires were fully completed, in 15.1% of cases only part of the questionnaires were completed, but the participants were retained in the dataset and included in the analysis, to avoid bias caused by attrition. A check if missing values were in any way systematic found no association with age, sex, relationship status or income. Incomplete cases were therefore retained where statistical methods allowed it. A majority of the sample was female (65.9%) with a mean age of 31.42 years (SD_{age} = 12.05; range 18–91 years).

2.4. Sample 2 (couples)

Sample 2 consisted of romantic dyads and was taken from an ongoing study for couples. Participants came in for testing together. In total there were n = 188 participants, of which 51.60% were female. The average age was M_{age} = 32.82 (SD_{age} = 13.43) and ranged from 18 to 80 years. Participating couples were hetero- (94.68%) or homosexual (5.32%), and needed to be in a steady relationship for at least 6 months.

2.5. Sample 3 (inpatients)

The third sample was recruited among patients presenting for inpatient treatment to the psychiatric hospital in [removed for blinding] who took part in a longitudinal observation study (n = 47, 51.06% female, 20–64 years old, M_{age} = 38.96 years, SD_{age} = 13.6; [reference removed for blinding]). The participants' primary diagnoses were Major Depressive Disorder (42.45%), followed by Panic Disorder (19.15%), Obsessive-Compulsive Disorder (12.77%), and Social Phobia (10.64%). Of the sample, 63.83% fulfilled the criteria for at least one comorbid disorder. The patients stayed in the clinic for 66.32 (SD = 22.7) days on average.

2.6. Sample 4 (in- and outpatients)

The fourth sample was recruited among patients presenting for inpatient and outpatient treatment within an ongoing longitudinal clinical trial ([reference removed for blinding]) with an inpatient and an outpatient group from the psychiatric hospital in [removed for blinding]

(n = 163, 50.92% female, 18–65 years old, M_{age} = 35.60 years, SD_{age} = 11.63). The participants presented with a range of diagnoses, mainly mood disorders (33.13%) and anxiety disorders (33.13%), and 71.17% of the patients had at least one comorbid disorder.

2.7. Measures

2.7.1. Acceptance and Action Questionnaire (AAQ-II)

The AAQ-II (Bond et al., 2011) is a 7-item questionnaire assessing experiential avoidance and psychological inflexibility. The AAQ-II has been widely used and has been shown to have acceptable internal consistency with α-coefficients of 0.78–0.87 (Bond et al., 2011; Fledderus, Oude Voshaar, ten Klooster, & Bohlmeijer, 2012; Krafft, Ong, Twohig, & Levin, 2019; Monestès et al., 2018) and between α = 0.87–0.93 in the present samples. It correlates highly with measurements of depression (BDI, r = 0.71) and is not dependent on age, gender, race, or clinical status (Karekla & Michaelides, 2017). Items are rated on a scale from 1 ("never true") to 7 ("always true") and encompass descriptors of negative internal experiences. A sum score is interpreted such that higher scores indicate higher levels of inflexibility (Bond et al., 2011). The authors demonstrate that the AAQ-II is associated with other measures of psychopathology, such as anxiety and depressive symptoms (Bond et al., 2011).

2.7.2. Cognitive Fusion Questionnaire (CFQ-7)

The CFQ-7 (Gillanders et al., 2014) is a short, 7-item questionnaire that assesses cognitive fusion. The CFQ has shown good psychometric properties with an excellent Cronbach's Alpha of α = 0.94–0.95 (China, Hansen, Gillanders, & Benninghoven, 2018) and between α = 0.91–0.95 in the present samples. It has been shown to be invariant across genders and clinical status, with good convergent and discriminant validity (Zacharia, Ioannou, Theofanous, Vasiliou, & Karekla, 2021). Items are rated on a scale from 1 ("never true") to 7 ("always true") and include statements that represent cognitive fusion such as "I get so caught up in my thoughts that I am unable to do the things that I most want to do".

2.7.3. Five Facets Mindfulness Questionnaire - short form (FFMQ-SF)

The FFMQ-SF (Bohlmeijer, ten Klooster, Fledderus, Veehof, & Baer, 2011) is a 24-item questionnaire that assesses five separate but related factors of mindful behavior (Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006), which constitutes another component of psychological flexibility. All factors (observing, describing, acting with awareness, nonjudging of inner experience, and nonreactivity to inner experience) have shown to have adequate internal consistency, with Cronbach's Alpha ranging from α = 0.75–0.87 (Bohlmeijer et al., 2011) and between α = 0.82–0.84 in the present samples. Items are rated on a scale from 1 ("never or very rarely true") to 5 ("very often or always true") and a sum score is interpreted such that higher scores equal more mindfulness.

2.7.4. Mental Health Continuum Short Form (MHC-SF)

The MHC (Keyes, 2002) is a widely used measure of well-being. The short version of the MHC (MHC-SF; Lamers, Westerhof, Bohlmeijer, Ten Klooster, & Keyes, 2011) consists of 14 items that show high total internal consistency (Cronbach's Alpha of $\alpha > 0.89$) and between α = 0.90–0.92 in the present samples. Each item assesses how often a statement was true during the past month, ranging from 0 ("never") to 5 ("almost every day"). Statements include emotional (e.g., "During the past month, how often did you feel happy?"), social (e.g., "During the past month, how often did you feel that you had something important to contribute to society?"), and psychological (e.g., "During the past month, how often did you feel that you liked most parts of your personality?") aspects of well-being.

2.7.5. Meaning in Life Questionnaire (MLQ)

The MLQ (Steger, Frazier, Oishi, & Kaler, 2006) is a 10-item questionnaire that assesses two factors: the extent to which meaning in life

and searching for meaning in life are present. Its two subscales show good internal consistency: Presence $\alpha = .86$; and Search: $\alpha = 0.89$ and between $\alpha = 0.87–0.89$ (Presence) and $\alpha = 0.90$ (Search) in the present samples. The questionnaire assesses its two factors by having respondents answer statements such as "I understand my life's meaning" or "I am seeking a purpose or mission for my life" on a scale from 1 ("absolutely untrue") to 7 ("absolutely true") with higher values reflecting higher experienced meaning or higher searching for meaning in life.

2.7.6. Brief Symptom Checklist (BSCL)

The BSCL (Franke, Stenzel, Rank, Herbold, & Küch, 2015) is a broad 53-item self-report inventory assessing levels of psychopathology on nine different subscales. Subscales show sufficient to good internal consistency ranging from Cronbach's Alpha of $\alpha = 0.75$ to Cronbach's Alpha of $\alpha = 0.90$ and between $\alpha = 0.74–0.86$ in samples 3 and 4. Each symptom is rated on a scale from 0 ("not at all") to 4 ("extremely"). The BSCL is a standard measure administered in most psychiatric hospitals in [removed for blinding] and is used for treatment quality assessment (Nationaler Verein für Qualitätsentwicklung in Spitälern und Kliniken (ANQ), 2016).

2.7.7. Brief Symptom Inventory (BSI-18)

The BSI-18 (Derogatis, 1993) is a short version of the BSCL consisting of only three subscales: somatization, depression, and anxiety. The BSI has been found to have high internal consistency ($\alpha = 0.91$) for a heterogeneous sample of psychotherapy clients (Franke et al., 2011) and $\alpha = 0.93$ in sample 1. Symptoms are rated on a scale from 0 ("not at all") to 4 ("extremely") and higher scored equate to higher symptom severity.

2.8. Data analysis

Psychometric analyses were conducted using the 6-item version of the Psy-Flex. To check construct validity, the inter-item-correlation was calculated for each subsample and the sample as a whole using Pearson's correlation coefficient or point-biserial correlations for dichotomous variables (i.e., sex). The reliability of the scale was tested based on Raykov's reliability coefficient for multi-level data (Raykov, 2001). Internal factorial validity was calculated using confirmatory factor analysis (CFA). With the CFA it was tested whether the Psy-Flex factor structure was one-dimensional as posited. Initially, an unrestricted one-factor model was fit and examined for sources of possible model misspecification, as indicated by the modification index (Kline, 2015). The one-factor model was tested for each subsample separately, in a first step. Acceptable model fit was assessed with root means square error of approximation (RMSEA; ≤ 0.08) with 90% confidence interval, standardized root-mean-square residual (SRMR; ≤ 0.08) and comparative fit index—Tucker Lewis index (CFI–TLI; ≥ 0.9), applying the guidelines for these fit indices suggested by (Hu & Bentler, 1999). CFA and multiple-group CFA were conducted in MPLUS 5.1 (Muthén & Muthén, 2018), applying the weighted least squares estimator based on a diagonal weight matrix. Acceptable model fit of RMSEA values between 0.05 and 0.08 (Fabrigar, Wegener, MacCallum, & Strahan, 1999) in each subsample provided the base for further steps. A multiple-group CFA was used to test the stability, generalizability, and population homogeneity of the one-factor solution across the four samples. A configural invariance model was specified to test whether the one-factor structure was valid across the four subsamples, followed by the metric invariance model (all factor loadings held equal across groups) and scalar invariance model (all factor loadings and intercepts held equal across groups). The improvement in model fit was tested by the scaled chi-square difference test for model comparison. In addition, the rule of thumb of Chen (2007) was used in order to compare the improvement in model fit. If the CFI decreases by more than 0.02 and the RMSEA increases by more than 0.015 than the two compared models differ remarkably in model fit.

Hypotheses 1 and 2 (convergent and divergent validity) were calculated by correlating the Psy-Flex total score with other measures. A high correlation with measures of specific components of psychological flexibility was expected (i.e., AAQ-II, CFQ-7, FFMQ-SF, MHC-SF, MLQ). Moderate, negative correlations with measures of symptomatology was excepted (i.e., BSCL, BSI-18). Using the interpretation guidelines of Cohen (1977), a correlation over 0.5 was considered high, between 0.3 and 0.5 was considered moderate, and between 0.1 and 0.3 was considered low, while any correlations below 0.1 were deemed negligible.

For Hypothesis 3, to check whether the Psy-Flex can accurately differentiate between a non-clinical and a clinical sample, the sum scores of the Psy-Flex of each sample were compared using an analysis of variance, with a priori defined contrasts. First, the normative samples as well as the clinical samples were compared to each other to check for between-group differences. Then, the two clinical samples were compared against the two normative samples, where a pronounced difference was expected.

For Hypothesis 4 (unique variance), a hierarchical regression model was set up to calculate the beta weights of symptomatology (BSCL and BSI scores respectively). In a next step, the Psy-Flex scores were added as a predictor and the two resulting models were compared and the difference in r^2 was tested for significance.

3. Results

3.1. Psychometric properties

Across all four samples, the average total Psy-Flex score was 19.9 (SD = 4.92). Participants responded using the full response range, from the minimum score of 6 to the maximum score of 30. Standard deviations per item ranged from 0.94 to 1.15, indicating similar variation among all items (see Table 3). All scales showed an approximately symmetric distribution, with the exception of BSI-18 who showed a right-skewed distribution. Only very few outliers were observed in our samples. For more details on basic statistics, total and item scores, please refer to Supplementary Table A.

All items showed sufficient variation, meaning they were able to capture individual differences in answering patterns. Inter-item correlation ranged from 0.38 (Item 2 and 6) to 0.69 (Item 3 and 4). All items were positively correlated and showed good coefficients of homogeneity, with no item below the threshold of 0.30 and an overall scale homogeneity of $H = 0.63$ (SD = 0.02). All items were locally independent. Detailed results of the individual items across all samples can be found in the supplementary materials (Supplementary Table A).

3.2. Factor structure and confirmatory factor analysis

We tested the one-factor solution for the Psy-Flex in the four subsamples. However, the one-factor model did not fit the data well (SRMR = 0.052; RMSEA = 0.17; CFI = 0.91; TLI = 0.85). Localized area of strain in that solution suggested correlated residuals between Items 5 and 6 (modification index 136). The model fit of the model with correlated residuals significantly improved ($-2\Delta LL(1) = 136.2, p < .001$,

Table 3

Mean and median values for Psy-Flex total score across samples.

	N	Mean (SD)	Median (P25, P75)
Psy-Flex			
Across All Samples	744	19.9 (4.9)	20.3 (16.3, 24.3)
Sample 1 (community)	346	20.4 (4.3)	20.3 (17.3, 23.3)
Sample 2 (couples)	188	23.4 (3.4)	23.6 (21.4, 25.7)
Sample 3 (patient)	47	16.1 (3.8)	16.3 (14.3, 18.3)
Sample 4 (patient)	163	15.9 (4.3)	15.3 (13.3, 18.3)

Note. P25 = 25th percentile; P75 = 75th percentile.

LL = log likelihood), with an acceptable overall model fit for the one-factor solution (SRMR = 0.031; RMSEA = 0.058; CFI = 0.98; TLI = 0.96). The one-factor structure could be replicated in each of the four subsamples and in the combined clinical and non-clinical samples (CFA fit indices in Table 4).

In the multiple group CFA an initial configural invariance model was established. Initially, a configural invariance model was established with a residual covariance between the items 5 and 6 to test whether the hypothesized one-factor structure was valid across the two subsamples. This model resulted in an acceptable model fit (χ^2 (16) = 45.2, $p < .001$, CFI = 0.98; TLI = 0.96, RMSEA = 0.073 (90% CI = 0.048, 0.098), SRMR = 0.04). The metric invariance model was tested in the next step. The metric invariance model fitted the data well (χ^2 (22) = 50.4, $p < .001$, CFI = 0.98; TLI = 0.97, RMSEA = 0.061 (90% CI = 0.039, 0.083), SRMR = 0.04). The scaled chi-square difference test showed that the metric invariance model was not statistically significantly worse than the configural invariance model, χ^2 (6) = 5.49, $p = .483$. This suggests that the factor loadings to the latent construct are comparable between samples. In a next step a scalar invariance model was tested. The scalar invariance model did not fit the data (χ^2 (28) = 238.4, $p < .001$, CFI = 0.82; TLI = 0.80, RMSEA = 0.15 (90% CI = 0.13, 0.17), SRMR = 0.25). The scaled chi-square difference test showed that the scalar invariance model was significantly worse from the metric invariance model, χ^2 (6) = 182.7, $p < .001$. In addition, the change in CFI and RMSEA was greater than 0.02 and 0.015 confirming the worse fit of the scalar invariance model. This suggests that clinical and non-clinical samples have different intercepts, as would be expected. We did not test stricter forms of measurement invariance such as residual invariance model based on the worse model fit for the scalar invariance model.

3.3. Reliability

The CFA showed that the Psy-Flex fits a one-factor model as hypothesized. We also analyzed the internal consistency of the factor by calculating the Raykov's coefficient (2001) for the whole sample. Reliability of Psy-Flex scores based on the Raykov estimation (2001) per sample was as follows: 0.90 (Sample 1); 0.91 (Sample 2); 0.78 (Sample 3); and 0.97 (Sample 4). Across all samples the reliability was 0.91. Corrected inter-item correlations according to Raykov estimation ranged between 0.63 and 0.79, supporting the results of the CFA that they measure the same construct.

3.4. Tests of construct and incremental validity

Hypothesis 1. Zero Order Correlations With Similar Constructs
To measure convergent validity, we compared the Psy-Flex to other scales that measure comparable constructs, such as elements of psychological flexibility and well-being. Detailed results of the convergent validity measures can be found in Table 5.

3.4.1. Measures of psychological flexibility and its components

The Psy-Flex and the AAQ-II, which measures experiential avoidance, correlated highly and negatively, with r = −0.71. The Psy-Flex also correlates highly and negatively with the CFQ-7, which measures cognitive fusion (r = −0.71). Furthermore, the Psy-Flex correlated with all subscales of the FFMQ-SF. The correlation was the highest with the Non-react subscale (r = 0.53) and lowest for the Observe subscale (r = 0.25).

3.4.2. Measures of well-being and meaning

The Psy-Flex also correlated positively with the MHC-SF and MLQ, both measures of well-being. Higher scores on the Psy-Flex coincided

Table 4
Model fit of the one-factor model of the Psy-Flex assuming correlated residuals for the items five and six using Confirmatory Factor Analysis.

Sample	χ^2 (df), p value	CFI	TLI	RMSEA (90% CI)	SRMR
Total	47.3 (8), p < .001	.98	.96	.076 (.05–.11)	.031
Sample 1	23.7 (8), p = .001	.98	.96	.080 (.04–.12)	.036
Sample 2	13.2 (8), p = .104	.98	.96	.059 (.001–.11)	.035
Sample 3	7.2 (8), p = .518	.98	.96	.001 (.001–.17)	.043
Sample 4	13.7 (8), p = .091	1.00	1.00	.067 (.001–.13)	.051
Non-clinical samples (Sample 1 and Sample 2)	41.2 (8), p < .001	.97	.94	.079 (.04–.11)	.037
Clinical samples (Sample 3 and Sample 4)	13.2 (8), p = .105	.99	.97	.057 (.001–.11)	.037

Note. Sample 1 = community; Sample 2 = couples, Sample 3 = inpatients; Sample 4 = in- and outpatients; SRMR = standardized root-mean-square residual; CFI = comparative fit index; TLI = Tucker Lewis index.

Table 5
Convergent validity of Psy-Flex with related constructs.

	Mean (SD)	Pearson's r	Samples used
ACT Measures			
AAQ-II	24.10 (10.49)	-.71*	1,3,4
CFQ-7	26.96 (10.82)	-.71*	1,3,4
FFMQ-SF			3,4
- Non-react	2.46 (0.67)	.53*	
- Non-judge	2.79 (0.80)	.32*	
- Describe	3.25 (0.86)	.33*	
- Observe	3.52 (0.87)	.25*	
- Act aware	3.04 (0.70)	.34*	
Well-Being			
MHC-SF	52.42 (14.91)	.59*	1,2,3,4
- Emotional	12.39 (3.91)	.59*	
- Social	16.04 (5.86)	.40*	
- Psychological	23.98 (6.85)	.60*	
MLQ			2,3,4
- Presence of Meaning	22.73 (7.92)	.61*	
- Search for Meaning	20.88 (8.67)	-.33*	
Symptomatology			
BSI-18 or BSCL	27.21 (14.85)	-.53*	1, 3, 4
- Somatization	8.55 (4.48)	-.42*	
- Depression	10.82 (5.29)	-.54*	
- Anxiety	10.86 (4.84)	-.53*	

Note. Sample 1 = community; Sample 2 = couples, Sample 3 = inpatients; Sample 4 = in- and outpatients; as not every measure was collected in every sample, the column "Samples used" indicates which samples were used for the calculation; Psy-Flex = Psychological Flexibility; AAQ = Action and Avoidance Questionnaire; CFQ-7 = Cognitive Fusion Questionnaire; FFMQ = Five Facet Mindfulness Questionnaire; MHC-SF = Mental Health Continuum Short Form; MLQ = Meaning in Life Questionnaire; BSI-18 = Brief Symptom Inventory; BSCL = Brief Symptom Checklist, but only the 3 Subscale (18 Items) used in the BSI-18.
*p < .001.

with higher scores on the emotional, social, and psychological well-being subscale of the MHC-SF respectively (r = 0.59, 0.40, and 0.60). The Psy-Flex correlated moderately with the MLQ subscale of presence of meaning (r = 0.61), while the Psy-Flex showed a low negative correlation with the MLQ subscale of search for meaning (r = −0.33).

3.4.3. Measures of symptomatology

Psy-Flex scores were also associated with overall symptomatology. In the analysis, we concentrated on the selection of subscales relevant to the most prevalent disorders (i.e., somatic, mood, and anxiety disorders). For the clinical samples and the community sample, Psy-Flex scores were moderately negatively associated with the depression subscale, the anxiety subscale, and the somatization subscale, indicating an interdependence between the Psy-Flex score and the amount and severity of symptoms that a person reported on the subscales. Detailed results of Pearson's correlation with measures of symptomatology can be found in Table 5. Detailed results for all measures can be found in the Appendix (Supplementary Table A).

Hypothesis 2. Divergent Validity for Non-Psychological Variables

Commonly tested non-psychological variables (sex and age) were, overall, not related to the construct of Psy-Flex. Across all samples, neither age (F (1, 690) = 0.537, p = .464) nor sex (F (1, 692) = 3.348, p = .068) were significantly associated with Psy-Flex outcomes. When examining each subsample for significant correlations, age was significantly associated with Psy-Flex for sample 1, and sex was significantly associated with Psy-Flex for sample 2 and sample 4 (for details see Table 6).

Hypothesis 3. Differences Between Non-Clinical and Clinical Populations (Discriminant Validity)

The non-clinical and the clinical population's Psy-Flex scores were compared using an Analysis of Variance (ANOVA). Consistent with our hypothesis, the different samples had significant between-group differences of mean Psy-Flex sum scores, F (1, 634) = 71.69, p < .001, η^2_p = .32. Post-hoc analyses using Tukey's HSD indicated significant differences on the Psy-Flex between all 4 samples (p < .001) except for the comparison of the two non-clinical samples (sample 1 and sample 2, p = .998). When comparing the non-clinical samples (sample 1 and sample 2) to the clinical samples (sample 3 and sample 4), the non-clinical group had higher scores on the Psy-Flex than the clinical group, as was expected: 23.32 (4.39) for the non-clinical population of sample 1 and sample 2, versus 17.59 (4.24) for the clinical population of sample 3 and sample 4, F (1, 632) = 242.9, p < .001, η^2_p = .25.

Hypothesis 4. Prediction of Unique Variance Above Established Constructs (Incremental Validity)

Results from the hierarchical regression analyses showed that the Psy-Flex predicted unique variance in scores of well-being above and beyond the variance explained by overall symptomatology. This pattern was observed for all samples tested, for the total score of the MHC as well as all subscales of the MHC except for the subscale psychological well-

Table 6
Divergent validity of Psy-Flex with non-psychological variables.

Construct	Measure	Sample	r with Psy-Flex
Age	Years	1	**.14***
	Years	2	.10
	Years	3	.13
	Years	4	.09
Sex	Female or Male	1	.06
	Female or Male	2	**.19***
	Female or Male	3	-.22
	Female or Male	4	**.13***

Note. Sample 1 = community; Sample 2 = couples, Sample 3 = inpatients; Sample 4 = in- and outpatients; significant correlations are indicated in bold, sex was coded 1 = female, 2 = male.
*p < .05.

being in sample 4. See Table 7 for details.

4. Discussion

Psychological flexibility has been called a fundamental component of mental health (Kashdan & Rottenberg, 2010). This study presented the results from the development and validation of a short, contextually sensitive questionnaire of psychological flexibility, called the Psy-Flex. Across four diverse clinical and non-clinical samples we observed that participants' responses were consistent with indices that support a one-factor structure and reliability, as well as solid evidence of convergent, divergent, and incremental validity. The Psy-Flex's ease of administration and ability to differentiate clinical and non-clinical samples as well as predicting unique variance of well-being is promising.

The one-factor solution was consistent with a priori conceptualization and theory (Hayes, Strosahl, & Wilson, 2012). Likewise, nearly all observed correlations were consistent with theoretical predictions, both in terms of direction and magnitude. It is noteworthy that the Psy-Flex was most strongly related to the AAQ-II but showed weaker inverse correlations with indices of symptomatology. This is important because the AAQ-II has been criticized for being too strongly correlated with symptomatology and as a result too much a measure of distress (Wolgast, 2014). Consistent with theoretical predictions, the Psy-Flex was also moderately correlated with measures of mindfulness, well-being, and meaning in life. The Psy-Flex was not related or only very weakly related to the non-psychological variables of age and sex. This is consistent with previous research (Christodoulou, Michaelides, & Karekla, 2018; Gloster et al., 2011; Karekla & Michaelides, 2017).

The ultimate goal of interventions targeting psychological flexibility is to help people increase their well-being. The Psy-Flex is advantageous in this regard, as it predicted unique variance in well-being above and beyond symptomatology. Previous investigations using the AAQ-II have shown that of psychological flexibility moderates the relationship between stressors and well-being in a representative population sample (Gloster, Meyer, & Lieb, 2017) and the relationship between somatization and quality of life in a community sample (Leonidou, Panayiotou, Bati, & Karekla, 2019). This study demonstrated prediction of unique variance in well-being in both clinical and non-clinical samples, which is a further indication of the questionnaire's utility. Future studies should consider how well measures of psychological flexibility predict well-being across various samples and settings and explore which measure or measures of psychological flexibility best predict well-being.

The Psy-Flex utilized situational and temporal specifiers to increase its context-sensitivity. The lack of such specifiers has been found to increase bias and inaccuracies (Fernandez-Ballesteros, 2004; Sudman et al., 1996) and lead to a lack of treatment sensitivity (Benoy, Knitter, Schumann, et al., 2019) and lower validity (Ong et al., 2019). Inclusion of contextual cues helps participants understand the items the way they were intended, thereby reducing confusion and ambiguity. Using such cues likely help respondents reflect on the utility of the respective skills (e.g., can I really use this skill when it matters most and in spite of difficulties), which may account for findings showing that including contextual cues increases the treatment sensitivity. Future developments of the measurement of psychological flexibility should include such cues and explore how they can be further used in ways that measure this in changing situations (e.g., event sampling methodology).

In order for assessments to be used in clinical settings, they must be economical in administration, understandable by participants with a wide variety of educational levels and have clinical utility. The Psy-Flex was designed for this purpose and several indices suggest it is recommendable for this context. First, the Psy-Flex is only six items long, which makes it extremely time-efficient to complete, as well as easy to evaluate and interpret. Inspection of each item can help identify whether one or more of the six skills is particularly low. This is especially true because the items focus on a short time span and the contexts when participants actually need the skills. This could potentially be useful in

Table 7
Hierarchical multiple regression analyses.

Predictor	Measure	Sample 1 (community) n = 346					Sample 3 (inpatients) n = 163					Sample 4 (in- and outpatients) n = 47				
		Model 1		Model 2		Δ in R² (p)	Model 1		Model 2		Δ in R² (p)	Model 1		Model 2		Δ in R² (p)
		β	p	β	p		β	p	β	p		β	p	β	p	
MHC-SF - Total score																
Brief Symptom Inventory	BSI	−0.26	<0.001	0.04	0.502		−0.31	0.045	−0.16	0.323		−0.27	0.001	−0.07	0.310	
Psychological Flexibility	Psy-Flex			0.47	<0.001	0.14 (**<0.001**)			0.38	0.019	0.12 (**0.014**)			0.63	<0.001	0.35 (**<0.001**)
MHC-SF - Emotional well-being																
Brief Symptom Inventory	BSI	−0.39	<0.001	−0.12	0.059		−0.44	0.003	−0.28	0.059		−0.44	0.003	−0.28	0.059	
Psychological Flexibility	Psy-Flex			0.42	<0.001	0.11 (**<0.001**)			0.40	0.009	0.13 (**0.006**)			0.40	0.009	0.13 (**0.006**)
MHC-SF - Psychological well-being																
Brief Symptom Inventory	BSI	−0.10	0.068	0.09	0.199		−0.09	0.589	0.00	0.986		−0.26	0.002	−0.09	0.219	
Psychological Flexibility	Psy-Flex			0.31	<0.001	0.20 (**<0.001**)			0.22	0.210	0.04 (0.191)			0.50	<0.001	0.23 (**<0.001**)
MHC-SF - Social well-being																
Brief Symptom Inventory	BSI	−0.26	<0.001	0.07	0.265		−0.44	0.003	−0.28	0.059		−0.30	<0.001	−0.10	0.142	
Psychological Flexibility	Psy-Flex			0.53	<0.001	0.17 (**<0.001**)			0.40	0.009	0.13 (**0.006**)			0.62	<0.001	0.35 (**<0.001**)

Note. Sample 1 = community; Sample 3 = inpatients; Sample 4 = in- and outpatients; Sample 2 did not have a questionnaire of symptomatology and is therefore excluded from this analysis. Psy-Flex = psychological flexibility, BSCL = Brief Symptom Checklist, BSI-18 = Brief Symptom Inventory. Significant changes changes have been marked in bold.

treatment planning for both respondents and clinicians alike. Some research has even suggested that tracking changes in ACT skills might also be useful for other psychotherapy schools (Probst et al., 2020). In addition, a short questionnaire like the Psy-Flex may easily be administered repeatedly to check for changes and to help the patient to quickly guide their development of psychological flexibility skills over time, and preliminary data suggests the Psy-Flex is more sensitive to treatment change than implicit trait measures of psychological flexibility (Benoy, Knitter, Knellwolf, et al., 2019). Finally, results showed that the Psy-Flex could discriminate between clinical and non-clinical samples and predict unique variance in well-being across samples. As such, the Psy-Flex appears to be relevant for multiple populations, including clinical settings and the general population (e.g., couples and community members).

This study needs to be interpreted in light of several limitations. First, with the exception of sample 1 (community sample), the samples were convenience samples and not specifically collected for the purpose of this study. This concern is mitigated, however, by the use of diverse clinical and non-clinical samples. Second, and related, not all constructs were collected for all samples. In particular, no symptomatology measures were collected for the couples (sample 2). Future research should address this. Third, other constructs that might inform on other aspects of the questionnaire such as intelligence level, reading level, etc., were not included. Once again, the use of four diverse samples helps mitigate this concern somewhat. Fourth, all data were collected using questionnaires and were therefore subjected to the biases inherent therein (Gloster et al., 2008; Rinner et al., 2019). Future research should consider using alternative methods such as experience sampling. Fifth, the errors in items 5 (values) and 6 (committed action) were correlated. Since committed action is not possible without values, these skills are theoretically connected, which might explain the error-correlation. Furthermore, both items use the word "important", linking them linguistically. We retained the error-correlation as we think this result is in line with theoretical and analytical assumptions. Further research should investigate how the skills of defining one's values and acting in line with them overlap. Finally, during the development of the scale, it was given to patients of various education levels to assess understanding of the items, but that information was not formally collected. Even though we found no indication in our data to suggest that participants were struggling to understand the Psy-Flex, future research could collect this data.

These limitations notwithstanding, the Psy-Flex is a short, context-sensitive questionnaire that measures the skills of psychological flexibility. The items of the Psy-Flex are sensitive to the circumstances and situations in which psychological flexibility skills are especially important to utilize. Results from this study show good psychometric properties and indices of clinical utility. Based on these results, the Psy-Flex may be economically and validly used in both research and clinical care. Future research should measure the Psy-Flex in additional populations and contexts.

Declaration of competing interest

Declaration of conflicts of interest: Given his role as Associate Editor, Andrew T. Gloster had no involvement in the peer-review of this article and had no access to information regarding its peer-review. Full responsibility for the editorial process for this article was delegated to an Associate Editor of JCBS and blinded reviewers. Other conflicts of interest: None.

Appendix Psy-Flex

The questions refer to your experiences in the **last seven days**.

1. Being present.

Even if I am somewhere else with my thoughts, I can focus on what's going on in important moments.

very often (5)	often (4)	from time to time (3)	seldom (2)	very seldom (1)

2. Being open for experiences.

If need be, I can let unpleasant thoughts and experiences happen without having to get rid of them immediately.

very often (5)	often (4)	from time to time (3)	seldom (2)	very seldom (1)

3. Leaving thoughts be.

I can look at hindering thoughts from a distance without letting them control me.

very often (5)	often (4)	from time to time (3)	seldom (2)	very seldom (1)

4. Steady self.

Even if thoughts and experiences are confusing me I can notice something like a steady core inside of me.

very often (5)	often (4)	from time to time (3)	seldom (2)	very seldom (1)

5. Awareness of one's own values.

I determine what's important for me and decide what I want to use my energy for.

very often (5)	often (4)	from time to time (3)	seldom (2)	very seldom (1)

6. Being engaged.

I engage thoroughly in things that are important, useful, or meaningful to me.

very often (5)	often (4)	from time to time (3)	seldom (2)	very seldom (1)

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jcbs.2021.09.001.

Funding information

References

ANQ Swiss National Association for Quality Development in Hospitals and Clinics. (2016). *Nationaler Verein für Qualitätsentwicklung in Spitälern und Kliniken: Publikationskonzept Stationäre Psychiatrie – Erwachsene*.

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

Benoy, C., Knitter, B., Knellwolf, L., Doering, S., Klotsche, J., & Gloster, A. T. (2019). Assessing psychological flexibility: Validation of the open and engaged state questionnaire. *Journal of Contextual Behavioral Science, 12*, 253–260. https://doi.org/10.1016/j.jcbs.2018.08.005

Benoy, C., Knitter, B., Schumann, I., Bader, K., Walter, M., & Gloster, A. T. (2019). Treatment sensitivity: Its importance in the measurement of psychological flexibility. *Journal of Contextual Behavioral Science, 13*(July), 121–125. https://doi.org/10.1016/j.jcbs.2019.07.005

Bohlmeijer, E., ten Klooster, P. M., Fledderus, M., Veehof, M., & Baer, R. (2011). Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. *Assessment, 18*(3), 308–320. https://doi.org/10.1177/1073191111408231

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., et al. (2011). Preliminary psychometric properties of the acceptance and action questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*(4), 676–688. https://doi.org/10.1016/j.beth.2011.03.007

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

China, C., Hansen, L. B., Gillanders, D. T., & Benninghoven, D. (2018). Concept and validation of the German version of the cognitive fusion questionnaire (CFQ-D). *Journal of Contextual Behavioral Science, 9*(June), 30–35. https://doi.org/10.1016/j.jcbs.2018.06.003

Christodoulou, A., Michaelides, M. P., & Karekla, M. (2018). Greek version of the avoidance and fusion questionnaire for youth: Psychometric evaluation and gender multigroup invariance in adolescents. *Journal of Psychoeducational Assessment, 36*(8), 844–849. https://doi.org/10.1177/0734282917713500

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* Academic Press.

Dawson, D. L., & Golijani-Moghaddam, N. (2020). COVID-19: Psychological flexibility, coping, mental health, and wellbeing in the UK during the pandemic. *Journal of Contextual Behavioral Science, 17*(June), 126–134. https://doi.org/10.1016/j.jcbs.2020.07.010

Derogatis, L. R. (1993). *Brief symptom inventory (BSI): Administration, scoring, and procedures manual* (3rd ed.). NCS Pearson.

Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., et al. (2020). The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment, 7001.* https://doi.org/10.1177/1073191120957102

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Fernandez-Ballesteros, R. (2004). Self-report questionnaires. In S. N. Haynes, E. M. Heiby, & M. Hersen (Eds.), *Comprehensive handbook of psychological assessment: Volume 3: Behavioral assessment* (pp. 194–221). John Wiley & Sons, Ltd.

Fledderus, M., Oude Voshaar, M. A. H., ten Klooster, P. M., & Bohlmeijer, E. T. (2012). Further evaluation of the psychometric properties of the acceptance and action questionnaire–II. *Psychological Assessment, 24*(4), 925–936. https://doi.org/10.1037/a0028200

Francis, A. W., Dawson, D. L., & Golijani-Moghaddam, N. (2016). The development and validation of the comprehensive assessment of acceptance and commitment therapy processes (CompACT). *Journal of Contextual Behavioral Science, 5*(3), 134–145. https://doi.org/10.1016/j.jcbs.2016.05.003

Franke, G. H., Ankerhold, A., Haase, M., Jäger, S., Tögel, C., Ulrich, C., et al. (2011). Der Einsatz des Brief Symptom Inventory 18 (BSI-18) bei Psychotherapiepatienten. *PPmP - Psychotherapie · Psychosomatik · Medizinische Psychologie, 61*, 82–86. https://doi.org/10.1055/s-0030-1270518, 02.

Franke, G. H., Stenzel, S., Rank, C., Herbold, D., & Küch, D. (2015). Die Brief Symptom Checklist (BSCL) im Einsatz bei Patientinnen und Patienten der orthopädischen Rehabilitation. In *Conference paper: Jahrestagung Des AK Klinische psychologie* (pp. 1–9). Berlin, September.

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

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

Gillanders, D. T., Bolderston, H., Bond, F. W., Dempster, M., Flaxman, P. E., Campbell, L., et al. (2014). The development and initial validation of the cognitive fusion questionnaire. *Behavior Therapy, 45*(1), 83–101. https://doi.org/10.1016/j.beth.2013.09.001

Gloster, A. T., Klotsche, J., Chaker, S., Hummel, K. V., & Hoyer, J. (2011). Assessing psychological flexibility: What does it add above and beyond existing constructs? *Psychological Assessment, 23*(4), 970–982. https://doi.org/10.1037/a0024135

Gloster, A. T., Klotsche, J., Ciarrochi, J., Eifert, G., Sonntag, R., Wittchen, H. U., et al. (2017). Increasing valued behaviors precedes reduction in suffering: Findings from a randomized controlled trial using ACT. *Behaviour Research and Therapy, 91*, 64–71. https://doi.org/10.1016/j.brat.2017.01.013

Gloster, A. T., Meyer, A. H., & Lieb, R. (2017). Psychological flexibility as a malleable public health target: Evidence from a representative sample. *Journal of Contextual Behavioral Science, 6*(2), 166–171. https://doi.org/10.1016/j.jcbs.2017.02.003

Gloster, A. T., Richard, D. C. S., Himle, J., Koch, E., Anson, H., Lokers, L., et al. (2008). Accuracy of retrospective memory and covariation estimation in patients with obsessive-compulsive disorder. *Behaviour Research and Therapy, 46*(5), 642–655. https://doi.org/10.1016/j.brat.2008.02.010

Gloster, A. T., Walder, N., Levin, M. E., Twohig, M. P., & Karekla, M. (2020). The empirical status of acceptance and commitment therapy: A review of meta-analyses. *Journal of Contextual Behavioral Science, 18*(March), 181–192. https://doi.org/10.1016/j.jcbs.2020.09.009

Hayes, S. C. (2006). The six core processes of ACT. https://contextualscience.org/the_six_core_processes_of_act#.

Hayes, S. C., Barnes-Holmes, D., & Wilson, K. G. (2012). Contextual Behavioral Science: Creating a science more adequate to the challenge of the human condition. *Journal of Contextual Behavioral Science, 1*(1–2), 1–16. https://doi.org/10.1016/J.JCBS.2012.09.004

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy, 44*(1), 1–25. https://doi.org/10.1016/j.brat.2005.06.006

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). Acceptance and commitment therapy. In *The process and practice of mindful change* (2nd ed.). The Guilford Press. https://doi.org/10.1017/CBO9781107415324.004.

Hofer, P. D., Waadt, M., Aschwanden, R., Milidou, M., Acker, J., Meyer, A. H., et al. (2018). Self-help for stress and burnout without therapist contact: An online randomised controlled trial. *Work & Stress, 32*(2), 189–208. https://doi.org/10.1080/02678373.2017.1402389

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Karekla, M., & Michaelides, M. P. (2017). Validation and invariance testing of the Greek adaptation of the Acceptance and Action Questionnaire -II across clinical vs. nonclinical samples and sexes. *Journal of Contextual Behavioral Science, 6*(1), 119–124. https://doi.org/10.1016/j.jcbs.2016.11.006

Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review, 30*(7), 865–878. https://doi.org/10.1016/j.cpr.2010.03.001

Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior, 43*(2), 207–222. https://doi.org/10.2307/3090197

Kline, R. B. (2015). Prinicles and practice of structural equation modeling. In D. A. Kenny, & T. D. Little (Eds.), *Methodology in the social sciences* (4th ed.). The Guilford Press.

Krafft, J., Ong, C. W., Twohig, M. P., & Levin, M. E. (2019). Assessing psychological inflexibility in hoarding: The acceptance and action questionnaire for hoarding (AAQH). *Journal of Contextual Behavioral Science, 12*, 234–242. https://doi.org/10.1016/j.jcbs.2018.08.003

Lamers, S. M. A., Westerhof, G. J., Bohlmeijer, E. T., Ten Klooster, P. M., & Keyes, C. L. M. (2011). Evaluating the psychometric properties of the mental health Continuum-Short Form (MHC-SF). *Journal of Clinical Psychology, 67*(1), 99–110. https://doi.org/10.1002/jclp.20741

Leonidou, C., Panayiotou, G., Bati, A., & Karekla, M. (2019). Coping with psychosomatic symptoms: The buffering role of psychological flexibility and impact on quality of life. *Journal of Health Psychology, 24*(2), 175–187. https://doi.org/10.1177/1359105316666657

Levin, M. E., Hildebrandt, M. J., Lillis, J., & Hayes, S. C. (2012). The impact of treatment components suggested by the psychological flexibility model: A meta-analysis of laboratory-based component studies. *Behavior Therapy, 43*(4), 741–756. https://doi.org/10.1016/j.beth.2012.05.003

Menon, G. (1994). Judgments of behavioral frequencies: Memory search and retrieval strategies. In *Autobiographical memory and the validity of retrospective reports* (pp. 161–172). New York: Springer. https://doi.org/10.1007/978-1-4612-2624-6_11.

Monestès, J.-L., Karekla, M., Jacobs, N., Michaelides, M. P., Hooper, N., Kleen, M., et al. (2018). Experiential avoidance as a common psychological process in European cultures. *European Journal of Psychological Assessment, 34*(4), 247–257. https://doi.org/10.1027/1015-5759/a000327

Muthén, L. K., & Muthén, B. O. (2018). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Ong, C. W., Lee, E. B., Levin, M. E., & Twohig, M. P. (2019). A review of AAQ variants and other context-specific measures of psychological flexibility. *Journal of Contextual Behavioral Science, 12*, 329–346. https://doi.org/10.1016/j.jcbs.2019.02.007. November 2017.

Probst, T., Mühlberger, A., Kühner, J., Eifert, G. H., Pieh, C., Hackbarth, T., et al. (2020). Development and initial validation of a brief questionnaire on the patients' view of the in-session realization of the six core components of Acceptance and Commitment Therapy. *Clinical Psychology in Europe, 2*(3). https://doi.org/10.32872/cpe.v2i3.3115

Raykov, T. (2001). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology, 54*(2), 315–323. https://doi.org/10.1348/000711001159582

Rinner, M. T. B., Meyer, A. H., Mikoteit, T., Hoyer, J., Imboden, C., Hatzinger, M., et al. (2019). General or specific? The memory–experience gap for individuals diagnosed with a major depressive disorder or a social phobia diagnosis, and individuals without such diagnoses. *Memory, 27*(9), 1194–1203. https://doi.org/10.1080/09658211.2019.1640252

Rolffs, J. L., Rogge, R. D., & Wilson, K. G. (2018). Disentangling components of flexibility via the hexaflex model: Development and validation of the multidimensional psychological flexibility inventory (MPFI). *Assessment, 25*(4), 458–482. https://doi.org/10.1177/1073191116645905

Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology, 53*(1), 80–93. https://doi.org/10.1037/0022-0167.53.1.80

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology.* Jossey-Bass Inc.

Thompson, M., Bond, F. W., & Lloyd, J. (2019). Preliminary psychometric properties of the everyday psychological inflexibility checklist. *Journal of Contextual Behavioral Science, 12*, 243–252. https://doi.org/10.1016/j.jcbs.2018.08.004. August 2018.

Walentynowicz, M., Schneider, S., & Stone, A. A. (2018). The effects of time frames on self-report. *PloS One, 13*(8), 1–18. https://doi.org/10.1371/journal.pone.0201655

Wilson, D. S., Hayes, S. C., Biglan, A., & Embry, D. D. (2014). Evolving the future: Toward a science of intentional change. *Behavioral and Brain Sciences, 37*(4), 395–416. https://doi.org/10.1017/S0140525X13001593

Wolgast, M. (2014). What does the acceptance and action questionnaire (AAQ-II) really measure? *Behavior Therapy, 45*(6), 831–839. https://doi.org/10.1016/j.beth.2014.07.002

Zacharia, M., Ioannou, M., Theofanous, A., Vasiliou, V. S., & Karekla, M. (2021). Does cognitive fusion show up similarly across two behavioral health samples? Psychometric properties and invariance of the Greek–cognitive fusion questionnaire (G-CFQ). *Journal of Contextual Behavioral Science*. https://doi.org/10.1016/j.jcbs.2021.01.003. January.